Identifying Bacterial Markers of Antimicrobial-Mediated Intestinal Dysbiosis

By

Duncan James Ewin

Submitted in accordance with the requirements of

Doctor of Philosophy

The University of Leeds,

Leeds Institute of Medical Research,

Faculty of Medicine and Health

December 2024

IPR Statement

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. The jointly authored publication and the contributed work are listed below.

Primary Research papers

Chapter 3: Gut Model

The work in this chapter partly forms the jointly authored paper.

- Davis Birch WA, Moura IB, Ewin DJ, et al. MiGut: A scalable in vitro platform for simulating the human gut microbiome-Development, validation and simulation of antibiotic-induced dysbiosis. Microb Biotechnol. 2023;16(6):1312-1324. doi:10.1111/1751-7915.14259

DJE conducted the Leeds HGM experiment, isolation, and enumeration of key taxa, and performed MALDI-TOF analysis. The Leeds HGM was supplemented with additional MiGut replicates for further metagenomic and metabolomic analysis. The Leeds HGM work was designed by JF AB and DJE

WDB and IM led the MiGut experiment, with WDB running the experiment, assisting with bacterial quantification, and performing data analysis. IM led bacterial quantification and assisted with running the experiment. AB designed the study. The report was written by WDB with contributions from AB, NK and PC. The work was overseen by MW.

Summary of Contributions by Author

	Leeds HGM	MiGuT
Experimental design	DJE, JF, AB	WDB, IM, AB
Oversight and experimental lead	DJE	WDB
Culture and identification	DJE	N/A
Metagenomics (including data analysis)	DJE	DJE*
16s Analysis	WBD, I M	WDB, IM
Metabolomics (including data analysis)	DJE	DJE*
Engineering design and approach	N/A	WDB

*Work presented in this thesis but not in the publication listed above

Conference Proceedings

Work in this thesis was also disseminated through the following conference posters and presentations.

- Duncan Ewin, William Davis Birch, Ines B Moura, Mark H Wilcox, Jane Freeman, Anthony Buckley. Suppression of butyric acid levels following multiple antibiotic treatments in an *in vitro* gut model of gut dysbiosis. Presentation at Anaerobe July 2023 by DE

- Duncan Ewin, Anthony Buckley, Mark H Wilcox, Jane Freeman, James Ault. P010 I still haven't found what I'm looking for - optimising methods for metabolomic investigations of the microbiome. Poster at Anaerobe July 2023 by DE

Perspective Articles

The work and experience gathered in the course of this project have significantly contributed towards the writing of the following perspective article

- Ewin, D., Birch, W. D., & Moura, I. B. (2023). *In vitro* models to study *Clostridioides difficile* infection: current systems and future advances. Current opinion in gastroenterology, 39(1), 23–30. https://doi.org/10.1097/MOG.0000000000000893

Acknowledgements

This project would not have been possible without the incredible support I have received. Throughout this journey, I have been fortunate enough to be surrounded by an exceptional group of people. To those who have helped me along the way, I cannot hope to fully convey what your support and guidance have meant to me.

First and foremost, I would like to thank Jane, Tony, and Mark for helping me get to where I am today. Not only throughout the course of my PhD but also for guiding me to the point where I even considered taking on this challenge. Thank you for supporting me and believing in me, especially during the times when I struggled to believe in myself.

I would like to thank Applied Microbiology International, for believing in my potential and funding this work.

I would like to thank Will and Ines, for their work with the MiGut, without which this project would not have been possible. It has been a pleasure working with you both.

Charmaine and Chris, thank you for your support and friendship throughout my PhD

James Ault and the Mass Spectrometry Facility at the Astbury Centre for their assistance with my LCMS, as well as the Next Generation Sequencing Facility at the University of Leeds for their contributions

To all the members of the HCAI team with whom I've had the pleasure of working, thank you for the friendship and support you have shown me truly invaluable, Bill, Kerrie, Emma, James, Karen, Kim, Georgina and Sharie.

My friends Matt and Nick who have supported me throughout, and occasional coding help along the way.

Finally, I would like to thank my family; Mum, Dad, Laura, Roslyn, my partner Sammy, for being my rock and helping my get through the most stressful days I could not have done it without you.

Abstract

Background: The gastrointestinal tract microbiome has significant functional roles in maintaining host health, maintained through a delicate balance. Disruptions to this balance may lead to diseases such as antibiotic-associated diarrhoea and *Clostridioides difficile* infection. Taxonomic analyses can only partially explain this dysbiotic states due to the inter-individual microbiome variability. Functional and metabolic analyses may provide more a unified description of the microbiome and dysbiosis.

Aims: Identify bacterial markers of antibiotic induced dysbiosis utilising i*n-vitro* gut models.

Methods: *In-vitro* gut models (Leeds Human Gut Model, MiGut) were used to simulate antibiotic-induced dysbiosis using three clinically reflective broad spectrum antibiotic dosing regimens (Amoxicillin, Ciprofloxacin, Piperacillin/Tazobactam). Gut models were challenged with two pathogens (*Clostridioides difficile* (027), KPC-positive *Klebsiella pneumoniae*) to demonstrate post-antibiotic functional dysbiosis. Shotgun metagenomic analysis was used to taxonomically and functionally profile the microbiome. Liquid Chromatography-Mass Spectrometry was used to identify metabolic markers of dysbiosis.

Results: *Cultural analysis*: Antibiotic instillation resulted in wide scale disruptions to the microbiome coinciding with antibiotic therapy. This resolved post-antibiotic instillation with monitored taxa returning to similar abundance and composition pre-antibiotic instillation. However, functional dysbiosis was demonstrated by the colonisation and proliferation of *Clostridioides difficile* and *Klebsiella pneumoniae* within the models.

Metagenomic analysis: Reductions in species diversity and a loss of key genera was associated with short chain fatty acid (SCFA) production. KEGG pathway analysis identified changes in abundance to pathways associated with SCFA metabolism for acetate, propanoate and butyrate. SCFA producers *Faecalibacterium and Coprococcus were reduced in* abundance following antibiotic instillation.

Metabolomic analysis: Reduced SCFA levels within the metabolome coincided with antibiotic instillation that persisted beyond recovery of key microbial

genera. Putative Metabolic Markers; Butyrate, Valerate, Lactate, 2-Hydroxyglutarate, 2-Oxoglutarate. Acetate and Propanoate.

Conclusions: A putative panel of nine biomarkers of dysbiosis, consisting of 2 taxonomic species and 7 metabolites, were identified. *In vitro* models are effective tools to screen for marker of antibiotic induced dysbiosis.

Table of Contents

IPR Statement ..... **i**
Acknowledgements ..... **iii**
Abstract ..... **iv**
Table of Contents ..... **vi**
Table of Figures ..... **xii**
Table of Tables ..... **xvi**
List of Abbreviations ..... **xviii**
Chapter 1 Introduction **1**
1.1 Composition 4
1.2 Metabolic Function 6
1.2.1 Non-digestible polysaccharides 6
1.2.2 Lipids 8
1.2.3 Vitamins 9
1.2.4 Bile acids 9
1.3 Colonisation Resistance 11
1.3.1 Direct colonisation Resistance 11
1.3.1.2 Antimicrobial Inhibition 12
1.3.1.3 Inhibitory Metabolites 13
1.3.1.4 Type VI Secretion Systems 13
1.3.2 Indirect Colonisation Resistance 14
1.3.2.1 Immune System Conditioning 14
1.3.2.2 Microbial Modification of Host-Derived Compounds 15
1.4 Dysbiosis 16
1.4.1 Age-Related Microbial Decline 16
1.4.2 Antibiotic-induced dysbiosis 18
1.4.2.1 Factors Influencing Antibiotics impact on the Gut Microbiome 19
1.4.2.2 . Effect of prophylactic Antibiotics on the gut microbiome 20
1.4.3 Clostridioides difficile infection 21
1.4.3.1 CDI Recurrence 22
1.4.4 Colonisation by Multidrug-resistant Organisms 22
1.4.5 Recovery from Dysbiosis 24
1.4.5.1 Microbiome therapeutics 25
1.5 Identifying dysbiosis 26

1.5.1 Biomarkers of Dysbiosis 26
1.6 Approaches to identifying markers of dysbiosis 29
1.6.1 Meta-omics 29
1.6.2 In-vitro Model systems 33
1.6.2.1 Batch Fermentation models 33
1.6.2.2 Continuous single-stage models 33
1.6.2.3 Continuous multi-stage models 34
1.7 Thesis outline 36
1.8 Thesis Aims 37
Chapter 2 Metabolomic Methods Development 38
2.1 Introduction 38
2.2 Chapter Aims 40
2.3 Experimental outline 41
2.3.1 High-Performance Liquid Chromatography, Sample Fractionation 41
2.3.2 Full Scan and Deep-Scan Experiments 42
2.3.2.1 Full Scan 44
2.3.2.2 Deep Scan 45
2.4 Sample Processing 47
2.4.1 Ethical approval 48
2.4.2 Solvent-Based Metabolite Extraction 49
2.4.3 Comparative Analysis of Faecal Sample Processing Methods: Bead-Beating vs. Manual Extraction 50
2.4.4 Comparative Evaluation of Extraction Solvents for Faecal and In-vitro model Samples 53
2.5 LCMS Parameters 55
2.5.1 High-Performance Liquid Chromatography 55
2.5.1.1 C18 Chemsitry 57
2.5.2 Mass Spectrometry 61
2.6 Data Processing and Analysis 62
2.6.1 Input 62
2.6.2 Spectrum Processing 63
2.6.3 Compound Detection and QC 63
2.6.4 Compound Identification 64
2.7 LCMS Deep Scan Experiments: Determining Optimal Coverage 65
2.8 Local Library Development – Processing and Annotation 68
2.9 Metabolite Identification Limitations 70

2.9.1 Impact of Polarity on Compound Recovery....70
2.9.2 Compound Classes Associate by Retention Time....72
2.9.3 Compound Annotation Confidence....73
2.10 Discussion....76
2.11 Summary....82
Chapter 3 Gut Model....83
3.1 Introduction....83
3.2 Experimental Design....84
3.3 Methods....85
3.3.1 *In-vitro* models....85
3.3.1.1 Ethical approval....85
3.3.1.2 Leeds Human Gut Model....86
3.3.1.3 MiGut....91
3.3.2 Bacterial Culture & Enumeration....92
3.3.2.1 MALDI-TOF Identification of Cultured Strains....94
3.3.3 Antibiotic Instillation....95
3.3.3.1 Antibiotic Dosage & Schedule....95
3.3.4 Pathogen Challenge....96
3.3.5 Cell Cytotoxicity Neutralisation Assay....96
3.3.5.1 Vero Cell Preparation....97
3.3.5.2 CCNA Assay Protocol....98
3.3.6 Metagenomic and Metabolomic Sample Collection....101
3.4 Results....101
3.4.1 Impacts of Antimicrobials on the Gut Microbiome....101
3.4.1.1 Facultative Anaerobes....101
3.4.1.2 . Obligate Anaerobes....105
3.4.1.3 Predominant Species Diversity Estimates (MALDI-TOF)107
3.4.2 Antimicrobials and Colonisation Resistance....110
3.4.2.1 *Clostridioides difficile*....110
3.4.2.2 Klebsiella pneumoniae....111
3.1 Discussion....112
3.1.1 Establishing Baseline Microbiome....113
3.1.2 Impacts of Antimicrobials on the Gut Microbiome....114
3.1.3 Recovery....116
3.1.4 Pathogen Expansion....118
3.2 Summary....119

Chapter 4 Metagenomic investigation of the *In-vitro* Intestinal microbiome. 120
4.1 Introduction 120
4.2 Chapter Aims 121
4.3 Experimental Outline 122
4.4 Methods 124
4.4.1 DNA Extraction 124
4.4.2 Bioinformatics software 124
4.4.3 Sequencing 126
4.4.3.1 Metagenomic Controls 126
4.4.3.2 Library preparation 126
4.4.4 Preparing Paired End Reads for MEGAN 127
4.4.4.1 DIAMOND alignment comparison 129
4.4.5 Microbiome Data Processing and Statistical Analysis 130
4.4.5.1 Distance Matrix Computation Methods 130
4.4.5.2 Genus Ordination 130
4.4.5.3 Differential abundance analysis 130
4.5 Results 131
4.5.1 Taxonomic analysis 131
4.5.1.1 Microbiome Compositional Shifts as a Result of Antimicrobial Instillation 131
4.5.1.2 Identifying Minority Genera 135
4.5.1.3 Antibiotic Impact on Microbiome Diversity 137
4.5.1.4 Differential Abundance Analysis 142
4.5.1.5 Species Diversity Across Experimental Phases. 144
4.5.1.6 Impact of Antibiotic Instillation on Species Diversity in Highly Diverse Genera 146
4.5.2 Metagenomic Functional Profiling of the intestinal microbiome 150
4.5.2.1 KEGG Pathway 151
4.6 Discussion 160
4.6.1 Steady state microbiome 160
4.6.2 Impacts of Antimicrobials on the Intestinal Microbiome. 161
4.6.2.1 Compositional Shifts 161
4.6.2.2 Diversity 162
4.6.2.3 Functional Analysis of the Metagenome 167
4.6.3 Summary 173

Chapter 5 Metabolomics 174
5.1 Background 174
5.2 Chapter Aims 175
5.3 Experimental Outline 175
5.4 Methods 178
5.4.1 Sample Collection and Processing 178
5.4.2 LCMS Analysis 178
5.4.2.1 High-Performance Liquid Chromatography 178
5.4.2.2 Mass Spectrometry 179
5.4.2.3 Data Processing 179
5.4.2.4 Statistics and Analysis 180
5.5 Results 181
5.5.1 Overview of Compound Features Detected by LCMS 181
5.5.2 Compositional Alterations to the *In-Vitro* Metabolome 182
5.5.2.1 Compositional Alterations Detected by Negative LCMS 183
5.5.2.2 Compositional Alterations Detected by Positive LCMS 184
5.5.2.3 Comparing the Resolution of Metabolomic Differences: PERMANOVA Analysis of Positive and Negative Ionisation Modes 185
5.5.3 Antibiotic-Induced Alterations in Short-Chain Fatty Acid Pathways 185
5.5.3.1 Acetate & Propanoate 186
5.5.3.2 Lactate 186
5.5.3.3 Butyrate 190
5.5.3.4 Associated Changes in Butyrate Metabolism 193
5.5.3.5 Valerate 203
5.6 Discussion 207
5.6.1 Variations to the Microbiome 209
5.6.1.1 Negative NMDS analysis 209
5.6.1.2 Positive NMDS analysis 212
5.6.2 Putative biomarkers of Antibiotic Induce dysbiosis 213
5.6.2.1 Butyrate 213
5.6.3 Summary 217
Chapter 6 Conclusions 219
6.1 Concluding remarks 219
6.1.1 Pipeline Development & Validation of LCMS 220

6.1.2 Utility of Leeds HGM and MiGut Platform for In Vitro Biomarker Screening 221
6.1.3 Proposed Bacterial Markers. 221
6.1.4 Limitations 223
6.1.5 Future Work 225
Bibliography 227
Appendix A Supplementary Material: Metabolomic Methods Development 253
A.1 Equipment 253
A.2 Mass Spectrometry Metabolite Library of Standards 253
Appendix B Supplementary Material: In-Vitro Gut Models 254
B.1 Bacterial Culture Media 254
B.1.1 Supplier Poured Media 254
B.1.2 In-house Media Used for Bacterial Enumeration 254
B.1.2.1 Formula for In-house Media 254
B.1.2.2 Agar Base Powders 256
B.1.2.3 Media Supplements & Antibiotics 257
B.1.3 Antimicrobials 258
B.2 Strains & Cell Lines 258
B.2.1 Bacterial Strains 258
B.2.2 Eukaryotic Cell lines 259
B.3 Equipment 259
Appendix C Supplementary Material: Metagenomic Data and Analysis for In-vitro Intestinal Microbiome Investigation 261
3.1 NGS Data Analysis and QC 261
3.1.1 Raw Sequence General Statistics 261
Raw Sequence Mean Quality Score 262
3.1.2 Trimmomatic 263
3.1.3 Paired-End Read Merger (PEAR) 265
Appendix D Supplementary Material: Metabolomics 266
4.1 Log-Transformed Intensity Distributions in Negative and Positive Modes (Steady State) 266

Table of Figures

Figure 1-1 Meta-omics Approaches Along the Central Dogma 30

Figure 1-2 Thesis Structure: Method development, In-vitro modelling and multi-omic approaches for exploring the intestinal microbiome 36

Figure 2-1 Overview of LCMS Metabolomic workflow with Validation steps for analysis 39

Figure 2-2 LCMS Experimental Outline for Full Scan and Deep Scan Experiments. 42

Figure 2-3 Examples of some common positive and negative molecule (M) adducts 43

Figure 2-4 MS processes during Full Scan LCMS experiments 44

Figure 2-5 MS processes during Targeted Deep Scan LCMS experiments 46

Figure 2-6 Outline of Key Steps in the Extraction of Metabolites from Faecal and *In Vitro* Samples 47

Figure 2-7 Overview for Metabolite Extraction for Faecal and *In Vitro* Gut Model Samples 49

Figure 2-8 Comparison of Coefficient Variability in Bead Beaten and manually processed Samples: Positive vs. Negative Ionization ...53

Figure 2-9 PCoA Assessment of Solvent Effects on Metabolite Profiles 54

Figure 2-10 Reverse Phase Chromatography, an overview. 56

Figure 2-11 C18 Column Chain Chemistry 57

Figure 2-12 Organic Phase (Solvent B) proportion in HPLC Chromatography.....59

Figure 2-13 Histogram of Compound Retention Times (minutes) of Pooled Faecal Samples.....60

Figure 2-14 Compound Discoverer Untargeted Metabolomics data processing workflow. 62

Figure 2-15 MS2 Spectra Distribution: Preferred Ion Coverage vs DDA Iteration, by Ionization Mode 66

Figure 2-16 Spectral Library Annotations vs. DDA Iteration: Negative Ionization.....67

Figure 2-17 Distribution of Compound Classes by Recovery Status in GUTMLS and BACSMLS Libraries 69

Figure 2-18 GUTMLS and BACSMLS libraries, XLogP vs Molecular Weight.....71

Figure 2-19 Prominent Compound Classes Analysed by XLogP and Retention Time.....72

Figure 2-20 Hierarchy of compound annotation in metabolomic analysis.....73

Figure 2-21 Compound annotations by highest level identification.....74

Figure 2-22 Final LCMS Metabolomic Workflow for Untargeted Analysis 76

Figure 3-1 Dysbiosis Model Timeline. 84

Figure 3-2 The Leeds human gut model (HGM) and configuration overview 86

Figure 3-3 Details of the Migut platform and model.....91

Figure 3-4 CCNA Vero Cell Seeding Layout 98

Figure 3-5 CCNA Assay variations for Detection and Quantification of *C. difficile* Toxin.....99

Figure 3-6 Vero cells displaying Positive and Negative reactions 100

Figure 3-7 Facultative Anaerobes: Leeds HGM Vessel 1 (Proximal)...102

Figure 3-8 Facultative Anaerobes: Leeds HGM Vessel 3 (Distal). 103

Figure 3-9 Obligate Anaerobes: Leeds HGM Vessel 1 (Proximal). 105

Figure 3.3-10 Obligate Anaerobes: Leeds HGM Vessel 3 (Distal).....106

Figure 3-11 Leeds HGM - Clostridioides difficile.....110

Figure 3-12 Leeds HGM - Detection of KPC positive *K. pneumoniae* (yellow) and *E. coli* (blue).....111

Figure 4-1 NMDS Ordination of 16s sequencing from Leeds HGM and MiGut platforms. 123

Figure.4-2 Pre-processing pipelines considered for Metagenomic data preparation. 128

Figure 4-3 Genus Level Compositional Changes in Vessel 1 132

Figure 4-4 Genus Level Compositional Changes in Distal Colon.....133

Figure 4-5 Genus Level Composition alterations, Vessel 1 vs Vessel 3 134

Figure 4-6 Distribution of assigned reads across all samples at the genus-level 136

Figure 4-7 Vessel 1 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus.....138

Figure 4-8 Vessel 3 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus.....140

Figure 4-9 Comparison of Unique and Shared Species Counts Across Proximal (vessel 1) and Distal (vessel 3) Regions at Steady State, Post-Antibiotic Recovery, and Post-Pathogen Challenge.....145

Figure 4-10 Heat map of species counts by genus across Leeds HGM and MiGut Platforms.....149

Figure 4-11 Distribution of Reads by Pathway Categories at Steady-state for both Leeds HGM and MiGut platforms 150

Figure 4-12 Total Reads Assigned to Subpathways within KEGG Pathways 151

Figure 4-13 Outline of major Short-chain Fatty acid production 152

Figure 4-14 Differentially Abundant Enzymes and Their Roles in Butyrate Metabolism 155

Figure 4-15 Differentially Abundant Genes and Their Roles in Propanoate Metabolism. 157

Figure 4-16 Differentially Abundant Genes and Their Roles in Acetate/Pyruvate metabolism.....159

Figure 4-17 Total Assigned Reads at Steady State for Leeds HGM and MiGut 160

Figure 4-18 Glutarate Pathway for Butyrate Synthesis 168

Figure 5-1 Experimental Outline: Metabolomic investigation into antibiotic-induced dysbiosis 176

Figure 5-2 Experimental compound annotations by highest-level identification 182

Figure 5-3 NMDS Analysis of Metabolome Changes Detected in Negative Ionisation mode 183

Figure 5-4 NMDS Analysis of Metabolome Changes Detected in Positive Ionisation mode 184

Figure 5-5 Extracted Ion Chromatogram (XIC) for Compound Identified as Lactate (m/z 88.05246) In Negative Ionisation 187

Figure 5-6 MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Lactate. Showing matched peaks and their associated fragments.....187

Figure 5-7 Lactate Abundance Across Experimental Stage 188

Figure 5-8 Lactate Abundance Across Experimental Stage and Model Replicates.....189

Figure 5-9 Extracted Ion Chromatogram (XIC) for Compound Identified as Butyrate (m/z 88.05246) In Negative Ionisation 190

Figure 5-10 Butyrate Abundance Across Experimental Stage 191

Figure 5-11 Butyrate Abundance Across Experimental Stage and Model Replicates.....192

Figure 5-12 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Hydroxyglutarate (m/z 148.03714) In Negative Ionisation....193

Figure 5-13 MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Hydroxyglutarate 194

Figure 5-14 2-Hydroxyglutarate Abundance Across Experimental Stage 195

Figure 5-15 2-Hydroxyglutarate Abundance Across Experimental Stage and Model Replicates 196

Figure 5-16 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Oxoglutarate (m/z 145.0143) In Negative Ionisation 198

Figure 5-17 2-Oxoglutarate MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Oxoglutarate. 199

Figure 5-18 2-Oxoglutarate Abundance Across Experimental Stage ..200

Figure 5-19 2-Oxoglutarate Abundance Across Experimental Stage and Model Replicates 201

Figure 5-20 Extracted Ion Chromatogram (XIC) for Compound Identified as valerate (m/z 249.13569) In Negative Ionisation 203

Figure 5-21 Valerate MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Valerate. 204

Figure 5-22 Valerate Abundance Across Experimental Stage 205

Figure 5-23 Valerate Abundance Across Experimental Stage and Model Replicates. 206

Figure C-1 Raw Sample Mean Quality Scores. 262

Figure C-2 Proportion of Raw Sequences with Illumina Universal Adapter Contamination prior to Trimming 263

Figure D-1 Log-Transformed Intensity Distributions for Negative and Positive ionisation modes (Steady State).....266

Table of Tables

Table 1-1 Bacterial Taxa Frequently Observed in the Gut Microbiome...4

Table 3-1 Leeds HGM vessels and their respective conditions. 88

Table 3-2 Leeds HGM Growth Media 89

Table 3-3 Key bacterial Strains and Culture Media 93

Table 3-4 Introduced Pathogens and Media for Culture 93

Table 3-5 Quality Scores for MALDI-TOF Identification.....94

Table 3-6 Antibiotics, Dosing Frequency and Leeds HGM Vessel 1 target concentration 95

Table 3-7 Vero Cell Growth Media 97

Table 3-8 Predominant Bacterial species recovered by bacterial culture by agar media type Leeds HGM Vessel 1 + Vessel 3 (MALDI-TOF) 109

Table 4-1 Metagenomic Sample Timepoints for Leeds HGM and MiGUT 122

Table 4-2 Queries Aligned in DIAMOND with Processing Pipelines....129

Table 4-3 Vessel 1 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery 142

Table 4-4 Vessel 1 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation 143

Table 4-5 Vessel 3 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery 143

Table 4-6 Vessel 3 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation 143

Table 4-7 Comparison of Species counts in Leeds HGM and MiGut platforms 144

Table 4-8 Differentially Abundant Genes Within Butyrate Metabolism153

Table 4-9 Differentially Abundant Genes Within Propanoate Metabolism 156

Table 4-10 Differentially Abundant Genes Within Acetate/Pyruvate Metabolism 158

Table 5-1 Key Metabolites Presented and Discussed in This Chapter, Along with Their Protonated and Deprotonated Form 186

Table 5-2 Differential Abundance of Lactate, P. Abx and P. Path vs SS 189

Table 5-3 Differential Abundance of Butyrate, P. Abx and P. Path vs SS 192

Table 5-4 Differential Abundance of 2-Hydroxyglutarate, P. Abx and P. Path vs SS 197

Table 5-5 Differential Abundance of 2-Oxoglutarate, P-Abx and P-Path vs SS 202

Table 5-6 Differential Abundance of Valerate, P. Abx and P. Path vs SS 206

Table 6-1 Putative Biomarkers Identified in Antibiotic-Induced Gut Dysbiosis 222

Table A-1 Equipment used for LCMS Workflow 253

Table A-2 LCMS Consumables 253

Table A-3 Fridges and Freezers 253

Table A-4 MSMLS Used in Validation and Library Generation 253

Table B-1 Supplier Poured Media Used for Bacterial Enumeration 254

Table B-2 Braziers CCEY + Lysozyme agar (CCEYL) 254

Table B-3 Braziers CCEY + Lysozyme, moxifloxacin, amphotericin B and colistin agar (MXF) 255

Table B-4 Bacteroides Bile Aesculin agar (BBE) 255

Table B-5 Nutrient agar: 255

Table B-6 Kanamycin Aesculin Azide, Nalidixic Acid, Aztreonam agar (KAA) 255

Table B-7 Beerens agar 256

Table B-8 LAMVAB 256

Table B-9 Agar Powders used in In-House Media 256

Table B-10 Supplements Used in the Creation of In-house Media[1] 257

Table B-11 Antimicrobials Used for In-vitro Study 258

Table B-12 Reference Bacterial Strains with Catalogue Numbers 258

Table B-13 Eukaryotic Cell Lines Used for Cell Cytotoxicity Neutralisation Assays. 259

Table B-14 Anaerobic Workstation 259

Table B-15 Aerobic Culture 259

Table B-16 CO_2 Culture 259

Table B-17 Balances 259

Table B-18 Autoclave & Sterilisation Cycle 259

Table B-19 In-Vitro Models (Leeds-HGM & MiGut) 260

Table B-20 Other Equipment 260

Table C-1 General Statistics for NGS data for miGUT and Leeds 261

Table C-2 Trimmomatic Summary Data for Pair-End Trimming 264

Table C-3 Summary of Assembled Reads in Millions (M Seq) and Percentage of Total Paired Reads 265

List of Abbreviations

ADD	Antibiotic-associated diarrhoea
AMC	Co-amoxyclav
AMX	Amoxicillin
AWaRe	Access, Watch, and Reserve
BACSMLS	Bile Acid Carnitine Sterol Library
BBE	Bacteroides Bile Esculin
BEER	Beerens Agar
BSH	Bile Salt Hydrolases
C18	Carbon-18
CARBA	CHROMID CARBA SMART Agar
CAWG	Chemical Analysis Working Group
CCNA	Cell Cytotoxicity Neutralization Assay
CDI	Clostridioides difficile infection
cdtA	*Clostridioides difficile binary toxin*
CFU	Colony-forming unit
ChEBI	Chemical Entities of Biological Interest
CI	Confidence interval
CID	Collision-induced fragmentation
CIP	Ciprofloxacin
CMS	Continuous multi-stage
CPE	Carbapenemase-producing Enterobacteriaceae
CRC	Colorectal cancer
CRO	Ceftriaxone
CV	Coefficient of variation
DDA	Data-dependent acquisition

DDD	Defined daily doses
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray Ionisation
EtOH	Ethanol
FAA	Fastidious Anaerobe agar
FISh	Fragment Ion Search
FMT	Faecal microbiota transplantation
GC	Gas Chromatography
GCMS	Gas Chromatography Mass pectrometry
GF	Germ-free
GMHI	Gut Microbiome Health Index
GUTMLS	Microbiome Metabolite Library of Standards
HBSS	Hanks Balanced Salt solution
HCAI	Healthcare Assocaited Infection
HGM	Human Gut Model
HMDB	Human Metabolome Database
HPLC	High-Performance Liquid Chromatography
HSCT	Haematopoietic stem cell transplant
IDB	Inflammatory bowel disease
KAA	Kanamycin Aesculin Azide Agar
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KPC	Klebsiella pneumoniae carbapenemase
LC	Liquid Chromatography
LCA	Lithocholic acid
LCMS	Liquid Chromatography Mass Spectrometry
LF	Lactose-fermenting

LOD	Limit of detection
m/z	Mass-to-charge ratio
MAC	MacConkey
MALDI-TOF	Matrix-assisted laser desorption/ionisation – Time of Flight
MDI	Microbial Dysbiosis Index
MDRO	Multidrug-Resistrant Organisms
MeCN	Acetonitrile
MeOH	Methanol
MS	Mass Spectrometry
MS^2	Tandem mass spectrometry
MSI	Metabolomics Standards Initiative
MXF	Brazier's CCEY agar + Moxifloxacin
NGS	Next-generation sequencing
NMDS	Non-metric Multidimensional Scaling
NUT	Nutrient Agar
PBS	Phosphate-buffered saline
PCOA	Principal Coordinates Analysis
PCR	Polymerase Chain Reaction
PE	Paired End
PERMANOVA	Permutational Multivariate Analysis of Variance
PES	Polyether sulfone
PUFA	Polyunsaturated fatty acids
QC	Quality control
RT	Retention time
rtPCR	Real Time PCR

RU	Relative units
SCFA	Short-chain fatty acids
SE	Single End
SHIME	Simulator of the Human Intestinal Microbial Ecosystem
SPF	Specific pathogen-free
T6SS	Type VI secretion systems
t*cdB*	*Clostridioides difficile* toxin B
TVC	Total viable counts
TZP	Piperacillin-Tazobactam
UDCA	Ursodeoxycholic acid
VDR	Vitamin D receptor
XIC	Extracted ion chromatograms

Chapter 1
Introduction

The human body is home to a thriving abundance of life; bacteria, archaea, viruses and eukaryotes all reside within its diverse range of ecosystems. This is known as the microbiome. As the human body has evolved, an intimate symbiosis with the microbiome has developed. It is now becoming evident that this relationship exists at a level where the human and microbial components can no longer thrive in isolation, leading to the emergence of a holobiont. There are distinct ecological niches that exist across the human body and reflect the abundance of environmental ranges found within the surrounding world. The largest and most complex of these exists in the nutrient-rich and densely populated intestinal tract. Within the intestinal tract, the microbiota strives to carve out an ecological niche that competes with endogenous and invading microorganisms whilst contending with the immunological defences of the host immune system. When successful, there exists a finely tuned balance between the host and microbiome known as homeostasis. During homeostasis, the microbiome and host exist in a mutually beneficial relationship, with each contributing to the well-being of the other.

The arrival of next-generation sequencing (NGS) brought about a new era in metagenomic studies, allowing for investigations of the microbiome beyond the capabilities of traditional culture-based methods. The Human Microbiome Project explored the microbiota of 252 healthy adults between the ages of 18 and 40, screening a total of 4,788 species (Human Microbiomes Project Consortium, 2012) and the American Gut Project (McDonald, D. et al., 2018) profiling 10,000 individuals from across the world were groundbreaking studies which demonstrated the potential for the future. These advances in sequencing technology, able to generate vast amounts of data, brought about a new understanding of the diversity and complexity of the microbiome, with the variations which exist between individuals and across the globe.

Collectively, the microbiome possesses a genome that far surpasses our own. It is estimated that the genes contained by the intestinal microbiome alone is approximately 150 times larger than that of the human genome (Qin et al.,

2010). Many of these genes work in complement with our own providing important metabolic functionality.

These large metagenomic studies have revealed that even between healthy individuals, there is a significant degree of microbiome compositional variability. In addition to individual variations, the microbiome has undergone significant shifts as human behaviour and lifestyles have changed, evidenced by shifts in composition as populations move from non-industrial to industrial societies (Keohane et al., 2020).

The compositional changes occurring between non-industrial and industrial societies highlight the impacts that human behaviour, the environment, and medical interventions can have in shaping the microbiome. Potentially, the greatest impact documented as shaping the microbiome comes through the use of antibiotics. Antibiotics have transformed the landscape of disease treatment. Despite their benefits, the use of antibiotics is not without drawbacks, as they can significantly impact the microbiome beyond their target organisms, leading to disruptions by reducing diversity and abundance.

As researchers continue to explore the human microbiome, the complex interactions which occur between host and resident microorganisms are being revealed, and it is becoming increasingly clear that if research is to continue exploring human health, it is imperative to consider the role of the microbiome.

As the microbiome has been increasingly studied, significant compositional differences between the microbiomes of healthy individuals have been observed. These differences highlight that there are multiple taxonomic pathways which lead to a healthy microbiome, with functional redundancy between bacteria allowing for different species to fulfil similar ecological roles. This redundancy, potentially being key in maintaining host health as the microbiome shifts over time in response to physiological and environmental changes.

These observations have prompted researchers to increasingly explore the functionality of the microbiome. The protection of the host from pathogens afforded by the microbiota through colonisation resistance has been extensively documented and reviewed (Caballero-Flores et al., 2023) (Hou et al., 2022) and serves to prevent the establishment and proliferation of potentially harmful

exogenous bacteria. The intestinal microbiome existing alongside the mucosal membrane of the gut has been shown to have an important role in conditioning the host immune system (Yamamoto et al., 2012; Graham and Xavier, 2023), as well as being responsible for the production of key metabolites within the gut (Miquel et al., 2013; Lee, J. et al., 2020).

Whilst the microbiome fulfils these important roles for the host, its proper function is reliant on the homeostatic balance of the microbiome between different microbial components and with the host immune system. When this balance is disrupted, either from external influencing factors such as antibiotics (Nel Van Zyl et al., 2022) or as the consequence of natural phenomena, such as the reduction diversity associated with ageing (Odamaki et al., 2016) it can have profound health consequences for the host.

Disruption of the microbiome and the homeostatic balance integral to optimal function of the microbiome is known as dysbiosis. Whilst predisposition to many of these conditions is associated with complex aetiologies which cannot be attributed solely to the microbiome, the malleable nature of the microbiome offers a route of functional amelioration through microbial contributions (Vrieze et al., 2012; Kelly, C.R. et al., 2016) when the genetic factors in play remain static

The purpose of this thesis is to explore intestinal dysbiosis through the application of *in vitro* models of the human intestinal tract. By employing a range of techniques, including bacterial isolation and culture, but with a significant emphasis on metabolomics and metagenomics, the primary aim is to characterise aspects of dysbiosis functionally. As a result of an increased understanding of dysbiosis, it may also be possible to derisk antibiotic selection by exploring function paucities associated with the side effects of antimicrobials and identify patients who may be at an increased of developing dysbiosis and the diseases associated with it, such as *Clostridioides difficile* infection

The Intestinal Microbiome and Dysbiosis

1.1 Composition

The microbiome does not occur as a fixed entity; microbial populations are dynamic, undergoing subtle changes over time. The microbiome of a healthy individual tends to exist within the confines of a specific range, remaining relatively stable over time. The degree of variation within the microbiome can differ significantly between body sites, and the gastrointestinal tract remains the most stable over time. (Human Microbiomes Project Consortium, 2012) Whilst stable, the gut microbiome varies in composition and density across different regions of the gastrointestinal tract. Changes in pH, nutrient availability, water absorption, bacterial growth, and the shedding of bacteria from sessile populations associated with the mucosal layers of the intestinal wall all contribute to the increasing density of bacteria along the intestinal tract. The concentration of bacteria is estimated to increase by 1,000-fold from the small intestine and colon to faeces (Sender et al., 2016). The two dominant phyla within the intestinal microbiota are the *Firmicutes* and *Bacteroidetes* , accounting for approximately 90% of bacteria within the large intestine. However, it is thought that the average intestinal microbiome is home to around 500 bacterial species (Sender et al., 2016)

Table 1-1 Bacterial Taxa Frequently Observed in the Gut Microbiome

Region	Bacterial Abundance[3]	Prominent bacterial Taxa
Upper Small Intestine[1] (Duodenum & Jejunum)	10^7	Streptococcus, Lactobacillus
Lower small Intestine[1] (Ileum	10^{11}	Streptococcus, Lactobacillus
Colon[2]	10^{14}	**Firmicutes** (Faecalibacterium, Clostridium, Ruminococcus), **Bacteroidetes** (Bacteroides, Alistipes, Prevotella)

[1](Ahmed et al., 2007), [1](Wang, X. et al., 2003), [1](Hayashi, H. et al., 2005), [2](Liu, C. et al., 2021), [3](Sender et al., 2016)

The degree of permanence exhibited by the microbiome over weeks, months, and years has been the topic of multiple investigations. Faith et al. (Faith et al., 2013) displayed that at a compositional level, the microbiome remains relatively stable over the course of time, with the majority of strains being predicted to remain as residents of an individual's intestinal microbiome throughout their life with higher levels of alpha diversity being associated with greater microbial stability. However greater variation exists at a metagenomic and metatranscriptomic level, possibly driven by strain replacement within the microbiota (Mehta et al., 2018; Chen, L. et al., 2021).

Faecal samples are the predominant sample used for the study of the intestinal microbiome, particularly with humans, meaning current data remains largely representative of the colon, with the small intestine being poorly characterised by comparison. Studies with ileostomy patients may provide some insights into the composition and function of the microbiome of the ileum, with Leimena et al. (2013) suggesting a dominance of *Streptococcus* species.

Ahmed et al. (2007) identified significant populations of *Bacteroides, Lactobacillus* and *Clostridium* species associated with the mucosa within the terminal ileum of patients undergoing emergency resection of the large bowel. Wang, X. et al. (2003) supported findings of Clostridium and Bacteroides from a colonic lavage patient

Hayashi, H. et al. (2005) detected populations of *Streptococci, Lactobacilli, Enterococcus and Escherichia* and a notable absence of *Clostridium* species within the ileal microbiota from autopsied individuals.

However, the underlying health conditions which may have contributed to patients being required to undergo ileostomy surgery or mortality may have influenced the composition of the intestinal microbiome and, as such, may not be representative of the healthy microbiome and explain some of the contrasting findings in previous studies.

Despite studies such as these, due to a current lack of non-invasive methods by which to obtain samples at a scale similar to faeces, the microbiome of the small intestine is likely to remain relatively elusive by comparison to the colon.

The colon is host to an incredible amount of bacterial density (Sender et al., 2016) and remains a focal point for research into the human microbiome. It is

largely an anaerobic environment with some diffusion of oxygen from surrounding tissue (Albenberg et al., 2014), and because of this, the microbiome is predominately occupied by anaerobes (Macfarlane et al., 1998; Arumugam et al., 2011; McDonald, D. et al., 2018).

1.2 Metabolic Function

The microbiome performs essential metabolic functions, significantly contributing to host physiology and health. These functions complement host metabolic pathways, providing additional capabilities that influence nutrient metabolism, immune modulation and gut homeostasis. While the taxonomic composition of the microbiome varies between individuals, the metabolic capacity it possesses remains stable between individuals (Human Microbiomes Project Consortium, 2012). This stable, functional component likely accounts for the repertoire of microbial metabolic pathways upon which humans rely.

1.2.1 Non-digestible polysaccharides

Research has revealed a wide range of dietary fibres which can be utilised by the microbiota, capable of supporting diverse microbial communities. By breaking down complex carbohydrates, the microbiota provides the host with additional resources to support physiological processes.

Plant and dietary fibres which are not processed by humans are able to be fermented by the gut microbiota. For example, oligofructose, an inulin-type fructan, being composed of fructose molecules, is a prebiotic fibre found in plants and capable of being processed but the gut microbiota. Falony, G. et al. (2006) demonstrated the complex interactions potentially occurring within fermentative pathways in the gut using oligofructose as a substrate. Some *Bifidobacteria* strains have the capacity to degrade inulin-type fructans, including oligofructose to fructose (Falony, Gwen et al., 2009) through a fermentative pathway producing acetate. When grown in co-culture with *Bifidobacteria longum, Anaerostripes caccae* was able to grow utilising fructose produced by *B.longum.* Similarly, *Roseburia intestinalis* was able to degrade oligofructose only in the presence of acetate, either supplemented or produced through fermentation. While neither *A.caccae* nor *R.intestinalis a*re capable of growing on oligofructose alone, the cross-feeding of metabolites by other microorganisms within the microbiome can initiate additional metabolic

pathways. Importantly, both *A.caccae and R.intestinalis* are butyrate-producing bacteria, a metabolite with an important role in host physiology as an energy source for colonocytes. Studies like this, highlighting the complex interactions between microbiota constituents, reveal how disturbances to the microbiome may have profound downstream metabolic consequences beyond those encoded by the populations directly impacted.

The importance of these bacterial pathways is highlighted by the utilisation of their products by the host. Short-chain fatty acids (SCFA) are produced through these fermentation pathways and are key metabolites contributing to host physiology. The significance of SCFA and the bacteria that produce them is well-documented (Sokol et al., 2008; Miquel et al., 2013; Canfora et al., 2015; Ríos-Covián et al., 2016; LeBlanc et al., 2017; Tsukuda et al., 2021; Visekruna and Luu, 2021; Frolova et al., 2022), underscoring their critical role in maintaining gut homeostasis. The predominant SCFAs within the gut are acetate, propionate and butyrate. Butyrate has been identified as having a pivotal role in gut homeostasis as an important energy source for colonocytes, contributing to approximately 70% of the cell's energy (Roediger, 1980)

Colonocytes of germ-free mice were shown to display a 56% reduction in the level of ATP and a 16-fold reduction in $NADH/NAD^+$ ratio when compared to conventionally raised mice. ATP production was able to be restored by colonising GF mice with either microbial populations derived from conventionally raised mice or by colonisation with known butyrate-producing strains of bacteria. (Donohoe et al., 2011)

Administration of ceftriaxone in rats has been shown to cause substantial changes to the microbiota, which persisted to 56 days post-antibiotic withdrawal. These changes in composition were associated with decreased concentrations of SCFAs and a shift in the colonic homeostasis, leading to increased epithelial permeability in the colon and increased bacterial translocation to the blood (Jernberg et al., 2007).

1.2.2 Lipids

Bacteria within the gastrointestinal tract have the capacity to metabolise dietary fat from the host diet, contributing to host health

Dietary changes over the last 30 years have resulted in an increase in the ratio of omega-6 fatty acid / omega-3 fatty acid from a historical 1:1 ratio observed in hunter-gatherer societies to a 20:1 ratio observed today (Simopoulos, 2016). Excessive intake of omega-6 fatty acids has been described as contributing to cancer and heart disease (German and Dillard, 2004). Miyamoto et al. (2019) described the protective effects of the microbiota to high-fat diets, by the metabolism of dietary polyunsaturated fatty acids (PUFAs), such as linoleic acid, to other beneficial metabolites. Mice colonised with PUFA metabolising strains of *Lactobacillus* were shown to have significant decreases in body weight and fat mass compared to the control group.

Kishino et al. (2013), demonstrated the capacity for Lactobacillus plantarum to metabolise unsaturated fatty acids. *L. plantarum* was demonstrated to possess the enzymatic pathway to saturate polyunsaturated fatty acids, generating hydroxy, oxo and conjugated fatty acids. In mice, there were notable differences in the levels of specific pathogen-free (SPF) mice and germ-free mice, with higher levels of hydroxy fatty acids present in SPF mice, demonstrating how the intestinal microbiome may play a role in shaping lipid metabolism.

The dietary habits of the host may impart compositional changes to the microbiome, in turn shaping its functional output. Watanabe et al. (2021) illustrated how diets high in cornstarch, fructose, branched-chain amino acids, soybean oil and lard could result in the formation of unique microbial profiles. Diets high in fats (Soybean oil and lard) were associated with obesity, glucose intolerance and insulin resistance in both groups. Glucose intolerance within mice fed the soybean oil diet was worse than those in the lard group. Faecal Microbiota Transplants to germ-free mice from soybean oil and lard groups were able to replicate the different glucose intolerance, suggesting that microbiota composition established by the diet was an influencing factor in glucose intolerance

1.2.3 Vitamins

The synthesis of vitamins is another important accessory function provided by the gut microbiome. While humans have the capacity to synthesise vitamin D, the gut microbiota and external dietary sources are relied upon as sources for other essential vitamins,

Magnúsdóttir et al. (2015) documented biosynthesis pathways for eight B vitamins within 256 common gut microbiota. They identified widespread production for each of the eight vitamins, approximately 40-60% coverage for each, and some bacteria possessing production pathways for all eight. Meanwhile, Wu et al. (2015) identified probiotic *Lactobacillus* strains *rhamnosus* GG, and *plantarum* as increasing Vitamin D receptor (VDR) protein expression in mouse and human intestinal epithelial cells. *In-vivo,* the increase in VDR protein expression induced by the *Lactobacillus* stains provided a protective function against *Salmonella*-induced colitis compared to wild-type strain and VDR knockout mice.

1.2.4 Bile acids

Bile salts are the result of the conjugation of host-derived bile acids with the amino acids taurine and glycine, being produced in the liver from cholesterol and excreted into the gut, with the majority being reabsorbed in the small intestine. Unabsorbed bile salt continues into the large intestine, where it is metabolised by the microbiota into secondary bile acids through the removal of hydroxyl groups. The conversion of primary to secondary acids is limited to a few bacteria within the gut, the majority of which belong to class *Clostridia* (Hirano et al., 1981). The primary bile acids within the human gut are cholate and chenodeoxycholate, which are then transformed by the microbiota.

The microbial transformation of primary bile acids into secondary bile acids diversifies the bile acid pool, broadening its functional capacity. Bile acids and secondary bile acids have been studied for their protective antimicrobial role in maintaining gut homeostasis (Metzendorf et al., 2022). Research has shown that centenarians possess distinct microbiotas, which have unique biosynthetic pathways for bile acids, *Sato et al. (2021)* describe pathways for the secondary bile acid isoallolithocholic acid by *Odoribacteraceae*, which exerted antimicrobial effects against gram positive bacteria.

The success of Faecal Microbiota transplants (discussed in detail, 1.4.5 Recovery from Dysbiosis) has been discussed as being partially mediated by the replenishment of Bile Salt Hydrolases (BSH) within the host (Mullish et al., 2019). A prominent reaction catalysed by BSH is the deconjugation of taurine and glycine from conjugated bile acids. The conjugated primary bile acid taurocholic acid has been identified as a trigger of *C. difficile* germination (Sorg Joseph and Sonenshein Abraham, 2008). Mullish et al. (2019) illustrated the potential importance of BSH in the resolution of recurrent *Clostridioides difficile* infection by demonstrating that *C. difficile* germination could be reduced when incubated with taurocholic acid when supplemented with the spent growth media of BSH expressing organisms.

The secondary bile acid deoxycholate, while positively influencing spore germination at levels similar to other cholate-derived bile acids, has an inhibitory effect on the growth of vegetive *C. difficile* (Sorg Joseph and Sonenshein Abraham, 2008). Thanissery et al. (2017) further documented the potential for secondary bile acids to inhibit taurocholic acid-mediated spore germination. While not a universal effect among *C. difficile, a variety of isolates (from ribotypes 027,017,012,001,078)*; deoxycholate, isodeoxycholate, lithocholate, isolithocholate, ursodeoxycholate, hyodeoxycholate inhibited spore germination on at least one of the clinical strains tested at physiological concentrations

Secondary bile acids have also been demonstrated to have anti-inflammatory action within the colon. Ward et al. (2017) documented the anti-inflammatory capacity of ursodeoxycholic acid (UDCA) and lithocholic acid (LCA). UDCA was able to inhibit the release of proinflammatory cytokines from colonic epithelial cells. In a mouse model of mucosal inflammation, UDCA was shown to reduce inflammation in a manner dependent upon the bacterial metabolism of ursodeoxycholic acid to lithocholic acid. Mice treated daily with UDCA lost less body mass and scored lower on the disease activity index than those without. Mice treated with LCA, a product of the bacterial metabolism of UDCA, were shown to have an enhanced capacity to inhibit the release of cytokines from colonic epithelial cells and were able to reverse changes to mucosal histology in vivo. 6α-methyl-Ursodeoxycholic acid, a nonmetabolising analogue of ursodeoxycholic acid, lacked its anti-inflammatory effects in vivo. The lack of

anti-inflammatory in 6α-methyl-Ursodeoxycholic acid suggests that the impact of UDCA in-vivo may be attributable to bacterial metabolism.

Ursodeoxycholic acid has also been demonstrated to possess chemopreventative properties on a variety of cancer cell lines as described by Akare et al. (2006). Ursodeoxycholic acid was shown to induce senescence in colon cancer cells and inhibit telomerase activity

1.3 Colonisation Resistance

Colonisation resistance refers to the protective effect that the microbiota provides the host (Freeman et al., 2003; Chilton, Caroline H. et al., 2012; Buckley et al., 2021). The mechanisms by which the microbiome confers colonisation resistance may be broadly characterised as being direct or indirect in nature. Direct colonisation resistance occurs due to interbacterial competition, often between similar bacterial strains (Lee, S.M. et al., 2013). Indirect colonisation resistance involves interactions between the microbiome and host pathways, such as the modulation of the immune system (Yamamoto et al., 2012) and metabolism of host-derived compounds such as mucin (Tailford et al., 2015) and bile salts (Sato et al., 2021; Metzendorf et al., 2022). Colonisation resistance is the first line of defence the human body possesses against invasion by exogenous bacteria and the suppression of potentially harmful pathobionts and quiescent pathogens.

1.3.1 Direct colonisation Resistance

1.3.1.1.1 Nutrient competition

The simplest method by which the microbiome provides colonisation resistance is through interbacterial competition for resources. This contest for nutrients and space excludes transient microorganisms from establishing themselves and prevents the outgrowth of opportunistic pathogens. These competitive interactions between species within the microbiome and the protective capacity they provide have been documented, notably in research by Deriu et al. (2013) and Momose et al. (2008), who demonstrated how endogenous *Escherichia* can suppress closely related species.

E. coli strain Nissile 1917 was initially isolated during an outbreak of *Shigella* during World War 1 from a soldier colonised with *Shigella,* but remained

asymptomatic (Sonnenborn, 2016). Deriu et al. demonstrated that Nissile 1917 is able to competitively suppress *Salmonella enterica serovar typhimurium* by out-competing *S. typhimurium* for iron. Competition for nutrients may encompass amino acids such as proline, which has been demonstrated to be a limiting factor in the growth of enterohaemorrhagic *E. coli* 0157:H7. In mice, the growth of 0157:H7 was significantly inhibited when inoculated into mice already colonised with a greater number of endogenous *E. coli* strains capable of proline metabolism.

The impact of the diet on the intestinal microbiome effectively illustrates how nutrient competition influences its composition, as the microbiota depend on nutrients derived from the host's dietary intake (David et al., 2014). Watanabe et al. (2021) effectively demonstrated how varying nutrient composition resulted in unique microbial profiles in mice. High carbohydrate fat and protein were shown to restructure the murine microbiome, additionally impacting the host metabolism (discussed previously in detail 1.2.2).

1.3.1.2 Antimicrobial Inhibition

Bacteria can inhibit the growth of closely related bacteria by the production of bacteriocins, short peptide antimicrobials broadly categorised by being produced by Gram-negative and Gram-positive bacteria. Perhaps some of the best characterised are lantibiotics produced by the *Lactobacillales.*

Nisin is a member of the lantibiotics class produced by *Lactococcus lactis,* which has antimicrobial activity against a relatively broad range of Gram-positive bacteria, such as *Staphylococcus aureus, Listeria monocytogenes* and *C. difficile* (Bartoloni et al., 2004). Nisin binds to lipid II, a precursor molecule in cell wall synthesis. By binding to lipid II, nisin induces the formation of pores in the membrane through the addition of lanthionine rings (Breukink et al., 1999) and by inhibiting cell wall biosynthesis (Asaduzzaman and Sonomoto, 2009) (Hasper et al., 2006).

Bacteriocins produced by Gram-negative bacteria are mainly limited to *Enterobacteriaceae* and are known as microcins. The probiotic strain Nissle 1917, which has been used as the active probiotic component in treatments of gastrointestinal diseases, has been identified as producing microcin's. Nissle

1917 has been shown to limit the expansion of *Enterobacteriaceae* (Sassone-Corsi et al., 2016).

Colicins are high molecular weight proteins toxic to some strains of *E. coli* as well as *Shigella* and *Citrobacter*. Colicins have a wide range of activities within the targeted cell. However, they rely on transportation across the periplasmic space unlike smaller peptide bacteriocins, which rely on specific receptors (Cascales et al., 2007).

A benefit of exploring the microbiome from a functional perspective is the identification and potential harnessing of its interactions, partially within the context of the rising concerns around multi-drug resistance. Bacteriocins, which differ from antibiotic due to their limited scope of activity, could offer new therapeutic options for targeted interventions.

1.3.1.3 Inhibitory Metabolites

Inhibitory metabolites are important considerations in the dynamic environment of the gut microbiome. They are the natural byproduct of bacterial growth and metabolism and may act to prevent the growth of competing organisms, helping to maintain homeostatic balance within the gut.

Beyond their role in contributing to gut homeostasis via host physiology, SCFAs have an inhibitory effect on other potentially harmful members of the intestinal microbiome. The inhibitory effect of SCFAs have been demonstrated on commensals, pathogens and opportunistic pathogens. (Chang Kai et al., 2024). One of the actions by which SCFA can inhibit other gut microbiome species is via intracellular acidification. This mechanism has been documented to inhibit antibiotic-resistant *Enterobacteriaceae* by potentially disrupting cell replication (Sorbara et al., 2019). High concentrations of SCFAs were shown to counter the competitive edge of O_2 and NO_3-based respiration by Enterobacteriaceae, possibly due to gas diffusion from surrounding tissue.

1.3.1.4 Type VI Secretion Systems

Type VI secretion systems (T6SS) allow for the export of molecules across the inner and outer membranes of gram-negative bacteria and into neighbouring bacteria and host cells (Chen, C. et al., 2019). T6SS can translocate effector proteins to adjacent cells, inhibiting growth. Components of T6SS appear to

share some relation to bacteriophage proteins and those in Type IV secretion systems (Zoued et al., 2013). To avoid self-targeting of the effector proteins cells expressing T6SS genes must also encode a protein which affords immunity to the cell. *Bacteroidetes* are one of the predominant phyla in the intestinal microbiome of healthy adults and show the widespread possession of T6SS genes (Russell et al., 2014). The predicted target of the effector proteins encoded by *Bacteroidetes* is predominately peptidoglycan. This suggests that the T6SS pathway is aimed at driving interbacterial antagonism rather than host-directed (Russell et al., 2014)

1.3.2 Indirect Colonisation Resistance

1.3.2.1 Immune System Conditioning

The integrity of the intestinal immune system and mucosal barrier is of vital importance as it is all that stands in the way of the translocation of potentially pathogenic bacteria to the rest of the body. The microbiome is a key component in the modulation of the immune system.

A comparison of germ-free (GF), specific pathogen-free (SPF), and gnotobiotic mice showed that in GF mice, there was little difference in the transcriptome of the small and large intestine, whereas in SPF mice, these areas possess direction transcriptional profiles (Yamamoto et al., 2012). This suggests that the microbiome could be driving host immune functionality by conditioning host gene expression. It was also demonstrated that timing of exposure is also essential for normal immune development in mice, as GF mice colonised after 5 weeks still failed to develop normal Toll-like receptor and Type-1 interferon pathways.

The effect of external influences on the microbiome and their potential impact on the microbiome's capacity to influence immune system modulation has been studied in mice. Low-dose penicillin exposure from birth alters the expression of genes exposed in immunity and results in a decreased production of antimicrobial peptides (Cox et al., 2014).

It has been suggested that members of the genus *Bacteroides* may be important to immune system function. *Bacteroides fragilis* is associated with increased expression of IgA-secreting cells, which play an integral role in the

immune function of mucous membranes and IgM-secreting cells, which form an important complement of the innate immune response (Grönlund et al., 2000)

Bacteroides thetaiomicon is associated with the expression of the angiogenin Ang4 by Paneth cells. Ang4 is a component of innate immunity within the gut, possessing microbicidal activity against bacterial and fungal pathogens (Hooper et al., 2003)

Clostridial species have also been described as having modulatory effects on the immune system. The preventive effect clostridial clusters IV and XIV have against inflammatory bowel disease has been associated with the accumulation of colonic T regulatory cells. Mice colonised with members of the clostridial cluster IV, and XI increase the presence of T regulatory cells (Atarashi et al., 2011).

1.3.2.2 Microbial Modification of Host-Derived Compounds

Bile acids are able to inhibit bacterial growth as they possess detergent activity, which damages cell membranes (Urdaneta and Casadesús, 2017) At high concentrations, bile acids will completely dissolve lipid membranes, leading to leakage and cell death, whilst at lower concentrations, bile acids alter membrane fluidity and permeability (Noh and Gilliland, 1993; Leverrier et al., 2003). In addition to the membrane disruptive effect, bile acids also induce damage to nucleic acids. In *Salmonella enterica,* bile exposure increases the frequency of nucleotide substitutions as well as frame shifts and chromosomal rearrangements (Prieto et al., 2004).

Primary bile acids can form soluble complexes with iron, increasing intestinal iron uptake by the host. Iron is a fundamental element for many cellular processes, and decreasing its availability in the lumen may inhibit bacterial growth as well as induce colicin production by E. coli (Sanyal et al., 1991).

The secondary bile acids deoxycholic acid and lithocholic acids have been found to have synergistic actions with 1-acetyl-b-carboline and turbomycin A, antibiotics which inhibit cell division of *C. difficile* (Kang et al., 2019) Secondary Bile acids have also been demonstrated to possess the ability to bind to and neutralise TcdB toxin one of the driving elements of CDI.

1.4 Dysbiosis

Disruptions to the gut's homeostatic balance can lead to large shifts in microbiota composition. These shifts often lead to the loss of colonisation resistance, which can have persistent immunological metabolic and physiological impacts on the host. Dysbiosis of the microbiome is often characterised by the loss of or proliferation of certain species within the gut. However, due to the significant variation in microbiome composition between individuals, defining dysbiosis purely through the perspective of microbial presence or absence presents significant challenges. Multiple microbial species may be capable of fulfilling similar ecological and functional roles within the microbiome, with each unique strain possessing the potential to contribute to the health and stability of both the microbiota and the host. Consequently, there may be greater benefit in exploring and defining dysbiosis from a functional perspective.

While natural factors, such as age-related microbial decline ,influence the microbiome composition, it is often the changes induced through external factors, such as diet, environmental exposures and xenobiotics such as food additives, drugs (Imhann et al., 2016) and importantly antibiotics that play a more obvious disruptive role (Muegge et al., 2011; Chilton, Caroline H. et al., 2012; Holmes et al., 2017; Singh et al., 2017; Gerassy-Vainberg et al., 2018; So et al., 2018; Macke et al., 2020; Venturini et al., 2021; Nel Van Zyl et al., 2022). These external factors are concerning due to their potential to rapidly and persistently alter microbiome composition, pushing it towards a dysbiotic state.

1.4.1 Age-Related Microbial Decline

As an individual reaches senescence, the microbiome begins to decrease in diversity. This reduction in diversity can lead to dysbiosis of the intestinal microbiota. Dysbiosis of the microbiome is common in the elderly and has been proposed as a potential biomarker of ageing (Bana and Cabreiro, 2019).

Age-associated microbial decline should be considered within the context of lifestyle changes that commonly accompany increasing age and frailty. Reduction in activity, increased medication, and diet are not universal amongst the elderly population and can, therefore, influence the trajectory of the microbiome towards homogeneity with increased age. Housing status within the

elderly impacts the rate of decline within the microbiome. Residents of long-term care facilities experience an increased decline in microbial diversity and populations associated with a more "youthful" microbiota within the population (Jeffery et al., 2016).

The influence of the microbiome on the ageing process itself is an emerging area of research. Bacterially produced indole has been demonstrated to increase health-span and change gene expression towards a younger profile in mice and *C. elegans.* This reflected finding further observations, where mice colonised with indole-producing *E. coli* displayed greater composite health scores (Sonowal et al., 2017).

Comparative taxonomic analysis of the microbiome with 16s by Xu et al. (2019) across 368 samples from participants ranging from newborn infants to the elderly put microbiome ageing on a continuum with gradual shifts occurring throughout the life.

Age-related microbial shifts have been described in associated alterations to specific taxa such as declining *Bifidobacteria* (Yatsunenko et al., 2012) and increasing abundances of *Bacteroidetes, Eubacterium* and *Clostridiaceae* (Odamaki et al., 2016).

Towards the extreme end of life, the microbiome has been shown to undergo significant changes in bacterial abundance. Luan et al. (2020) explored the microbial profiles of individuals at the extremes of life, noting that distinct changes occur in the months leading to death. The majority of these changes were associated with reductions in ten bacterial species, belonging to the genera: *Akkermansia*, *Alistipes*, *Bacteroides*, *Butyrivbrio* and *Prevotella*, with two species *Bifidobacterium longum* and *Ruminococcus bromii* increasing. These changes potentially preceding clinical symptoms of deterioration by up to 4 months.

The prolonged nature of the ageing process, combined with the relatively recent development omics fields such as metagenomics, has resulted in a scarcity of true longitudinal studies tracking the development of the microbiome across the human lifespan. Animal models, however, with shorter lifespans and comparatively accelerated ageing processes, offer opportunities to explore the

interplay between ageing and the microbiome (Sonowal et al., 2017; Catterson et al., 2018)

Alam et al. (2021) showed that some *Bacteroidetes* and *Firmicutes* populations remained relatively stable in ageing mice. This indicates that some bacterial populations may be more resistant to age-induced changes. Odamaki et al. (2016) noted that the observed alterations to *Bacteroidetes* levels occurred in a “stepwise” manner in individuals over 70 years old.

In mice, *Clostridiaceae* and *Lactobacillaceae* reduced during ageing (Alam et al., 2021). Interestingly these appear to be counter to observations in humans (Odamaki et al., 2016). These changes may be attributable to model organism differences or as Xu et al. (2019) demonstrated, that age-associated changes to the microbiome may not follow a fixed trajectory, and some taxa potentially increase before declining.

Reduction in bacterial stability and shifting composition may lead to dysbiosis and increase the risk of pathobiont expansion and colonisation by multidrug-resistant organisms (MDROs). Pathobionts are members of the microbiome that typically do not pose a risk to human health; however, when allowed to proliferate, they may prove detrimental to health. Dysbiosis may be associated with the creation of potential niches for colonisation by exogenous bacteria. Together, these factors may contribute to an increased risk of opportunistic infections, such as *Clostridioides difficile* infection (CDI), which is particularly associated with older adults (Eze et al., 2017).

1.4.2 Antibiotic-induced dysbiosis

Antibiotics can cause a profound loss in the abundance and diversity of the microbiome, disrupting its balance and potentially leading to dysbiosis. Despite the efforts being made with antibiotic stewardship programmes, global antibiotic use has continued to rise. Between 2000 and 2015, global antibiotic consumption increased by 65% in total doses and 39% in doses per 1,000 inhabitants per day (Klein et al., 2018).

Despite overall consumption remaining below levels seen prior to the implementation of COVID-19 measures, antibiotic consumption in England increased in 2023 to 17.61 defined daily doses per 1,000 inhabitants per day (DID) over the previous year 17.4. An increase in consumption was observed

across the majority of antibiotic groups, with the greatest proportional increase seen in anti-*C.difficile* agents (UKHSA, 2024).

In middle and low-income countries, antibiotic consumption is especially pronounced, heightening the risk of multidrug-resistant organism (MDRO) carriage (Andriatahina et al., 2010; Kiddee et al., 2019; Qureshi et al., 2021). Klein et al. (2018) estimated that if antibiotic consumption trends continue at their current pace, global consumption will reach 128 billion defined daily doses (DDD) (41.1 DID), a 202% increase over 15 years. Tackling antimicrobial resistance and antibiotic stewardship is particularly a problem in low- and middle-income countries. Iskandar et al. (2021) highlighted the specific obstacles encountered in these settings; being hindered by weak government oversight, inadequate healthcare infrastructure, additionally cultural practice, such as possible self-medication or over-the counter sale of antibiotics may lead to uncontrolled consumption. Insufficient laboratory infrastructure and diagnostic testing also leading to a reliance on empirical rather than targeted antibiotic usage.

1.4.2.1 Factors Influencing Antibiotics impact on the Gut Microbiome

The extent to which antibiotic disrupt the microbiome depends on a range of influencing factors. Variables such as pharmacokinetics, route of administration and spectrum of activity all have a role in the extent and nature of microbial disruption.

Oral antibiotics can affect the microbiome directly as they travel through the intestinal tract. Systemic antibiotics may reach the gut after being processed by the liver where they may be excreted into the intestine via hepatobiliary clearance or returned to the blood for clearance by the renal system. The hydrophobicity of an antibiotic can determine how it is cleared from the body. Hydrophilic antibiotics are more likely to excreted from the body via renal mechanisms whilst lipophilic are predominantly excreted via hepatic clearance (Vincent et al., 2016). Intestinal absorption is another important factor when considering what impact antibiotics may have upon the microbiome. For example, metronidazole is readily absorbed in the small intestine meaning that concentrations decrease along the gut. In contrast, vancomycin is poorly absorbed and therefore maintains high concentration (Kim, S. et al., 2017).

Poorly absorbed antibiotics may be expected to have greater potential to induce dysbiosis in the gut microbiome, clindamycin for example has poor intestinal absorption and is associated with several dysbiosis related conditions such as; antibiotic associated colitis, diarrhoea and *Clostridioides difficile* infection (CDI) discussed in detail in 1.4.3 Clostridioides difficile infection.

Studies in *in-vitro* models have explored the impact of antibiotics upon the microbiome. Antibiotic instillation at clinically reflective concentrations, with broad spectrum antibiotics; clindamycin, vancomycin, and piperacillin tazobactam have been shown to have considerable deleterious impacts on the intestinal microbiota, while this is less evident with more targeted antibiotics (Baines, S. D. et al., 2005; Baines, Simon D. et al., 2009; Rea et al., 2011; Chilton, C. H. et al., 2014).

1.4.2.2 . Effect of prophylactic Antibiotics on the gut microbiome

Prophylactic antibiotics have been used to prevent infections in immunocompromised patients such as those undergoing organ transplants (Centers for Disease Control and Prevention et al., 2000). Solid organ transplant and Haematopoietic stem cell transplant (HSCT) recipients are at an increased risk of developing infections due to immunosuppression as a result of anti-rejection drugs. The prophylactic antibiotic regimes required to defend the patient from pathogens and opportunistic infections can deplete the microbiota making their use controversial (Shono and van den Brink, 2017)

There is evidence that dysbiosis in HSCT patients is associated with an increase in adverse outcomes (Chong and Koh, 2020). The use of broad-spectrum antibiotics in the treatment of neutropenic fever in allo-HSCT recipients is associated with an increase of graft-versus-host disease related mortality at 5 years. The 3-year survival for allo-HSCT patients with high intestinal diversity is 67% compared to 36% in patients with low intestinal diversity (Taur et al., 2018)

1.4.3 Clostridioides difficile infection

Perhaps the most well-known outcome of antibiotic-induced dysbiosis is antibiotic-associated diarrhoea of which *Clostridioides difficile* infection accounts for 30% of cases (Bartlett et al., 1977; Anand and Glatt, 1993; Bartlett, 1996). CDI can cause a range of symptoms from mild antibiotic-associated diarrhoea to pseudomembranous colitis, and life-threatening toxic megacolon. CDI is a significant burden on healthcare systems. Tresman and Goldenberg (2018) estimate mean total healthcare costs associated with *C. difficile* infection at £12,710 (95% CI £9652-£15769) for primary infections, rising to £31,121 (95% CI 10792-42447) in recurrent cases.

Beyond the considerable financial burden, the impact of CDI on patient welfare can be immense. Patient and public involvement work indicates *C. difficile* infection patients suffer with debilitating impact to their quality of life and, mental health

In healthy individuals, the intestinal microbiota provides colonisation resistance to *C. difficile* colonisation and infection. CDI is particularly associated prior antibiotic use, disproportionately affecting older adults. The loss of diversity and key members of the microbiota, associated with advance age may play an additional, important role in the association of CDI with elderly patients.

Some antibiotics are associated with an increased propensity to induce CDI; clindamycin third generation cephalosporins and fluoroquinolones are all linked with CDI. A meta study by Sliming and Riley (2014) identified strong associations between: carbapenems,3rd and 4th generation cephalosporins, and moderate associations with: clindamycin, fluoroquinolones, and penicillins + beta-lactamase inhibitor combinations.

The risk associated with antibiotic use rises with dose, number of antibiotics received, and length of antibiotic therapy (Stevens, V. et al., 2011). Shifts in the microbiome as a result of antibiotic exposure can lead to changes in the functional capacity of the microbiome. These alterations can create a conducive environment for *C. difficile* expansion within the microbiome, leading to symptomatic disease.

Select members of the microbiota play a key role in the deconjugation of primary bile acids, transforming them into secondary bile acids (Hirano et al.,

1981) and disruption to the microbiota and therefore the pools of primary and secondary bile acids may play a role in CDI. Primary bile acids have been shown to contribute to the germination of *C. difficile (Wilson et al., 1982)*. The gut metabolome in *C. difficile* susceptible mice is associated with an increase in primary bile acids such as taurocholate, with a decrease in the abundance of secondary bile acids (Theriot et al., 2014). This dynamic, however, is complicated by the fact that the primary bile acid chenodeoxycholate is able to inhibit the germination of *C. difficile* spores through the competitive inhibition of taurocholate-mediated germination (Sorg and Sonenshein, 2010), highlighting the complexity of aetiology associated with the progression of microbiome associated diseases.

1.4.3.1 CDI Recurrence

Despite effective antimicrobial therapy, persistent dysbiosis following the treatment of primary CDI may lead to recurrence in 30% of cases, with the risk of recurrent CDI increasing with subsequent episodes ((NICE), 2021). This can be compounded by further use of antibiotics used to treat the infection (Eze et al., 2017). Recurrent CDI is defined by the additional instances of CDI which occur within 8 weeks of previous episodes (Debast et al., 2014)

Metronidazole, vancomycin and fidaxomicin are used to treat CDI. However, metronidazole is no longer recommended as a first-line treatment for CDI due to poor efficacy (McDonald, L.C. et al., 2018; (NICE), 2021; [NICE], 2021). Vancomycin remains widely used but has a profound effect on the microbiome, possibly perpetuating the dysbiosis and leaving patients vulnerable to further infection (Lewis, B.B. et al., 2015). Fidaxomicin has been identified as an effective treatment of CDI and, due to its narrower spectrum causes less collateral damage to the microbiome and is associated with lower rates of recurrence (Ajami et al., 2018)

1.4.4 Colonisation by Multidrug-resistant Organisms.

As antimicrobial resistance continues to rise, the perseveration of existing antibiotic agents has become a critical factor in the continuing efficacy of many clinical interventions that modern medicine has come to rely on. Organ transplantation, surgical intervention, and cancer treatments all require the support of antimicrobial therapy to ensure patient well-being. Perhaps the

greatest threat facing modern medicine is the rise of multi-drug-resistant organisms (MDROs) and the spread of antibiotic resistance genes. It has been suggested that in the next 50 years, antibiotic-resistant bacteria could be responsible for the death of over 10 million people per year (O'Neill, 2014). Dysbiosis and the loss of colonisation resistance increase the potential for an individual to become colonised by exogenous bacteria (Mahieu et al., 2017)

Venturini *et al.* (2021) explored the potential for antibiotics to alter the composition of the murine microbiome, leaving it susceptible to colonisation by MDROs. In this study, mice were treated with piperacillin-tazobactam (TZP) and ceftriaxone (CRO) for 5 days; at the height of antibiotic-induced dysbiosis, colonisation resistance was tested with multi-drug-resistant *Escherichia coli* and *Klebsiella pneumoniae.* Dysbiosis was associated with disruptions to the murine gut in composition and diversity, leaving it susceptible to colonisation by both multidrug-resistant organisms (Venturini et al., 2021).

Ducarmon et al. (2021) explored the prevalence and risk factors associated with MDRO colonisation among nursing home residents, finding that 44.4% (12) of the participants within the study were colonised at least one time-point during the study. ESBL-producing bacteria were the most common MDROs identified in the study within 22.2% (6) of participants. Antibiotic use was identified as the second highest risk factor (OR 2.84 95% CI 0.85-9.49) in MDRO carriage after hospital admittance in the past year (OR 3.78 95% CI 0.69-20.70). *Dorea, Lachnospiraceae_ND3007_group* and *Atopobiaceae* were identified as being consistently more abundant in patients never colonised by MDROs within the study.

In their systematic review, Sulis *et al.* (2022) examined the association between antibiotic exposure to Access, Watch, and Reserve (AWaRe) antimicrobials and MDRO colonisation. Previous exposure to any AWaRe group antibiotic was associated with increased colonisation with MDROs. Reserve-class antibiotics were associated with the greatest risk of MDRO colonisation. Linezolid (a Reserve antibiotic) was associated with the greatest risk of those antibiotics profiled (OR 2.56 [95% CI: 2.09 – 3.14). These findings highlight the potential for antibiotics (particularly those within the Reserve classification) to drive MDRO colonisation potentially as a byproduct of antibiotic-associated dysbiosis.

1.4.5 Recovery from Dysbiosis

Recovery of the microbiome following disruption has been shown to be heavily reliant upon environmental reservoirs within the host population. Ng *et al.* (2019) demonstrated the importance of community reservoirs, such as shared living environments and social interactions in the recovery of the microbiome following antibiotic treatment. Microbiotas of co-housed mice recovered rapidly following the withdrawal of antibiotics. Mice kept in isolation showed impaired recovery of the microbiome when compared to co-housed mice. This finding potentially suggests that critically ill patients, who are often kept in isolated environments to prevent further infections, could face challenges in microbiome recovery due to the lack of environmental exposure increasing recovery time.

Strategies are now being explored to aid in microbiome recovery, such as faecal microbiota transplantation (FMT), which has been used successfully to restore the microbiome in patients experiencing recurrent CDI. FMTs involve the transfer of faecal microbiota from healthy microbiomes to those who experience dysbiosis. These may be from transplant donors or autologous transplants where the recipient undergoes an infusion with their own microbiome with healthy stool prepared before treatment which may impact the microbiome. A systematic review by Shogbesan *et al*. (2018), explored FMT outcomes in immunocompromised patients. They Identified success rates of 88% to 93% following single and multiple FMTs within the general population, whereas patients who were immunocompromised were associated with lower success rates of 78% - 89%. Among the 303 immunocompromised patients, 2 deaths were identified within the following 30 days; however, they could not be directly attributed to FMT, CDI or underlying immunocompromised states. Additionally, they identified 5 instances of bacteraemia following FMT.

The transfer of undefined microbial populations such as with FMT may have unforeseen consequences. Following two instances of bacterial infections cause by multi-drug-resistant organisms, the US FDA issued additional recommendations for donor screening for FMT. Invasive infections caused by extended-spectrum beta-lactamase (ESBL)-producing *E. coli* in two immunocompromised patients, which proved fatal in one instance, were attributed to FMTs from a single donor; the donor stool and FMT had not been prior to transplantation (FDA, 2019).

Due to the risks associated with FMT, strict guidelines have been proposed for donor selection (Cammarota et al., 2017). FMT is regulated in the UK and only recommended for the treatment of severe and recurrent infections of CDI ([NICE], 2022). Despite the risks, donor FMT has been shown to be highly effective in the treatment of recurrent CDI (Kelly, C.R. et al., 2016),

The use of autologous FMT has been explored in allo-HSCT patients by Taur et al. (2014). 25 patients about to undergo antibiotic prophylaxis processing transplantation were asked to freeze and bank stool samples to aid the reconstruction of their microbiome post antibiotics.14 received autologous FMT, and 11 were placed in a control group. Prophylactic antibiotic treatment was associated with disruption of the microbiome, leading to a domination of enterococcus within the microbiome. Autologous FMT was associated with increased diversity, a 63.8% (95% CI 36.3 - 91.2%) increase to inverse Simpson Index) over those in the control group not receiving autologous FMT who saw a 38.7% (95% CI 17.9 – 59.6%) increase to their inverse Simpson index.

1.4.5.1 Microbiome therapeutics

As advancing knowledge of host-microbiome interactions enables better characterisation of dysbiosis, opportunities arise to develop more targeted restorative therapies. Researchers are beginning to explore how defined bacterial consortia can restore the microbiome to a healthy state, reducing the associated consequences of less defined treatments (Caballero et al., 2017). Given the ongoing burden of CDI and the challenges in treating recurrent infections, many targeted consortia interventions focus on enhancing the resolution of recurrent CDI.

SER-109 is a notable example of a restorative microbiome therapy targeted at treating recurrent CDI. Composed of live *firmicute* spores positioned to compete with *C. difficile* for essential nutrients and aid in the re-establishment of colonisation resistance by altering bile acid profiles. A clinical trial exploring the efficacy of SER-109 in preventing further recurrent CDI in patients who had previously experienced three or more episodes of CDI found that patients receiving SER-109 were found to have lower rates of recurrence compared to placebo (12% and 40% respectively) (Feuerstadt, P. et al., 2022).VE303 is a

similarly positioned therapeutic developed as a consortium of commensal clostridia, has shown similar promise, reducing recurrence in patients who have experienced one or more episodes of CDI to 13% in “high dose” administration of VE303, compared to 45.5% in the placebo group (Louie et al., 2023). In a diversion from the oral therapeutic SER-109 and VE303, RBX2660 (REBYOTA) is a rectally administered “microbiota product” also aimed at treating recurrent CDI. The largest study of live microbiota therapeutics to date, the PUNCH study explored the treatment and resolution of recurrent CDI with REBYOTA. The treatment success rate, defined as an absence of CDI diarrhoea through 8 weeks, was 73.8% and the clinical response rate, defined as resolution of current CDI episode and no new CDI episode at 6 months, was 91.0% (Feuerstadt, Paul et al., 2024)}.

1.5 Identifying dysbiosis

There is a complex relationship between the benefits of resolving infection and the potential to compound damage to the microbiome, potentially leading further down the path to relapse and antibiotic resistance. The immediate need to resolve life-threatening infections is now more carefully considered within the broader implications to microbial integrity, diversity and wider gut health. Identifying biomarkers of dysbiosis would aid in addressing this complex issue. Identifying at-risk compositional and functional profiles and the antibiotics potentially complicit in their destabilisation.

1.5.1 Biomarkers of Dysbiosis

Biomarkers are employed as clinical tools for assessing pathological conditions and physiological function. Acting as diagnostic indicators for biological processes (Biomarkers Definitions Working Group, 2001; Strimbu and Tavel, 2010), these markers originate from the host or external sources such as the microbiota and may fall within several classifications such as genetic, proteinaceous or metabolites.

Predictive indices are composite tools constructed from multiple variables, including biomarkers. While they do not fall explicitly within the classification of biomarkers, they are discussed here alongside single-point biomarkers

Compositional and metabolic markers from dysbiotic microbiomes have been identified as being associated with colorectal cancer (CRC) development and progression. Markers may be associated with the prevention of tumour growth or linked to microbiomes associated with CRC progression (Purcell et al., 2017; Avuthu and Guda, 2022).

The Gut Microbiome Health Index (GMHI) is a composite score assessing 50 microbial species associated with a healthy microbiome. It is able to distinguish between healthy and non-healthy Individuals with an accuracy of 73.7% (Gupta et al., 2020). The index is built on data from shotgun metagenomes from 4347 participants and compares the relative abundance of microbial populations at the species level, grouped by their associated good and adverse health conditions. The unhealthy classification is an aggregation of 12 disease states and abnormal body weight conditions. While the index may be successful in distinguishing large compositional changes between healthy and dysbiotic states, it lacks the power to differentiate between diverse diseased states. Importantly, it is unclear whether shared microbial characteristics across a range of conditions such as: overweight, underweight, Crohn's disease and colorectal cancer are sufficiently informative, or if state-specific indices provide more descriptive utility. Finally, since GMHI is based solely on population demographics and shotgun metagenomics, it may lack features required in diagnostic tests: metagenomic sequencing host and data processing times exceed what would be appropriate clinically and a diagnosis of "unhealthy" may not be appropriate or specific enough for therapeutic intervention.

The Microbial Dysbiosis Index(MDI) (Toto et al., 2024) was designed as a tool to assess disruptions to the intestinal microbiome associated with inflammatory bowel disease(IDB). MDI uses both clinical observations and taxonomic markers, with several bacterial genera described as being positively associated in stools and ileal biopsies of IDB patients. The metabolic state was assessed with predictive analysis based on 16s data and PICRUST and confirmed with indican levels in urine, a product of bacterial metabolism of tryptophan within the intestinal tract.

Predictive markers based on microbial populations at the point of admission have been explored and applied to Antibiotic-associated diarrhoea (ADD) and CDI (Berkell et al., 2021). Findings indicate that patients with ADD and CDI

commonly show higher abundances of pathobionts from the *Enterococcus* genus (*E.faecium* and *E.faecalis)* in contrast to reductions in taxa such *as Blautia, Ruminococcus, Porphyromonas, and Bifidobacterium* observed as being lower in abundance. The ratio of *Enterococcus*-to-*Ruminococcus* ratio was identified as being an effective marker for CDI.

These largely taxonomy-based indices show strong associations with specific populations with well-defined clinical statuses. However, there is little evidence supporting their effectiveness as predictive or diagnostic tools. Offering relatively binary approaches, they aim to define a broad dysbiotic state. In the case of MDI, this dysbiotic instance is specifically contextualised for its association with IDB patients but does not explicitly mention any capacity to differentiate IDB from other gastrointestinal diseases. Given these limitations, indices of this nature may currently have limited diagnostic potentials. However, they may be useful as predictive tools for disease progression.

Microbiome-based health indices offer significant promise in diagnosing and classifying dysbiosis. However, relying solely on taxonomic associations in their construction is inadequate, as it overlooks critical functional and metabolic factors. A multi-omic approach, incorporating more diverse biomarkers, may provide a more refined framework for assessing microbiota health and functionality.

Research has begun to explore the integration of metabolomic markers with taxonomic indicators. An investigation into hospital-acquired diarrhoea and CDI compared antibiotic-associated cases with non-antibiotic associated instances and identified with antibiotic-associated instances, there were unique microbiota profiles of *Enterococcus*-dominated and non-*Enterococcus* dominated microbiomes, and an increase in *Enterococcus* species was positively associated with CDI. Within both groups, *C. difficile* detection was associated with an increase in the presence of the SCFA butyrate (Bosnjak et al., 2023). Providing an interesting perspective on the integration of multi-omics in microbiota investigations, the methodology invites consideration regarding the practical application of identified biomarkers. *C.difficile* infection, in this instance, was diagnosed through the detection of GDH and Polymerase Chain Reaction (PCR) for Toxin B (t*cdB*) and binary toxin (*cdtA*), which is not sufficient to diagnose CDI. The utility of butyrate as a biomarker singular

biomarker of CDI is limited, as butyrate production is not limited to toxigenic *C. difficile*, particularly when effective biomarkers (such as *C. difficile* toxins) with validated assays for CDI exist. Despite these limitations, this study effectively demonstrates the potential of metabolic markers in identifying CDI and gut dysbiosis.

1.6 Approaches to identifying markers of dysbiosis

1.6.1 Meta-omics

Omics techniques transformed the understanding of genetic information flow and cellular processes within singular organisms; Meta-omics approaches apply the same framework to study complex microbial ecosystems, capturing the intricate dynamics and interactions that shape microbial communities (Barker-Tejeda et al., 2024). By applying these techniques to the gut microbiome, it is possible to study both the compositional and functional characteristics in a way that can contextualise microbial activity to human health outcomes (Yachida et al., 2019; Kappel et al., 2020; Liu, X. et al., 2023).

Meta-omics encompasses multiple omics modalities: genomics, transcriptomics, proteomics and metabolomics, each targeting distinct layers of microbiome function. Combining these approaches, it is possible to explore the microbiomes from complementary perspectives (Figure 1-1 Meta-omics Approaches Along the Central Dogma).

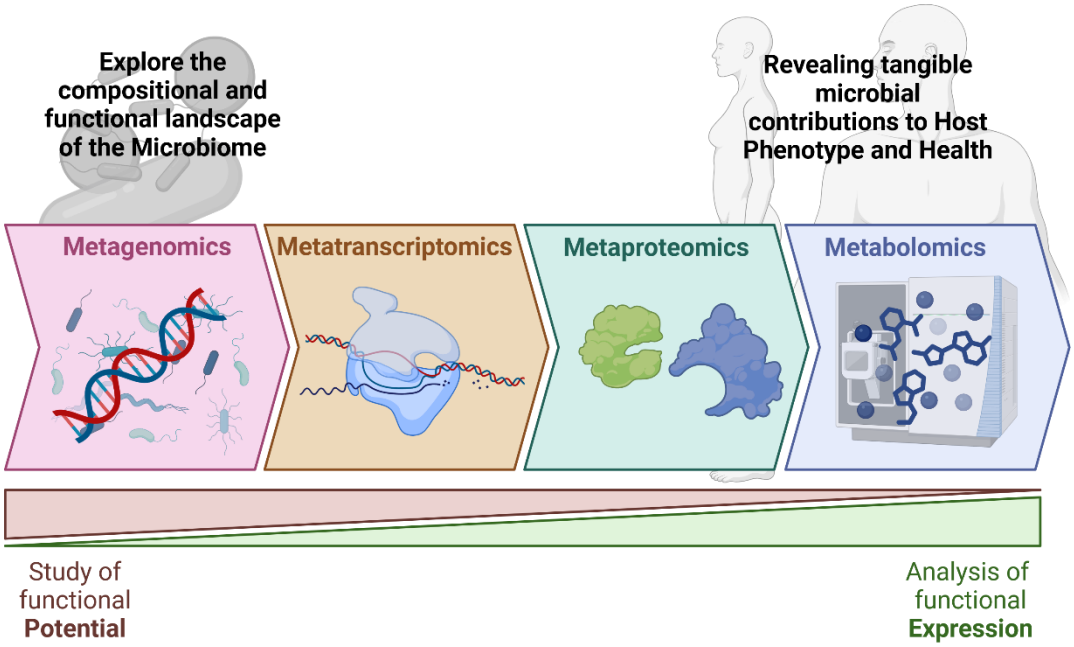

Figure 1-1 Meta-omics Approaches Along the Central Dogma. Illustrating how different modalities are applied from the study of functional potential to phenotypic expression

Metagenomic analysis has a well-established role in microbiome research, providing extensive compositional and functional information for the human intestinal microbiome (Human Microbiomes Project Consortium, 2012) (McDonald, D. et al., 2018) (Tigchelaar et al., 2016).

When designing multi-omic studies, it is important to consider the suitability of sequencing methodologies to the research question. For studies exploring compositional changes at a genus level 16s rRNA and shotgun sequencing may perform comparably (Pearson, 2013). However, shotgun sequencing has been shown to detect an increased range of the microbiota within the gut. Shotgun sequencing offers a greater resolution able to detect a broader range of the microbiota including less abundant taxa. This increased capacity to recover minor constituents of the microbiome may provide a more nuanced insight into compositional and functional shifts, enhancing the resolution of microbiota analysis.

Protein homology-based alignments are employed to enhance computational efficiency in metagenomic workflows by aligning translated protein sequences rather than genomic sequences, redundancy is reduced. These translated

microbial sequences are ideally matched to a non-redundant sequence database (O'Leary et al., 2016) to facilitate taxonomic and functional binning.

Metabolomics, a relatively new field, complements metagenomics by focusing on the characterisation of the metabolome, small molecules that are the result of the biological activity of both host and microbiota (Wishart et al., 2007; Wishart et al., 2022). In *in-vitro* studies, particular attention may be directed towards microbiota-specific mechanisms; while this may simplify the complexities of *in-vivo* symbiosis, *in-vitro* studies allow for the controlled study of microbial process in a more focused and manageable environment than can be achieved with *in-vivo.*

The integration of metagenomic and metabolomic approaches has shown promise in characterising the intestinal metabolome within animal models.

Research indicates that even with relatively short courses of single antibiotics, significant alterations can occur within the intestinal metabolome, affecting the abundance of 87% of identified compounds, implicating multiple metabolic pathways. Additionally, the effects of enrofloxacin and vancomycin have been observed to alter microbiome composition, host immune response and metabolome in mice (Sun et al., 2019).

Antibiotic-associated atherosclerotic lesions have been identified as being associated with a loss of microbial abundance and diversity. Compositional findings accompanied by untargeted metabolomics on serum samples indicated that antibiotic treatment-induced effects on the metabolome, particularly tryptophan and secondary bile acid metabolism (Kappel et al., 2020).

A ten day course of cefoperazone, induced dysbiosis in mice, making them susceptible to *C. difficile* infection, resulting in compositional changes and increases in key metabolites such as bile acids and carbohydrates conducive to bacterial growth (Theriot et al., 2014), cefoperazone associated dysbiosis in mice was similarly associated with *C. albicans (Gutierrez et al., 2020).*

While multi-omic investigations in animal models reveal valuable insights into the implications of antibiotic-induced gut dysbiosis, the dynamics in human patients are characterised by complex disease states with multifaceted aetiologies.

Multi-omic studies utilising metabolomics have been used to explore the microbiome of Crohn's disease and ulcerative colitis patients, associating metabolic shifts with host inflammation, with up to 31% of features differently abundant in IDB patients. Bile acids and long-chain fatty acids were identified as compounds of interest (Franzosa et al., 2019).

Efforts are being made to catalogue compounds associated with the human metabolome encompassing host and microbiota derive products. The Human Metabolome Database is one such project (Wishart et al., 2022), providing a resource of compound information and contextual and identification characteristics for metabolome derived compounds.

Due to the variety of statistical methodologies which underpin -omics studies, there are issues with comparability between studies. The gut microbiome-metabolome dataset is an aggregation of 14 human gut microbiome-metabolome studies, that seeks to address that through its unification and curation of datasets (Muller et al., 2022). With emerging fields, there will inherently be variations in the processing and utilisation of data sets, which can make comparisons between data and the identification of biologically meaningful truths challenging, which is evident from the evolving perspectives researchers have on the microbiome.

While human and in-vivo models offer powerful insights into the intestinal metabolome, there are practical and ethical limitations to the studies which can be performed. While in *in-vivo* models, it is common to conduct cross-sectional analysis on individual subjects, sampling and analysis methods often necessitate sacrifice, and due to this, longitudinal *in-vivo* studies explore the dynamics across populations over time. *In-vitro* models allow the study of single systems within highly controlled environments to explore the mechanisms of dysbiosis over long periods of time with multiple sampling opportunities.

1.6.2 In-vitro Model systems

The ability to accurately model the complex interactions of the microbiota is crucial to the future success of microbiome research. Model systems are required to be able to maintain complex and stable microbial communities for the duration of study, and while *in-vivo* models such as mice are common, particularly in studies implicating the host immune response, *in-vitro* models are a popular alternative for studying the microbiome due to ethical and practical considerations. The aim of using *in-vitro* models in human microbiome research is to accurately replicate the conditions of the human colon within a highly controlled system.

In-vitro models for studying the microbiome can be broadly categorised into three classifications: Batch fermentation, continuous single-stage and continuous multistage models. Model systems vary in complexity and suitability to research questions.

1.6.2.1 Batch Fermentation models

Batch fermentation models offer the simplest in-vitro system, in which bacteria are maintained in a single vessel with controlled internal conditions. They potentially offer an avenue for highly parallelised screening of the microbiome over relatively short time frames (Roupar et al., 2021). The simple nature of these models causes nutrients to deplete and the build-up of potentially inhibitory metabolites within the vessel, limiting the feasibility of longitudinal studies.

1.6.2.2 Continuous single-stage models

Continuous single-stage models represent the next conceptual development in the progression of *in-vitro* models, offering a higher level of complexity compared to batch fermentation methods (Gianetto-Hill et al., 2023).

Unlike batch fermentation systems which are closed environments, continuous single-stage models allow for the continuous replenishment with fresh growth media and the removal of potentially inhibitory metabolites produced by cultures. These factors have the potential to significantly extend experimental lifespan, enabling long-term cultivation of microbial community under conditions which may more closely mirror those found *in vivo*. In addition to nutrient

replenishment and removal, turnover rate can be used to simulate gut transit times, exploring transient colonisation and antibiotic washout from the intestinal microbiome. These factors can allow for more accurate monitoring of microbial growth rates and metabolic outputs than would be achievable with batch models.

Longer experimental time allows for continuous single-stage models to explore a wider range of research questions exploring how the microbiome changes over time, such as in response to antibiotic instillation.

Despite their ability to support longer experimental durations, due to their simplicity continuous single stage models fail to capture the physiological variable along the gastrointestinal tract (Procházková et al., 2024). This limitation prevents them from representing the distinct environmental conditions that shape the microbial communities across different regions of the gut. As a result of this single stage models cannot fully represent the complexity of the trophic interactions that emerge both within and between these environmental niches.

1.6.2.3 Continuous multi-stage models

Continuous multi-stage (CMS) models represent an expansion of the single-stage systems discussed earlier. Designed to study the intestinal microbiome, these most have been developed to more effectively replicate and support the diverse microbial ecology found within the human colon. The model utilised within this thesis is a development of those described and validated against the intestinal contents of sudden death victims by Macfarlane *et al.* (Macfarlane et al., 1998).

The conditions within the gastrointestinal vary significantly along its length, with considerable variations occurring within the colon alone. Multistage chemostats aim to provide an effective *in vitro* representation of this diverse range of environments.

Gastrointestinal transit, varying pH and nutrient availability can be replicated and maintained for extended periods of time with CMS in a manner not possible with batch or single stage systems.

The ability of these models to accurately maintain conditions and support microbial populations representative of the human colon for sustained periods

has led to their use across numerous studies. They have been used extensively to study the intestinal microbiome, being employed in pre-clinical research (Roberts et al., 2020) with notable studies investigating the impacts of antibiotic therapy and the potential to develop antibiotic-associated diseases (Saxton et al., 2009; Chilton, Caroline H. et al., 2012; Moura et al., 2022; Davis Birch et al., 2023), and exploring restorative therapies such as FMT (Moura et al., 2020). The versatility and modularity of these systems have enabled adaptation with biofilm-supporting structures to explore potential sources of recurrent infection in CDI (Crowther et al., 2014; Normington et al., 2021).

CMS models have been further expanded beyond modelling the colon. The Simulator of the Human Intestinal Microbial Ecosystem (SHIME) expands CMS systems to reactors representing the stomach and small intestine in addition to the colon (Van de Wiele et al., 2015). Notably all vessels within the SHIME model are seeded with faecal microbiota. While this does have distinct practical advantages, the suitability of faecal microbiota for upper gastrointestinal tract models is questionable due to variations in microbial composition and density between regions (Leimena et al., 2013) (Human Microbiomes Project Consortium, 2012) (Sender et al., 2016)

Understanding the limitations of any model system is an important step in applying that system to research. Within the realm of statistical models, the idea by George E. P. Box that “all models are wrong, but some are useful” is widely acknowledged. All model systems are inherently flawed, but some may be powerful research tools that can be applied to study the microbiome from varied perspectives.

In vitro models thus provide an avenue to explore trophic interactions of the microbiota under controlled environmental parameters. They allow for spatial investigations of the colonic microbiota that would require surgical procedures *in vivo* and involve fewer ethical constraints than animal models. The insights gained from *in vitro* studies can provide valuable insights to inform and effectively allocate resources in clinical investigations in patient populations.

1.7 Thesis outline

This thesis comprises four investigatory chapters in which distinct aspects of microbiome analysis are addressed using culture, metagenomics and metabolomics in the context of antibiotic-induced dysbiosis.

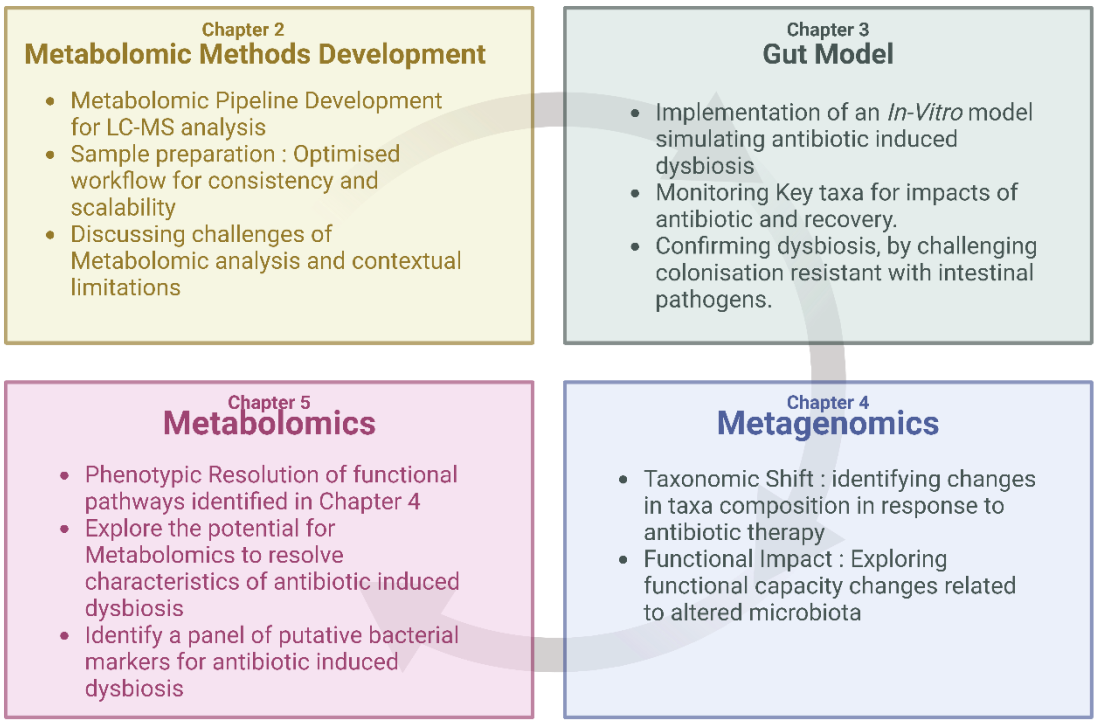

Figure 1-2 Thesis Structure: Method development, In-vitro modelling and multi-omic approaches for exploring the intestinal microbiome. Summary of key experimental chapters contained within this thesis, focusing on exploring the impacts and identification of potential biomarkers associated with antibiotic associated dysbiosis.

1.7.1.1.1 Chapter 2: Metabolomics Method Development

This chapter details method development for a unified Liquid Chromatography Mass Spectrometry pipeline facilitating untargeted metabolomics investigations across *in vitro* models and clinical patient samples.

1.7.1.1.2 Chapter 3: Gut Model

This chapter covers the implementation of in-vitro models for antibiotic-induced dysbiosis, assessing both the immediate impacts of antibiotics on the intestinal microbiome and potential for microbial recovery.

1.7.1.1.3 Chapter 4: Metagenomics

This chapter begins omics investigation into the microbiome, presenting a metagenomic analysis of the long-term effect of antibiotic therapy, comparing the bacterial community impact and recovery assessed via metagenomic to those identified via culture-based method in chapter two. Metagenomics allows for more comprehensive insights into the microbiome composition as well as being able to explore the potential functional capacity of the microbiome.

1.7.1.1.4 Chapter 5: Metabolomics

This chapter explores the phenotypic resolution of antibiotic-induced gut dysbiosis. Building on the microbial compositional and functional changes identified in Chapter 4, using the Liquid Chromatography Mass Spectrometry pipeline developed in Chapter 2.

1.8 Thesis Aims

This thesis aims:

- To develop an effective and scalable pipeline for metabolomic analysis of in-vitro gut models and human faecal samples.
- To determine whether *in-vitro* gut models can be used for the investigation of antibiotic induced dysbiosis
- To identify putative bacterial markers of antibiotic induced gut dysbiosis, using multi-omics approaches (metagenomics and metabolomics) in *in-vitro* gut models

Chapter 2
Metabolomic Methods Development

2.1 Introduction

Metabolomics is an emerging and rapidly evolving area of research, finding applications across a wide range of research questions (Yang et al., 2019). The field investigates the collective spectrum of small molecules, substrates, metabolic intermediaries and endpoints of biological processes providing insight into the cumulative expression of the genomic and functional capacity of a system. By studying the metabolomes of various biological matrices, such as faeces, blood, and tissue extracts, it is possible to gain critical functional insights into health and disease (Tsukuda et al., 2021; Dekkers et al., 2022). However, due to the field's evolving nature and the unique methodological challenges posed by the diverse range of sample types, there is little standardisation within the field, creating an obstacle to workflow selection (Wang, Z. et al., 2018; Erben et al., 2021; Kelly, P.E. et al., 2022).

This thesis utilises Liquid Chromatography Mass Spectrometry (LCMS) to explore the intestinal microbiome through in vitro gut models to identify possible functional biomarkers of antibiotic-induced dysbiosis, aiming to validate these putative markers in clinical samples.

The coupling of chromatography, either liquid (LC) or gas (GC), with Mass Spectrometry (MS) results in a versatile analytical tool which can simplify highly complex samples through the separation of components over time. LCMS has advantages in its board compatibility in compounds (Pitt, 2009), including those molecules with lower volatility, which may be less suited to GCMS. A large part of the versatility of LCMS can be attributed to the range of column chemistries that can be utilised for sample fractionation, which adds dimensionality to mass spectrometry data through chromatographic retention times, increasing the specificity of compound identification.

This chapter covers a spectrum of methodological considerations for metabolomics pipeline development (Figure 2-1 Overview of LCMS Metabolomic workflow with Validation steps for analysis), including LCMS Setup, optimisation of extraction methodology for high-throughput analysis,

validation of data-dependent Mass Spectrometry analysis, data annotation and post-processing.

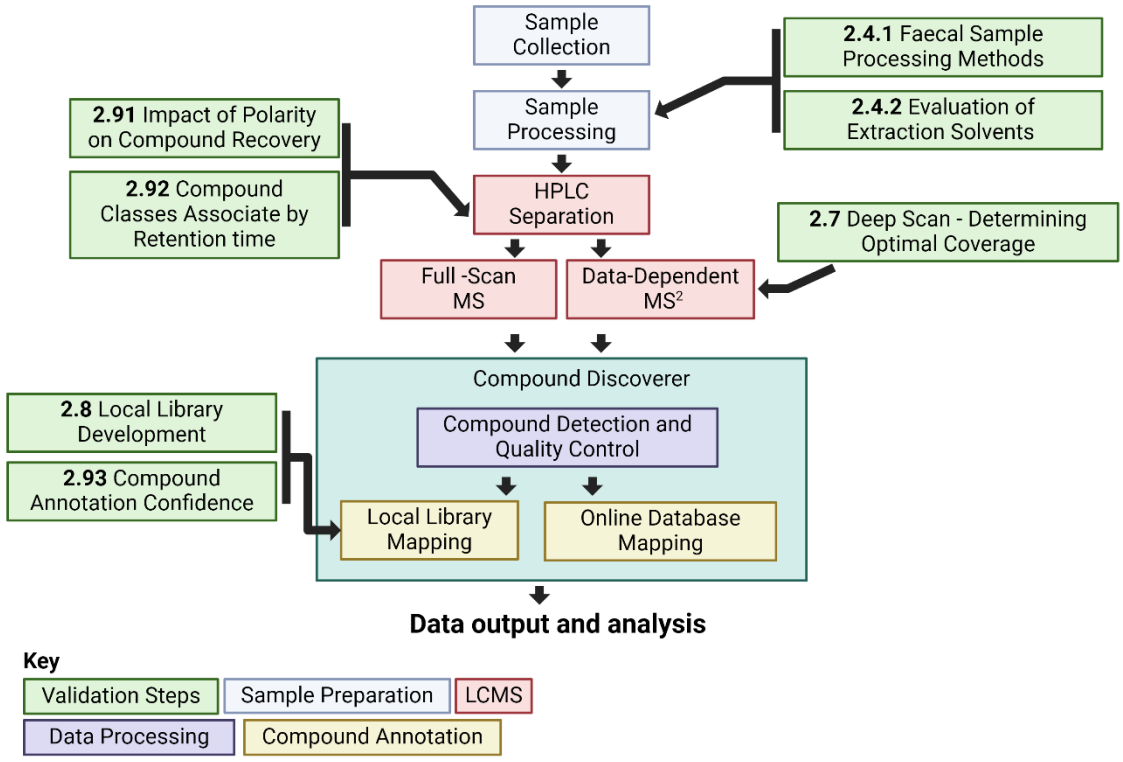

Figure 2-1 Overview of LCMS Metabolomic workflow with Validation steps for analysis. Outlining the key steps in metabolomic analysis, including sample collection (Blue), LCMS analysis (Red), data processing (Purple), and compound annotation (Yellow), is listed. Chapter subheadings that address the validation steps taken are outlined (Green).

Whilst metabolomics is widely used, its application to multistage chemostats such as the MiGut and Leeds human gut model represents a novel area of research. It warrants developing and validating a unified workflow for both in vitro samples and faeces; because of this, it was necessary to develop a pipeline suitable for use within the research group and the requirements of the University of Leeds community.

Standardising practices through developing a group-specific pipeline that addresses the analytical challenges associated with faecal and in vitro samples will ensure reproducibility and comparability for future investigations. When developing the workflow, it was important to consider the utility of the pathway for both sample types, as it is anticipated that metabolomic insights from in vitro

investigations may pave the way for translation biomarker validation studies on clinical samples.

2.2 Chapter Aims

- **Establish a reliable LCMS pipeline**: Sample extraction, LCMS analysis and the post-processing of *in vitro* and faecal samples in metabolomic analysis
- **Define the analytical boundaries of the pipeline**: identifying which compounds fall within the realm of whose which can be reasonably determined, assessing its suitability for exploring the intestinal microbiome both *in vitro* and via faecal samples.
- **To improve compound annotation depth and accuracy**: by developing an in-house LCMS spectral library.

2.3 Experimental outline

In the context of a novel application for the university and research group, the absence of established metabolomics expertise necessitated developing and validating a new pipeline. The steps taken in the following sections aim to ensure the reliability of the pipeline and begin to define experimental boundaries by exploring the analytical recovery of known metabolites.

2.3.1 High-Performance Liquid Chromatography, Sample Fractionation

High-Performance Liquid Chromatography (HPLC) is a powerful technique that enables the fractionation and analysis of highly complex samples based on the physiochemical properties of analytes within the sample. This process results in the gradual elution of compounds throughout a chromatographic gradient by gradually altering the composition of the mobile phase. The elution times of compounds can be used as data points, allowing for more specific matching of compounds to known standards. Chromatographic retention times can potentially be used to differentiate compounds of the same m/z through different chemical properties. Upon elution, analytes are directed to the MS; the reduction in the number of compounds entering the MS enhances the sensitivity and specificity of compound detection. Column chemistry selection may greatly impact the type of compounds which can be recovered; this factor is the reason for the validation steps discussed in 2.9 Metabolite Identification Limitations. This thesis utilises Carbon-18 (C18) reverse phase HPLC. Column chemistry is discussed in detail in the section. 2.5.1 High-Performance Liquid Chromatography.

2.3.2 Full Scan and Deep-Scan Experiments

Two LCMS experiments are conducted to maximise data collection whilst minimising machine time and data processing requirements. This is achieved by avoiding redundant data acquisition across samples by reducing the number of tandem mass spectrometry (MS^2) experiments performed (Figure 2-2 LCMS Experimental Outline for Full Scan and Deep Scan Experiments.).

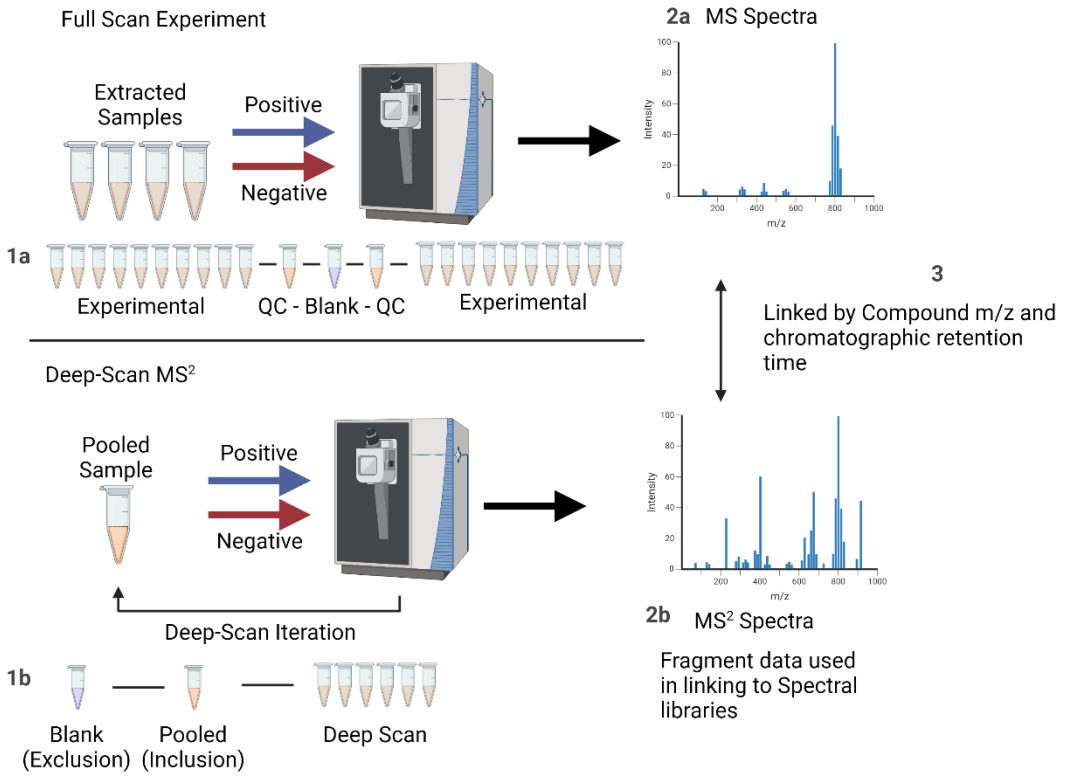

Figure 2-2 LCMS Experimental Outline for Full Scan and Deep Scan Experiments.. This figure outlines key steps in the MS analysis of extracted samples. The injection order (1a) illustrates how quality control (QC) samples and blanks for normalisation and carry-over checks are integrated. For Deep-Scan analysis, the injection process (1b) facilitates the creation of exclusion and inclusion lists for iterative MS^2 generation in the following steps. The output of the MS data (2a) can be linked to MS^2 spectra (2b) from the Deep-Scan experiment through the correlation of compound m/z values and retention times.

Experiments are performed in positive and negative ionisation modes, referring to the way in which molecules are ionised and selected before being transferred to the mass spectrometer. The decision to perform both positive and negative experiments was made due to the fact compounds may form ions more effectively in one mode over the other, influencing sensitivity of detection. Compounds may have a greater propensity to gain or lose protons ($[M+H]^+$ or $[M-H]^-$), or form adducts with other ions within the sample such as chloride (Cl^-) or sodium (Na^+) ions (Figure 2-3 Examples of some common positive and negative molecule (M) adducts).

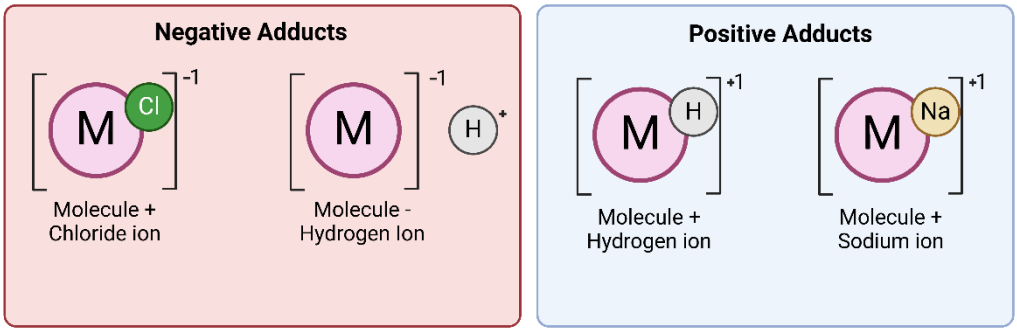

Figure 2-3 Examples of some common positive and negative molecule (M) adducts. Examples shown include protonation ($[M+H]^{+1}$), sodium adduction ($[M+NA]^{+1}$) in positive ionisation mode and deprotonation ($[M-H]^{-1}$) and chloride adduction ($[M+Cl]^{-1}$) in negative mode.

The rationale behind conducting experiments in both positive and negative ionisation modes to provide more comprehensive characterisation of the intestinal microbiome.

2.3.2.1 Full Scan

A Full-Scan experiment (Figure 2-4 MS processes during Full Scan LCMS experiments) aims for comprehensive metabolite coverage, capturing all ions across the defined full m/z range. Within the Full-Scan experiment, no ion fragmentation is performed, aiming to provide intact masses for all compounds within the sample. Focusing solely on the collection of single MS data for intact compounds, rather than conducting MS^2 experiments as well, more sample points can be included over the course of a compound's elution. This results in better peak integration and more accurate measurements of abundance.

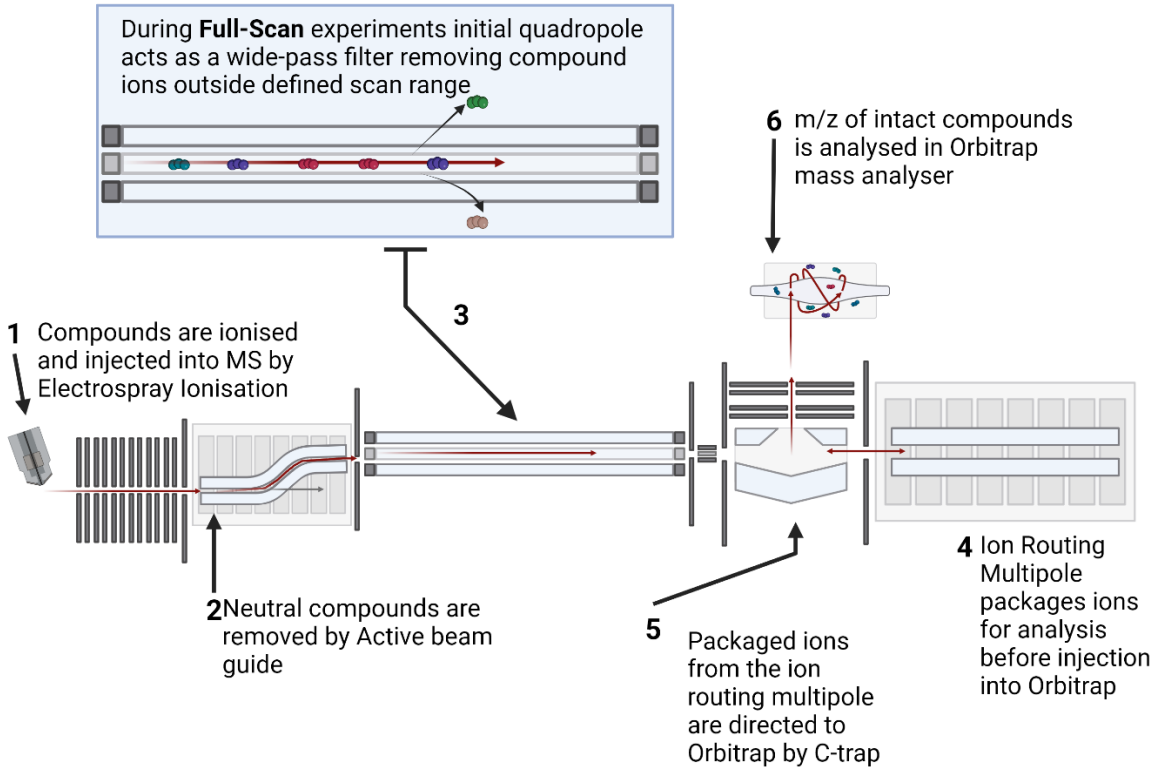

Figure 2-4 MS processes during Full Scan LCMS experiments. (1) Compound ionisation by Electrospray ionisation occurs in positive and negative moves. (2) Neutral compounds are removed from the ion beam; neutral compounds fail to be directed by the active beam guide. (3) Quadrupole filters out compounds which do not fall within the scan range. (4) The Iron routing multipole packages groups of compounds, which are routed to the C-Trap (5) The C-Trap injects intact compounds into the orbitrap (6) Compounds are analysed by the Orbitrap mass analyser.

During the Full-Scan experiment, samples were analysed with triplicate injections, and the sample order was randomised across the run. Quality control injections taken from the pooled samples serve as reference points for

normalisation throughout the run, as readings may fluctuate over time as conditions with the mass spectrometer change. Blanks are injected regularly and act as points to identify the carryover of analytes between chromatographic runs. Following blank injects, an additional QC injection is advised to maintain column conditions between samples. The rate at which QC and blank injections are performed is every 5-10 samples, depending on the length of the run, to balance data quality with machine run time and availability.

2.3.2.2 Deep Scan

Deep-Scan experiment (Figure 2-5 MS processes during Targeted Deep Scan LCMS experiments) is a targeted scan focusing on compounds within a list of pre-determined m/z values by iteratively working through a list of all compounds identified within the sample set from a pooled sample. The most abundant compounds are identified and added to a target list for further analysis. Collision-induced fragmentation (CID) breaks compounds into fragment ions. These fragment ions provide valuable information on compound composition and structure, used to differentiate compounds of similar mass and composition. Each iteration of the Deep-Scan experiment works down the list of compounds, digging deeper into the data set and providing fragmentation data for more compounds within the dataset.

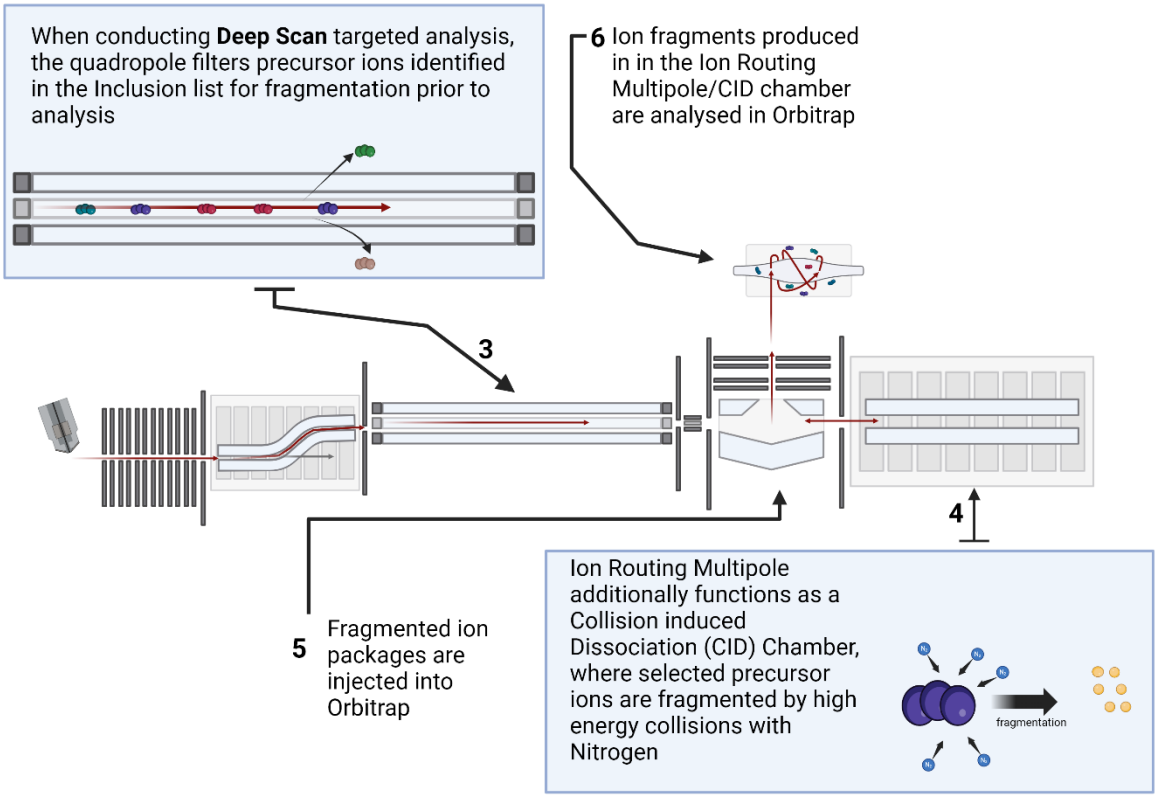

Figure 2-5 MS processes during Targeted Deep Scan LCMS experiments. (1) Compound ionisation by Electrospray ionisation occurs in positive and negative moves. (2) Neutral compounds are removed from the ion beam; neutral compounds fail to be directed by the active beam guide. (3) During Deep-Scan experiments, the quadrupole acts as a narrow filter, selecting compounds on the current iterations inclusion list and allowing them to pass to the Ion routing Multipole/CID Chamber (4) These precursor ions which arrive in the CID chamber, undergo fragmentation via high energy impacts with nitrogen, fragment ions are directed to the (5) C-Trap before being injected into the (6)Orbitrap for analysis.

Deep Scan experiments are conducted on pooled samples, which theoretically contain all metabolites within the experiment set. A blank sample is injected and analysed with a full scan experiment to generate an initial exclusion list of background compounds that will not be selected as precursor ions for MS2 experiments in the following injections. An initial full scan experiment on the pooled sample generates a list, which, in theory, contains all compounds within the sample set. This compound list generates target lists of precursor ions for MS2 experiments. The Deep-Scan workflow iterates through the compound list, selecting the analysis of the most abundant compounds before adding them to

an exclusion list for the next iteration. This prevents the continual reanalysis of the same compounds across iterations. The appropriate number of iterations necessary for comprehensive metabolite identification is discussed in 2.7 LCMS Deep Scan Experiments: Determining Optimal Coverage.

2.4 Sample Processing

In the development of a metabolite extraction pipeline, emphasis was placed on creating an optimised methodology for both faecal samples and *in vitro* model samples (Figure 2-6 Outline of Key Steps in the Extraction of Metabolites from Faecal and *In Vitro* Samples).

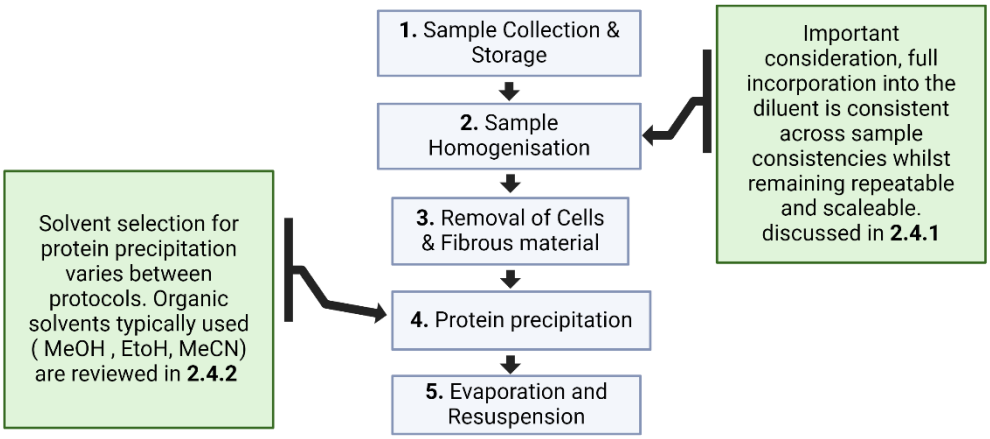

Figure 2-6 Outline of Key Steps in the Extraction of Metabolites from Faecal and *In Vitro* Samples.

In the development of a metabolite extraction pipeline, particular emphasis was placed on initial homogenisation of faecal specimens covered in 2.4.3 Comparative Analysis of Faecal Sample Processing Methods: Bead-Beating vs. Manual Extraction.

Faecal samples, owing to their heterogeneous composition, present distinctive challenges. Variations in diet, health status, age, medication history, and underlying gastrointestinal issues contribute to the challenging nature of establishing a standardised protocol for metabolite extraction suitable for a range of sample consistencies. Disruption and emulsification of faeces into a diluent such as phosphate-buffered saline (PBS) phosphate-buffered saline is necessary to aid in the removal of non-soluble fibrous material and remove host

and microbial cells. Incorporating solid faecal samples into a liquid state is integral to the centrifugation and filtration steps that facilitate this.

Due to the aqueous environment of the *in vitro* models, samples from the MiGut or Leeds Human Gut Model-(HGM) are readily processed without the need to be incorporated into a diluent. *In vitro* model samples can be analysed without any preprocessing steps, and following mixing by agitation samples can be processed immediately via centrifugation and filtration.

Extraction solvent choice was also a concern for developing a unified extraction process for faeces and *in vitro* samples. Several solvents are commonly used for metabolite extraction for LCMS analysis, the potential of which is discussed in 2.4.4 Comparative Evaluation of Extraction Solvents for Faecal and *In-vitro* model Samples

2.4.1 Ethical approval

The collection and use of human faeces in the model has been approved by the School of Medicine Research Ethics Committee, University of Leeds (MREC-15-070-Investigation of the interplay between Commensal Intestinal Organisms and Pathogenic Bacteria).

2.4.2 Solvent-Based Metabolite Extraction

Samples were all processed with a solvent-based extraction Figure 2-7 Overview for Metabolite Extraction for Faecal and *In Vitro* Gut Model Samples. Faecal samples underwent an initial homogenisation step to improve the efficacy to centrifugation and filtration in the follow steps and ensure uniform distribution of metabolites within the sample. Faecal sample homogenisation methodology is discussed in 2.4.3 Comparative Analysis of Faecal Sample Processing Methods: Bead-Beating vs. Manual Extraction.

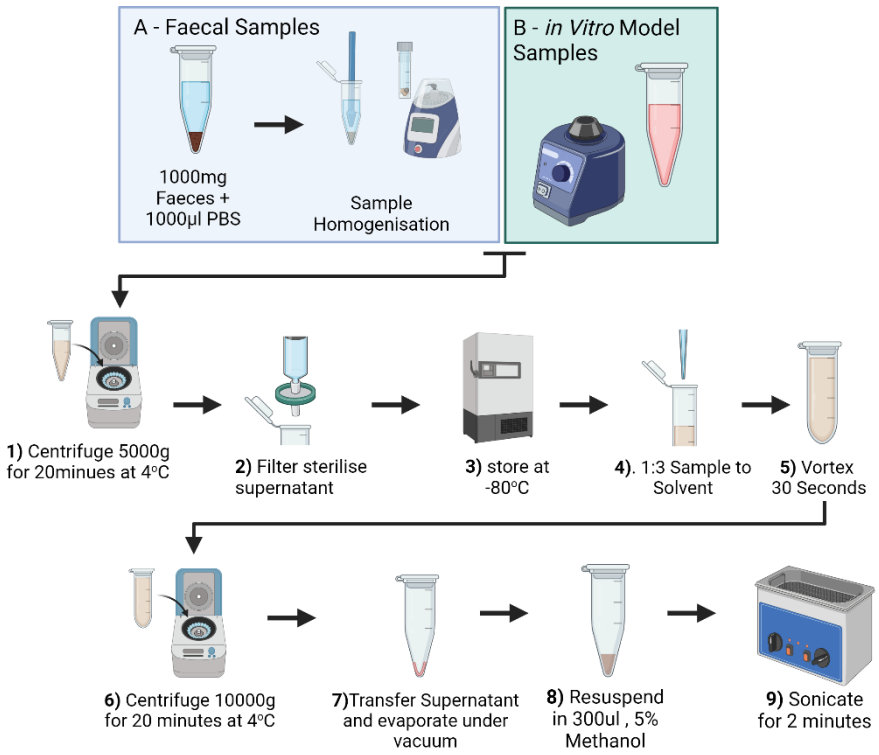

Figure 2-7 Overview for Metabolite Extraction for Faecal and *In Vitro* Gut Model Samples. (A) Faecal Samples undergo an initial homogenisation step to aid in extraction of metabolites (B) in contrast *in vitro* samples are aqueous in nature and more homogenous, therefore only require vortexing to ensure metabolites are uniformly dispersed within the sample prior to extraction

To remove bacterial and host cells prior to storage and extraction, samples are centrifuged at 5000g for 20 minutes and filter sterilised through a 0.22um polyether sulfone (PES) syringe filter. The Supernatants are then presumed to be sterile, free from non-cellular material, bacterial and host cells, and suitable

for long-term storage at -80°C, generally recommended for metabolomic samples (Smith et al., 2020; Valo et al., 2022).

If stored at -80°C prior to extraction, samples should be thoroughly defrosted and vortexed before beginning the solvent-based metabolite extraction. Samples are diluted in a 1:3 ratio of sample to solvent (Methanol (MeOH)) and vortexed for 30 seconds. They are then centrifuged at 10000g for 20 minutes at 4°C to remove precipitate, which should contain proteins and other macromolecules. The supernatant containing the metabolites of interest is transferred to a new Eppendorf tube and evaporated under a vacuum until dry. The dry material can then be resuspended in the original sample volume using 5% methanol. The sample should subsequently be vortexed and sonicated to aid in dissolution.

2.4.3 Comparative Analysis of Faecal Sample Processing Methods: Bead-Beating vs. Manual Extraction

The homogenisation process for faecal samples combines raw faecal samples in a 1:1 ratio with a diluent to create a faecal slurry. This is an important step as it facilitates the removal of cells and other non-soluble material from the aqueous phase. PBS is used instead of organic solvents, which have the potential to disrupt filter membranes used to remove bacterial cells and non-soluble material.

Faecal slurries are the processed in the same manner is *in vitro* samples (Figure 2-7 Overview for Metabolite Extraction for Faecal and *In Vitro* Gut Model Samples.)

Faecal material falls under the Human Tissue Act (UK Government 2004) and may contain pathogens hazardous to human health; therefore, removing bacterial and host cells before storage renders the samples safe to handle and compliant with regulations. Ensuring that sample processing aligned with compliance with regulations and ensuring sterility for processing in non-biosafety labs were both considered in protocol development.

Six faecal samples obtained from healthy donors were subjected to analysis in this study. Each sample underwent both processing methods: bead-beating and manual agitation. Following homogenisation, the samples were extracted and analysed in triplicate via LCMS.

To evaluate the effectiveness of the homogenisation procedures, the abundance of identified compounds was compared across 3 replicates for each extraction method. Variations in metabolite extraction were identified by calculating the coefficient of variation (CV) for each compound signal within the replicate. The uniformity of compound extraction was used to determine the consistency of homogenisation between bead-beating and manual processing techniques.

Manual Preparation:

Manual agitation, similar to grinding with a pestle followed by vortex mixing serves as an accessible method for sample homogenisation. The hands-on nature allows for case-by-case tailoring to samples of a heterogeneous nature, albeit with potential inconsistencies, time constraints, and scalability issues inherent in manual sample preparations. Initial manual agitation is necessary as vortexing alone is insufficient for the full incorporation of faecal samples into the diluent. This method becomes inefficient when working with larger sample sets, necessitating a more standardised and time-efficient homogenisation approach for faecal samples.

PBS was added to faecal samples in a 1:1 ratio. Samples were initially disrupted using a sterile loop inverted and used in a "pestle-like" manner to facilitate homogenisation. Following disruption, samples were vortexed for 2 minutes to further enhance mixing and solubilisation of metabolites.

Bead Beater Preparation:

Bead-beating examined as a scalable alternative to manual mixing and vortexing in sample preparation. Bead-beating offers shorter processing times, scalability with increased sample numbers, and enhanced control and repeatability. By comparing bead-beating with the baseline method, the aim was to assess whether bead-beating can match the sample consistency achieved with manual preparation. The experimental design was to evaluate manual homogenisation alongside bead-beating, for consistency w processing efficiency and scalability.

PBS was added to faecal samples in a 1:1 ratio. A sterile and acid-washed 6mm glass bead was added to each tube to aid in disruption. Homogenisation was performed using a FastPrep-24 5G homogeniser set to a speed of 4 m/s,

with 2 cycles of 20 seconds, interspersed with a 120-second pause between cycles. The selection of bead size and homogenisation cycle parameters was made to optimise emulsification of the sample while minimising cell damage and processing time.

Following homogenisation, samples processed by both methods were sub-aliquoted into three replicates. Metabolites were then extracted using the method described in 2.4.2 Solvent-Based Metabolite from each sub-sample and the relative abundance of compounds was assessed via LCMS

Compound relative abundance variation Analysis.

The relative abundance of compounds was measured, and the coefficient of variation (CV) of compound abundance between replicates of processing method recorded (Figure 2-8 Comparison of Coefficient Variability in Bead Beaten and manually processed Samples: Positive vs. Negative Ionization). The coefficient of variation provides a normalised metric of dispersion for readings within each set of replicates. A lower CV indicates that datapoints within the replicates are closer to the mean. Compound CV was explored in both positive and negative ionisation modes to ensure that impacts were consistent across compounds recovered across both ionisation modes.

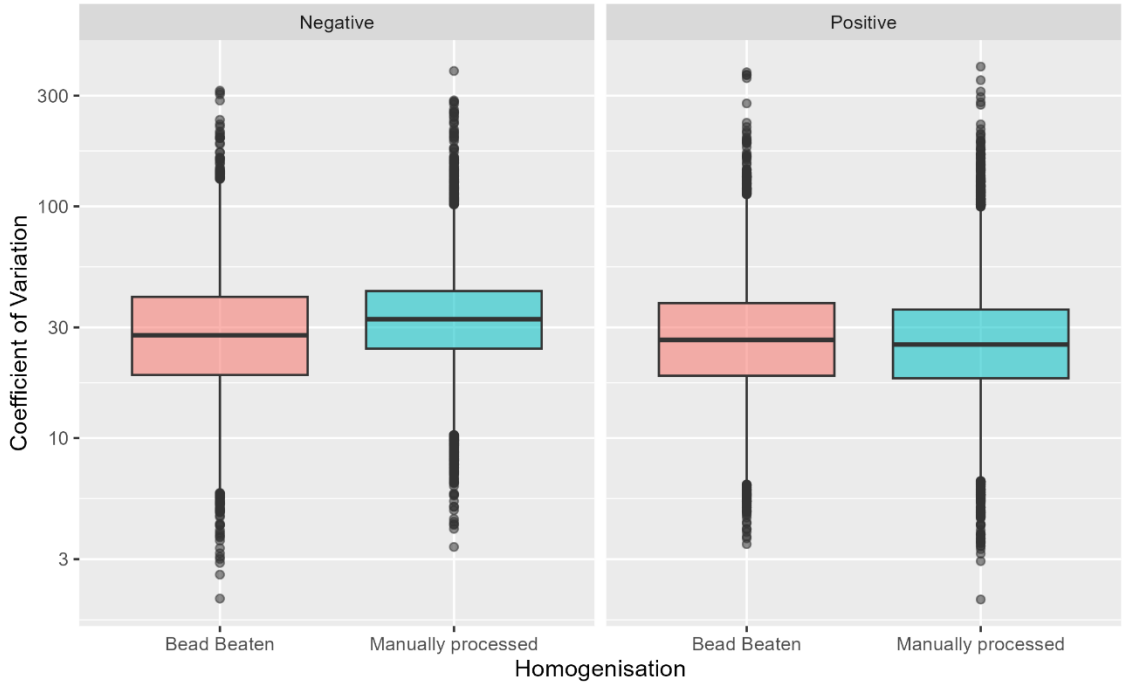

Figure 2-8 Comparison of Coefficient Variability in Bead Beaten and manually processed Samples: Positive vs. Negative Ionization

The combined mean for All Bead-beaten samples was 32.47 and manually extracted 33.50. The lower mean CV value implies there is less relative variability within the bead-beaten samples, suggest greater consistency. A welch Two Sample t-test was conducted using the `t.test()` function in R to compare the means of Bead Beaten and Manually processed groups. The small but statistically relevant difference in mean supported by a p-value of 5.785e-11 suggests that bead-beating samples as a method of sample homogenisation, is a viable and potentially preferable method compared to manual extraction.

2.4.4 Comparative Evaluation of Extraction Solvents for Faecal and In-vitro model Samples

There are a variety of extraction solvents used for metabolite extraction when processing samples for LCMS. Methanol (MeOH) Ethanol (EtOH) Acetonitrile (MeCN) are commonly used in extraction protocols for reverse phase chromatography analysis applications. During the development of an extraction protocol, it was deemed appropriate to identify if there were significant differences in the metabolome extracted from faecal and *in vitro* samples between solvents. Solvent choice may impact metabolite extraction efficiency, as different solvents could potentially extract different compound classes. Three

commonly used solvents (MeOH, EtoH and MeCN), a 1:1 mixture of Methanol: Ethanol (MeEt) as well as PBS for evaluation.

Pooled faecal sample from 4 healthy volunteers, processed with the bead beating method described previously and pooled *in vitro* material from the distal portion of a healthy Leeds Human Gut model during steady state were extracted with each solvent.

Untargeted LCMS analysis was conducted in both positive and negative ionization modes, resulting in the identification of 9393 and 9076 compounds across all samples, respectively.
Principal Coordinates Analysis (PCOA) was employed to visualize the extracted metabolomes, where sample types tended to cluster closely together rather than by solvent type (Figure 2-9 PCoA Assessment of Solvent Effects on Metabolite Profiles).

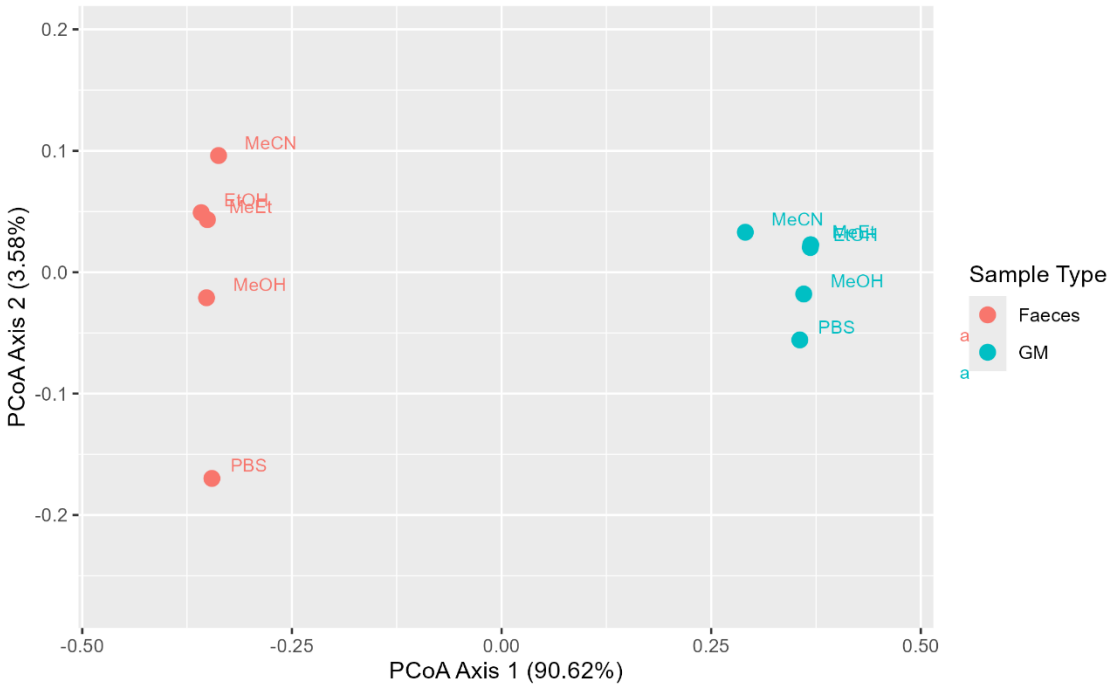

Figure 2-9 PCoA Assessment of Solvent Effects on Metabolite Profiles

Variations between the extraction solvents were more pronounced in faecal samples than *In Vitro* Model samples, potentially attributed to their increased complexity and diversity of compounds within starting material.

The three organic solvents and solvent mixture (MeOH, EtOH, MeCN, MeEt) display comparable clustering in both Faecal and *In vitro* extractions.

With this in mind and in the context of untargeted LCMS analysis where the focus is on capturing gross relative changes in the metabolome between Healthy and dysbiotic states, extraction solvent choice may have a diminished impact compared to targeted studies.

Considering the development of targeted studies in the future, with a focus on absolute quantification and heightened sensitivity, profiling how extraction solvents interact with specific compound classes of metabolites may potentially be advantageous going forward.

2.5 LCMS Parameters

High-performance liquid chromatography (HPLC) was coupled with mass spectrometry (MS) to resolve and analyse compounds. This coupling enables the analysis of complex samples by reducing the complexity of individual sample injections over time and adding another dimension of data through chromatographic retention times.

2.5.1 High-Performance Liquid Chromatography

High-performance liquid chromatography (HPLC) is characterised by its high working pressure and small particle size of the stationary phase. The smaller bead size used in HPLC offers a greater surface area for increased interactions between the analytes and stationary phase of the column, leading to increased separation efficiency, improving the resolution compounds eluting over similar time frames. The increased operational pressure of HPLC allows for enhanced separation speed and efficiency which would not be possible with traditional liquid chromatography.

Liquid Chromatography is often described as being Normal Phase or Reverse Phase. This distinction describes the polarity of the stationary (column) and mobile phases (solvent).

Normal phase chromatography has a polar stationary phase and non-polar mobile phase. The non-polar solvents used in normal phase are less suited for the analysis of polar aqueous samples typical in many areas of biological research (Jandera, 2013), alongside the environmental and safety concerns

inherent when working with solvents such as hexane or chloroform means that normal phase chromatography does not see as widespread use reverse phase.

In contrast to normal phase chromatography, reverse phase utilises a non-polar stationary phase and a polar mobile phase (Figure 2-10 Reverse Phase Chromatography, an overview.). The solvents used in reverse phase such as water, methanol and acetonitrile pose fewer safety concerns than those common in normal phase. This configuration is better suited to aqueous sample types, making reverse phase the more commonly employed methodology in biological and clinical research.

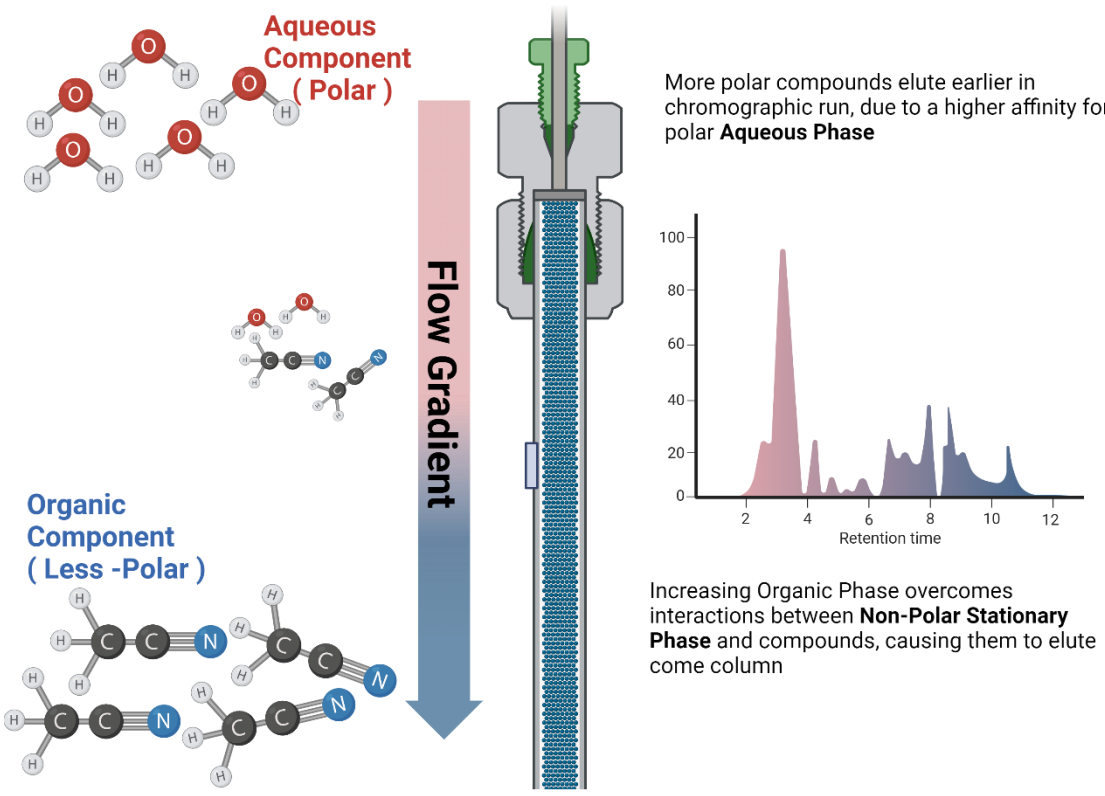

Figure 2-10 Reverse Phase Chromatography, an overview. The mobile phase consists of two components, Aqueous and Organic, which are relatively more and less polar, respectively. The proportions of each component change over time.

2.5.1.1 C18 Chemsitry

Sample analysis in this thesis is performed using C18 reverse phase columns. C18 column chemistry is defined by its octadecyl $(CH_2)_{17}$ chains bound to a silica support base (Figure 2-11 C18 Column Chain Chemistry). These long hydrocarbon chains provide the hydrophobic interaction points within the column, upon which separation is achieved. These Hydrophobic interactions cause non-polar compounds to elute later than the chromatographic run than polar compounds.

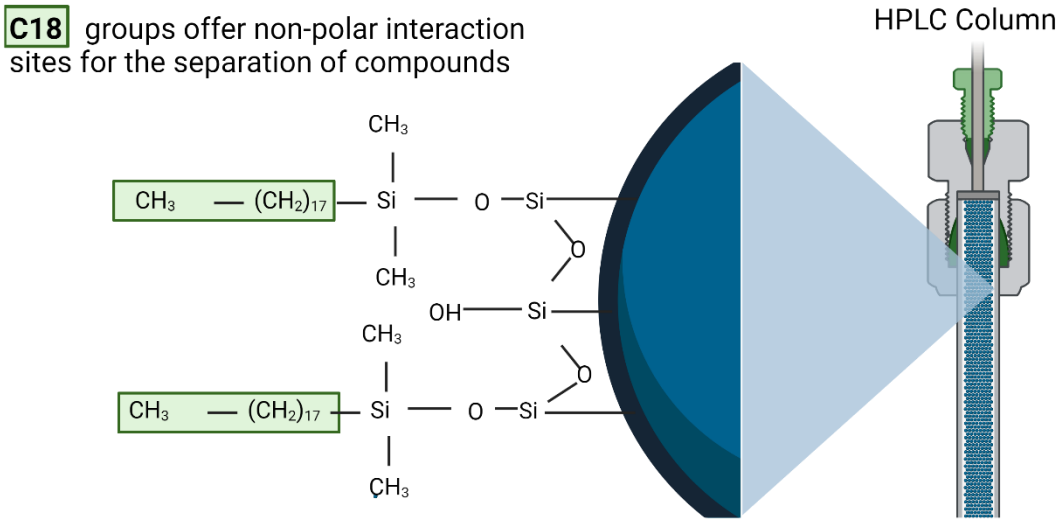

Figure 2-11 C18 Column Chain Chemistry . Silica Beads with Octadecyl Chain attachment

HPLC Column

C18 columns were utilised for chromatographic separation. C18 columns were chosen due to their versatility in analysing biological samples. Hypersil GOLD C18 selectivity Reversed Phase columns (2.1mm x 150mm) with corresponding guard columns (2.1mm x10mm) (Thermo Fisher Scientific) were used across all experiments.

Solvents

The aqueous phase (solvent A) consisted of high-performance liquid chromatography-grade water (Romia) with 0.1% v/v formic acid, while the organic phase (solvent B) comprised acetonitrile (Romia) supplemented with 0.1% v/v formic acid. Formic acid is added to the solvent to aid peak formation by enhancing peak shape and reducing chromatographic tailing (Souihi et al., 2023).

Chromatographic Gradient

A 15-minute chromatographic gradient was utilised to balance runtime and sample separation effectiveness. Flow rate was 0.3 mL/min. Data collection occurred between 0.5 minutes and 10 minutes of each sample run. The chromatographic gradient begins with a composition of 99% aqueous phase and 1% organic phase. Between 0.5 minutes and 5 minutes, the proportion of the organic phase increased to 60%; in minutes 5 to 6, the organic phase was further increased to 95%, which is maintained until minute 10 (Figure 2-12 Organic Phase (Solvent B) proportion in HPLC Chromatography).

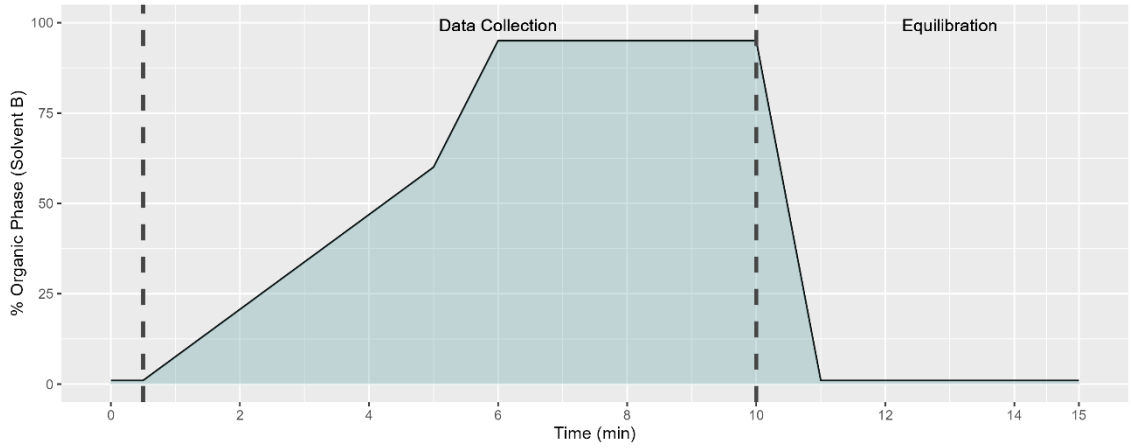

Figure 2-12 Organic Phase (Solvent B) proportion in HPLC Chromatography

After the data collection portion of the run, an equilibration step was performed to prime the column for the next sample. During the equilibrium phases (before 0.5 minutes and after 10 minutes), the column flow was diverted to waste to prevent sample build-up within the ionisation source and preserve the stability of LCMS across extended analysis runs. During the equilibration step, the organic phase component is reduced to 1% and maintained for 5 minutes.

The gradient was validated using pooled faecal material from healthy donors, chosen for its anticipated complexity and diverse compound composition. 5µl sample volumes were injected in triplicate and resolved according to the specified gradient. The purpose of this validation was to ensure there was minimal loss of metabolites during the equilibrium phase (Figure 2-13 Histogram

of Compound Retention Times (minutes) of Pooled Faecal Samples). Retention times for both positive and negative injections were recorded, with

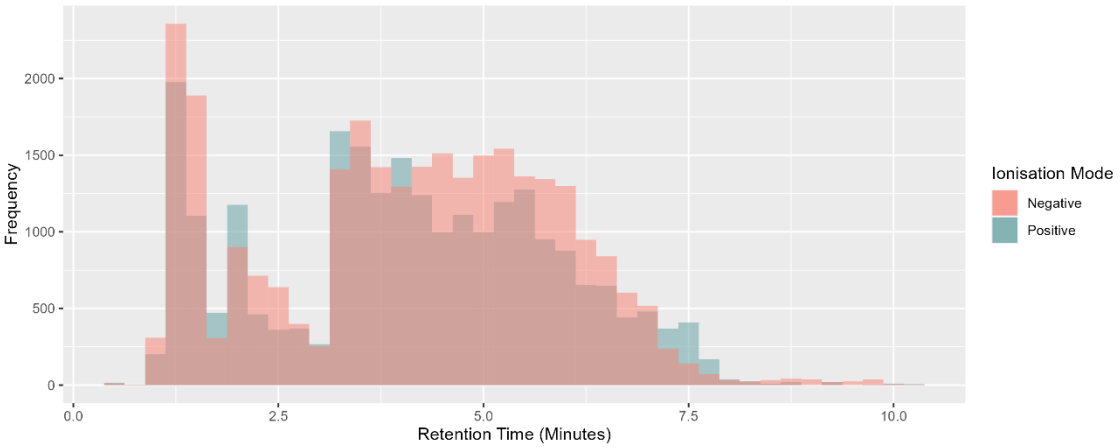

Figure 2-13 Histogram of Compound Retention Times (minutes) of Pooled Faecal Samples

of approximately 95% of compounds in positive and negative mode eluting at retention times of 7.0 minutes and 6.8 minutes, respectively. This assured that there would be minimal loss of recoverable compounds during the equilibrium phase.

2.5.2 Mass Spectrometry

Mass spectrometry analysis was conducted using an Exploris 240 Orbitrap system (Thermo Fisher Scientific). Heated Electrospray Ionisation (ESI) was used as the ionisation source. The positive ionisation mode was set at a voltage of 3.5 kV, while the negative ionisation mode was maintained at 2.5 kV. A lower negative voltage helps to reduce arching within the ion source. The vaporiser temperature was set to 320^{o}C. All data acquisition was performed with a scan range of 80 – 800 mass-to-charge ratio (m/z). For Full Scan experiments, a resolution of 120,000 was used. MS^2 experiments utilised stepped collision energy at collision energies of 30, 50, and 150 and a resolution of 15,000. Runs were managed with Xcalibur (ver 4.3) software (Thermo Fisher Scientific)

2.6 Data Processing and Analysis

Metabolomic data was processed using Compound Discoverer v3.3, a streamlined platform with integrated tools for post-processing, compound detection, and identification. It supports both untargeted and targeted metabolomics workflows. The workflow used in this study was adapted from a default metabolomic workflow to suit the sample type and study requirements (Figure 2-14 Compound Discoverer Untargeted Metabolomics data processing workflow. Key data processing steps and their significant Parameters are detailed below

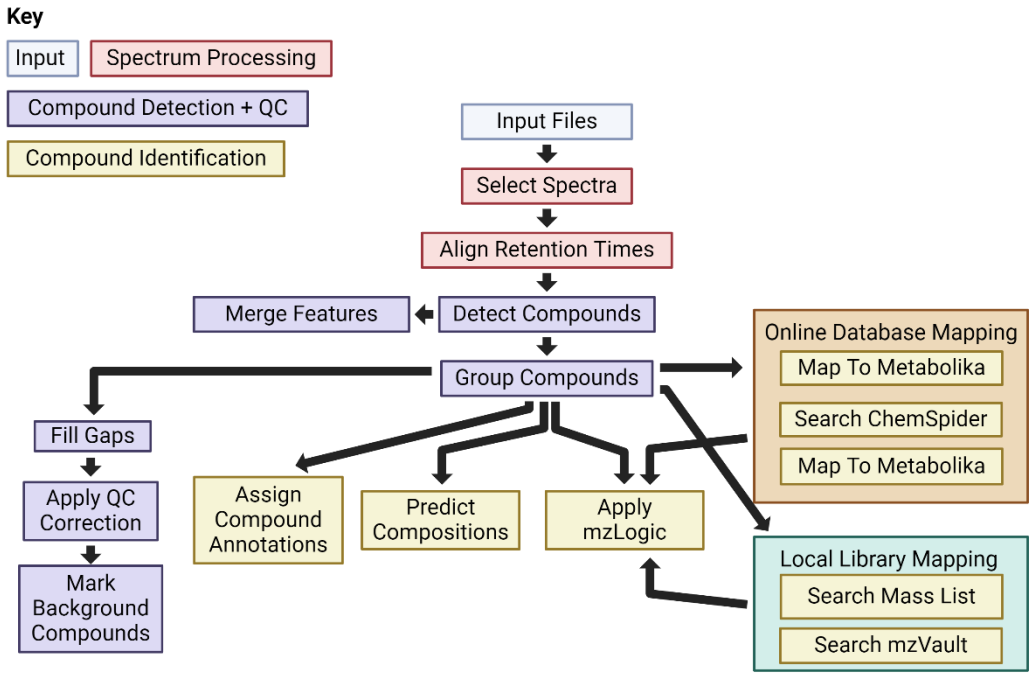

Figure 2-14 Compound Discoverer Untargeted Metabolomics data processing workflow.

2.6.1 Input

Raw files for processing are imported as Xcalibur RAW files from the data acquisition PC. They are categorised as ‘Samples,’ ‘QC,’ ‘Identification Only,’ or ‘Blank’ for Compound Discoverer’s use in subsequent processing steps.

2.6.2 Spectrum Processing

The Align Retention Times node adjusts for small retention time variances among compounds in sample runs. It employs the Chromalign algorithm (Sadygov et al., 2006) for the alignment of chromatographic and Mass spec data.

2.6.3 Compound Detection and QC

Detect Compounds filters out low quality extracted ion chromatograms (XIC) traces, default parameters were used; mass Tolerance 5ppm, minimum peak intensity 10000, and minimum 5 scans per peak.

Merge Features combines compounds identified by Detect compounds via m/z values and chromatographic retention time. Default parameters used; mass tolerance 5ppm, retention time (RT) tolerance 0.05 (min).

Group Compounds, groups compounds together based on similar characteristics with default parameters used; mass tolerance 5ppm, RT tolerance 0.2 (min), Preferred Ions [M+H] and [M-H] and most common ion selected as the area for integration.

Peak rating, a measurement of how “good” the peak is; contributed to by Area, Coefficient of Variation, full width at half maximum, Jaggedness (changes in peak direction), Modality (size of deepest valley), and “Zig-Zag Index” (mean of all valleys within a peak) Together these scores off insight into the overall shape and quality of the chromatographic peak. Base on scoring from the Peak Rating Filtering can be done during processing, where a minimum score “4” was selected across 3 files (all samples injected in triplicate so filtering out compounds which are not detected across at least 3 being filtered as “noise”

Fill Gaps an interpolation node which calculates the area of missing chromatographic peaks.

Apply QC Corrections Uses a Linear regression model to normalise compound abundances based on QC samples throughout the experiment.

Mark Background Compounds highlights compounds identified in sample blanks.

2.6.4 Compound Identification

Assign Compound Annotation sets the order in which annotations are applied. Ordering was curated to apply the highest quality compound annotations where possible.

1. mzVault - This was chosen to be top annotation as it contains locally generated data RT and Spectral Vault data. The combination of RT and special data enhances reliability of compound annotations.

2. mzCloud - mzCloud contains spectral data for compounds and therefore provides excellent fingerprint mapping of compounds detected to known spectra providing accurate compound detection.

3. Predicted Compositions - suggest likely compound formula based on mass spectrometry data.

4. Mass List Search - despite providing limited information, Mass list searches can be used to identify compounds have been run locally but lack MS2 spectral information. The inclusion of retention time data alongside compound masses provides an additional dimension to compound identification.

5. ChemSpider Search - Offers broad compound coverage, the reliance solely on compound formula or mass-based searches limits the specificity of compound annotation. It may, however, serve as a useful tool when other searches fail to provide potential identifications.

2.7 LCMS Deep Scan Experiments: Determining Optimal Coverage

The Deep Scan experiment relies on data-dependent acquisition (DDA) of MS^2 spectral data. In the outlined untargeted workflow, it is responsible for the generation of all necessary spectral data used for compound identification via spectral libraries. The deep scan analysis begins with the generation of a list of putative precursor ions to be selected for fragmentation and analysis. Experimental cycles then select the most abundant compounds, moving down the list with each successive cycle probing deeper into the dataset. This method requires the balancing of dataset coverage against machine time, prompting an evaluation of the cycles required for effective coverage of the intestinal metabolome.

A pooled faecal sample, derived from samples obtained from 4 healthy volunteers, underwent analysis. Faecal samples were selected due to their anticipated diversity and complex of compounds, making them a suitable medium for the evaluation of experimental performance and efficacy.

A data-dependent acquisition experiment was carried out, comprising 10 iterations of both positive and negative ionisation. Post acquisition, the data-dependent acquisition for preferred ions ([M+H]+1 and [M-H]-1) for each iteration was evaluated.

MS^2 coverage was broadly the same for positive and negative ionisation modes whilst being consistently lower for negative ions. This is likely attributed to the increased size of the negative dataset (Negative 95281 compounds vs Positive 80980 compounds) Figure 2-15 MS2 Spectra Distribution: Preferred Ion Coverage vs DDA Iteration, by Ionization Mode

Mean coverage for both ionisation modes was approximately 37% for the first iteration, 54% for the second, 64% for the first and 69% for the fourth. The fourth round of DDA marked the point of diminishing returns for generating new spectral data within the experiment.

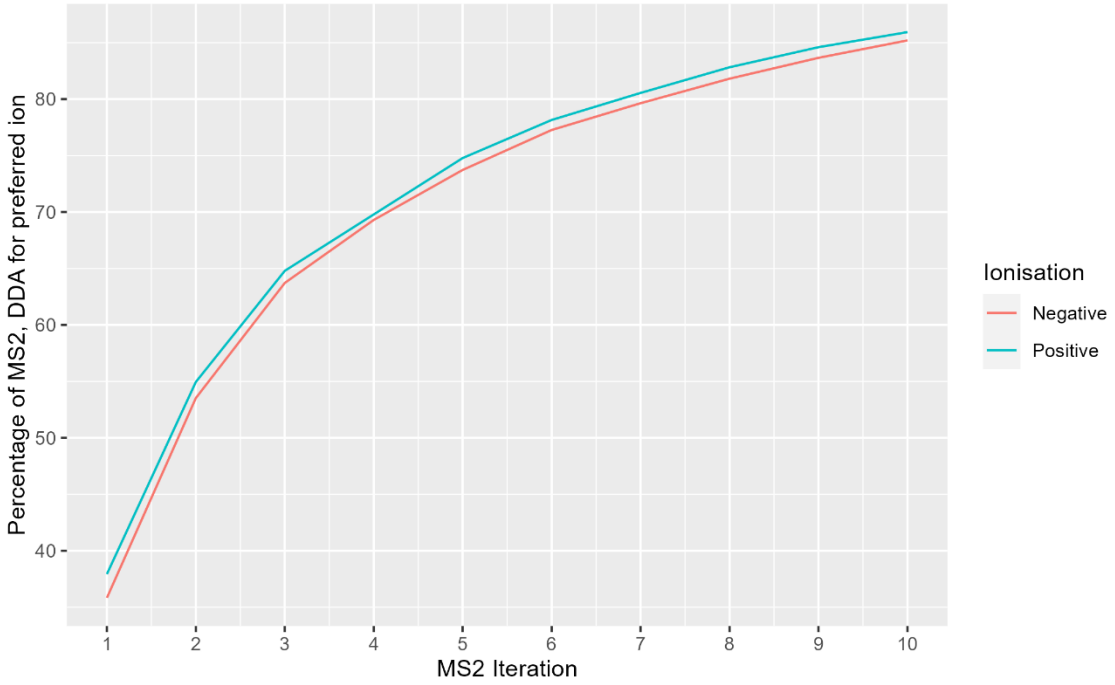

Figure 2-15 MS2 Spectra Distribution: Preferred Ion Coverage vs DDA Iteration, by Ionization Mode

MS^2 spectra are in matching compounds against spectral libraries, mzCloud and mzVault. Beyond the absolute coverage of the dataset, it is important to interpret this data within the context of how successive DDA iterations influence annotations, as these affect the practical utility of the data.

Assessing “Full Match” identifications against compounds within the mzVault and mzCloud, show a similar trend to total MS^2 coverage (Figure 2-16 Spectral Library Annotations vs. DDA Iteration: Negative Ionization).

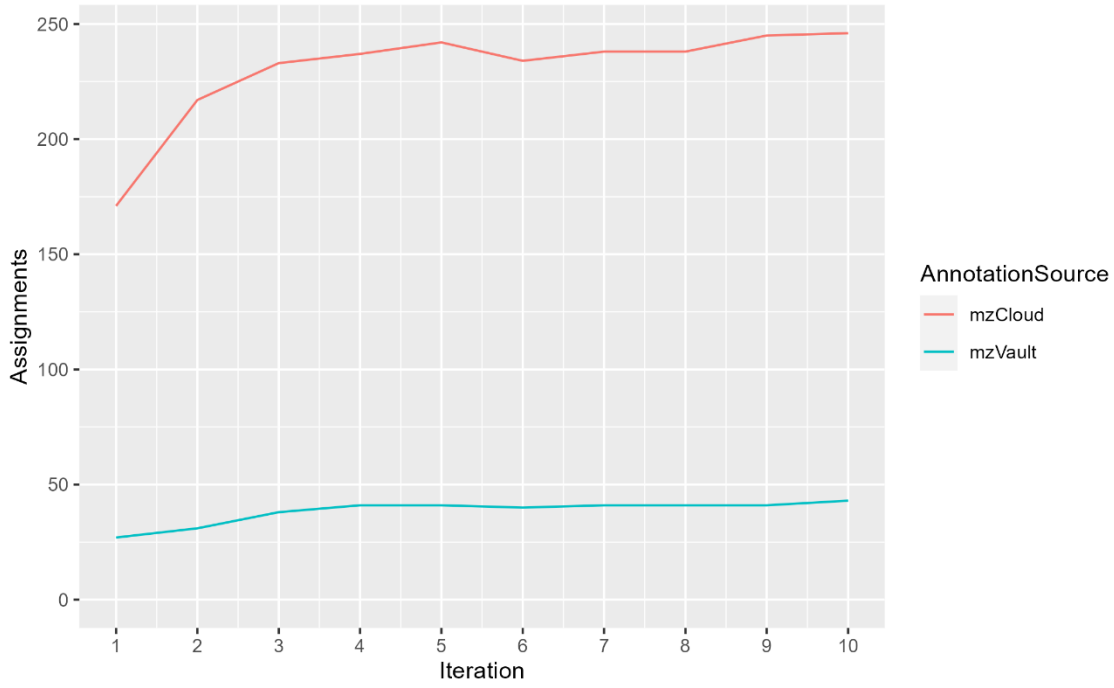

Figure 2-16 Spectral Library Annotations vs. DDA Iteration: Negative Ionization

For both total MS2 coverage of the data and assigned compound annotations, 4 DDA iterations seem sufficient; it is important to note, however, that the utility of additional DDA iterations may change in the future as local and online databases expand.

2.8 Local Library Development – Processing and Annotation

For the enhancement of compound annotation, the development of local metabolite libraries was pursued. Two commercially available libraries for mass spectrometry were selected; Microbiome Metabolite Library of Standards (GUTMLS), and Bile Acid Carnitine Sterol Library (BACSMLS) (IROA Technologies, USA). They were chosen for their relevance to the intestinal microbiome, as well as their potential to serve as benchmarks for the analytical workflow. BACSMLS – contains 96 compounds belonging to the classes; bile acids, carnitines, and sterols, and GUTMLS - contains 185 compounds belonging to a diverse range of important metabolic pathways associated with the intestinal microbiome.

The libraries were supplied as lyophilised 96 well plates with each well containing 5µg of compound per well. Compounds were solubilised as per manufacturer specifications using HPLC grade solvents (Methanol, Ethanol, Chloroform) and prepared for multiplex injection. The rows of each plate were pooled by row allowing BACSMLS to be analysed over 8 triplicate injections and GUTMLS 16. Libraries were both run using the MS^2 experiment previously described, due to the simplicity of the analyte being run, DDA was deemed unnecessary. Both Libraries were run in positive and negative ionisations.

MLSDiscovery (IROA Technologies, USA) was Initially used to generate spectral libraries. Despite the successful identification of compounds with the data and generation of a library, Compound discoverer was unable to match compounds with the Data set to online resources but not mzVault, despite being within the defined parameters. It was believed that this was due to a lack of standardisation within the NIST file format used to store LCMS data, resulting in much of the meta-data Compound discoverer utilises in library matching being lost. Despite multiple attempts to make the workflow function with reduced metadata, it was unsuccessful in mapping compounds within the data to mzVault and removal of metadata requirements from local library mapping resulting in the workflow failing entirely. It was decided to rebuild the library entirely. Compound Discoverer was instead used directly for library generation to retain meta-data lost previously. A mass List of all compounds within the library was used as the primary annotation source.

Out of 281 metabolites analysed across two libraries, 170 compounds were reliably identified: 132 from GUTMLS and 38 from BACSMLS. The ionisation mode influenced the detection of compounds. For GUTMLS, 112 were positive and 114 were negative, with 94 identified in both modes. For BACSMLS, 35 were positive and 26 were negative, with 23 identified in both modes. This highlights the importance of analysing samples in both ionisation modes when aiming to fully capture the metabolome during an initially exploratory analysis.

Compound identification was correlated to compound class likely due to shared properties Figure 2-17 Distribution of Compound Classes by Recovery Status in GUTMLS and BACSMLS Libraries. The LCMS protocol was particularly suited to resolving identifying carboxylic acids and derivatives (26/27 compounds Identified) and bile acids (21/33 compounds). It identified fewer Fatty Acyls (24/41 compounds), carnitines (21/33) and sterols (7/31). These findings are like those previously identified for the BACSMLS library, although there are some discrepancies resulting from LCMS methodology, data curation and the method of standard spiking (Pooled vs Individual).

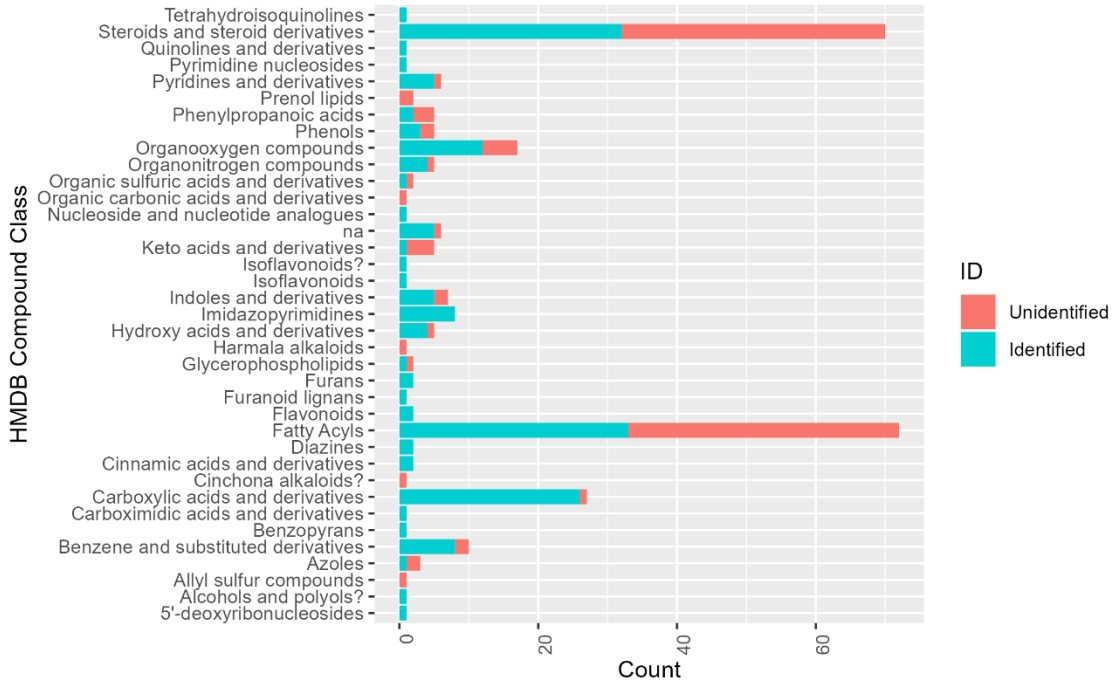

Figure 2-17 Distribution of Compound Classes by Recovery Status in GUTMLS and BACSMLS Libraries

The 287 spectra generated (147 Positive / 140 Negative) were all manually curated during library generation. Sense-checking identified compounds against pooled samples to ensure isomers within the libraries were correctly annotated. MS^2 fragment data was compared to experimental and theoretical spectral data in online resources were possible Human Metabolome Database (HMDB) (Wishart et al., 2022), PubChem (Kim, Sunghwan et al., 2023), mzCloud).

A mzVault library was constructed with the curated data for each library containing MS^2 spectral data, retention time and machine meta-data. The mzVault libraries could potentially be made retention time agnostic in future and could therefore, be utilised across multiple column chemistries using spectral data alone.

2.9 Metabolite Identification Limitations

2.9.1 Impact of Polarity on Compound Recovery

LCMS relies on the physiochemical properties of compounds for effective sample separation via HPLC. The C18 chromatographic method outlined in this thesis utilises hydrophobic interactions between compounds and the stationary phase for sample separation.

Non-polar compounds which have a greater affinity for the C18 stationary phase bind more strongly and consequently elute later in the chromatographic run. The recovery of highly nonpolar compounds may present a challenge as the organic solvent utilised in the mobile phase may lack the elution capability to overcome interaction between these compounds and the stationary phase causing them to remain bound to the column. Consequently, this method of analysis may not identify highly nonpolar compounds. To understand the implications of this limitation on metabolome and assess the role in compound polarity may play in compound recovery and identification within this workflow the metabolite libraries analysed in 2.8 Local Library Development – Processing and Annotation were selected as they consist of known compounds relevant to the area of study; as they contain all known compounds it is possible to identify which are outside the limitations of the analysis method.

XLogP (Wang, R. et al., 1997) is an algorithmic approach used for calculating the partition coefficients, known as LogP values, of compounds in an octanol-

water mixture. LogP values serve as an estimation of the compound's polarity. Higher XLogP values indicate a greater degree of hydrophobicity, while lower values suggest a higher degree of polarity.

The XLogP of call compounds within the two libraries were obtained. Values were pulled from entries in PubChem by matching compound IDs (CID) present in the supplied library meta-data. XLogP was analysed against the identification status within the library (Figure 2-18 GUTMLS and BACSMLS libraries, XLogP vs Molecular Weight).

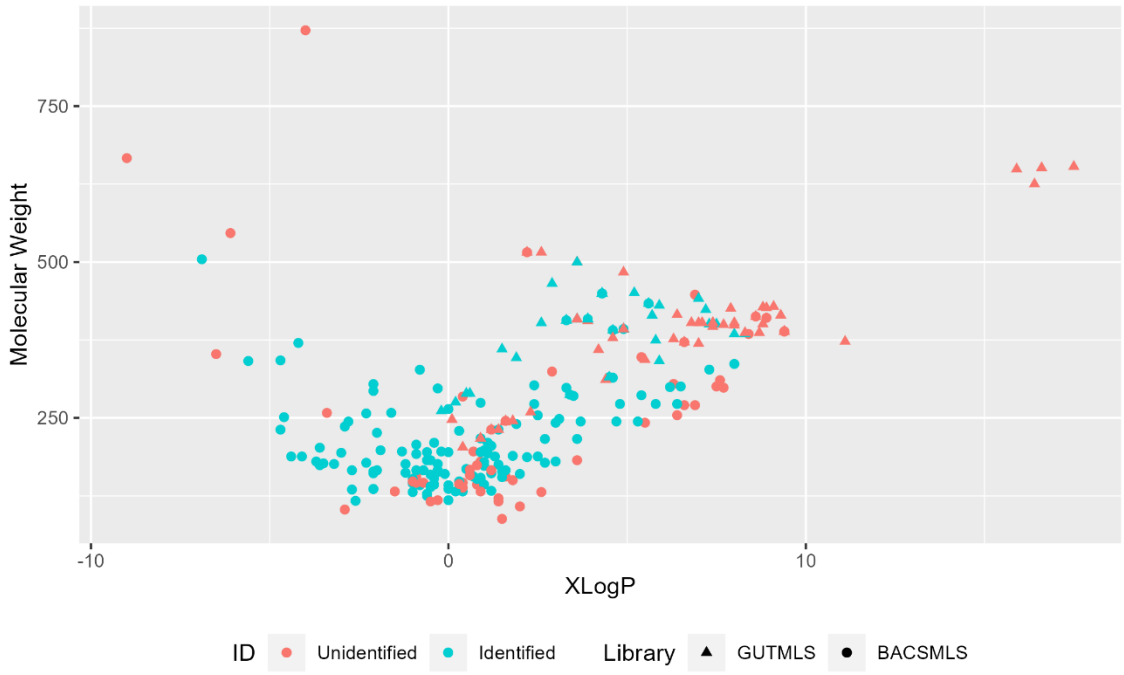

Figure 2-18 GUTMLS and BACSMLS libraries, XLogP vs Molecular Weight GUTMLS (triangles) and BACSMLS (circles) libraries, plotted based on XLogP and Molecular Weight. Identified compounds are marked in blue, while unidentified ones are marked in red.

Higher XLogP values are associated with unidentified compounds, the average XLogP of identified compounds being 0.459 and the average XLogP of unidentified compounds 2.117. Whilst XLogP and, therefore potential compound polarity alone does not entirely describe a compounds recovery potential, as other factors may influence library building such as fragmentation, peak quality, it does appear to be a limiting factor and worth considering in the future if moving towards a targeted/semi-targeted approach.

2.9.2 Compound Classes Associate by Retention Time

Compounds belonging to the same class are expected to share similar physiochemical properties. As shown in 2.8 Local Library Development – Processing and Annotation some compound classes were recovered at lower rates. To identify how compound classes are distributed throughout the chromatography run and identify potential implications this may have for compound recovery the 10 most abundant compound classes (benzene and substituted derivative, carboxylic acids and derivates, fatty acyls, hydroxy acids and derivatives, imidazopyridines, indoles and derivatives, organooxygen compounds, pyridine and derivatives, Steroids steroid derivatives, and unclassified "na") were evaluated for retention time and XLogP (Figure 2-19 Prominent Compound Classes Analysed by XLogP and Retention Time).

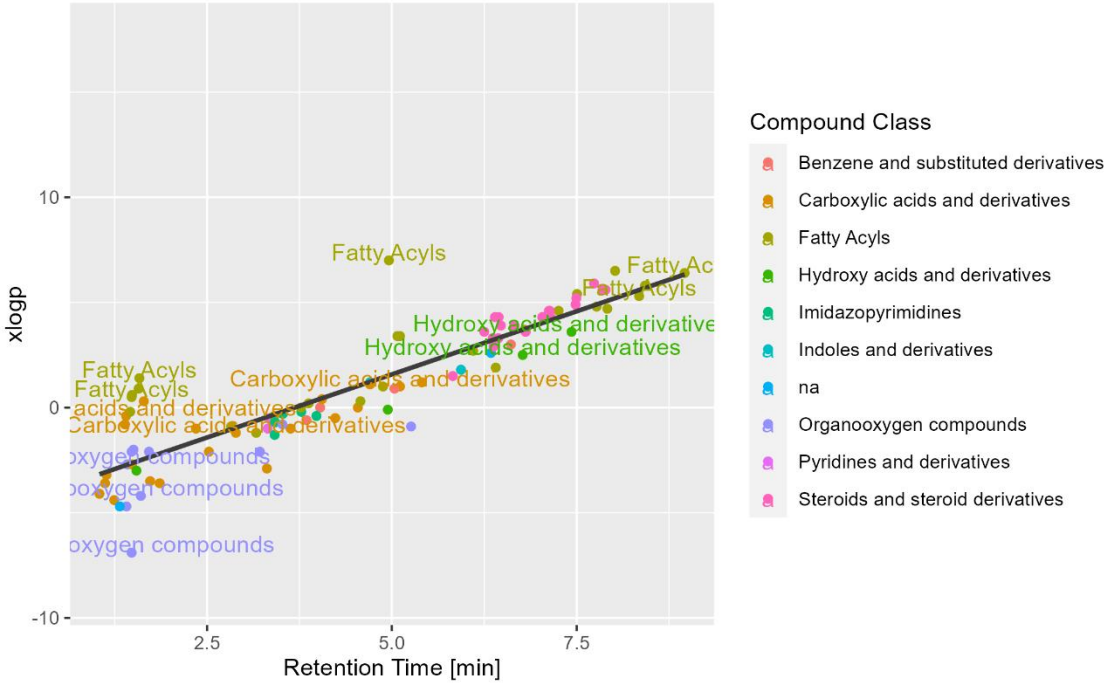

Figure 2-19 Prominent Compound Classes Analysed by XLogP and Retention Time

There is positive correlation between increased retention time and XLogP. Compound classes tend to cluster together when considering XLogP and retention time. Within the metabolite libraries tested, two compound classes, fatty acyls and steroids and steroid derivatives, were identified within the metabolite libraries as being problematic for recovery within this workflow.

These compound classes exhibited higher XLogP values and retention times, suggesting that their poor recovery may be associated with retention within the column. Organooxygen compounds were associated with very low retention times during chromatographic separation, potentially indicating a lack of effective binding to the stationary phase, causing them to elute before compound capture begins.

2.9.3 Compound Annotation Confidence

As shown throughout this chapter, compound Identification is a major challenge in LCMS. There are accepted levels in confidence in metabolite identification (Figure 2-20 Hierarchy of compound annotation in metabolomic analysis.) these consider the specificity of the data used in compound annotation. Compound annotation is a point of discourse within the literature (Reisdorph et al., 2019) (Schrimpe-Rutledge et al., 2016) with some variation between publications.

In this thesis, the evaluation of data annotations involves four levels of compound annotation (Figure 2-20 Hierarchy of compound annotation in metabolomic analysis.).

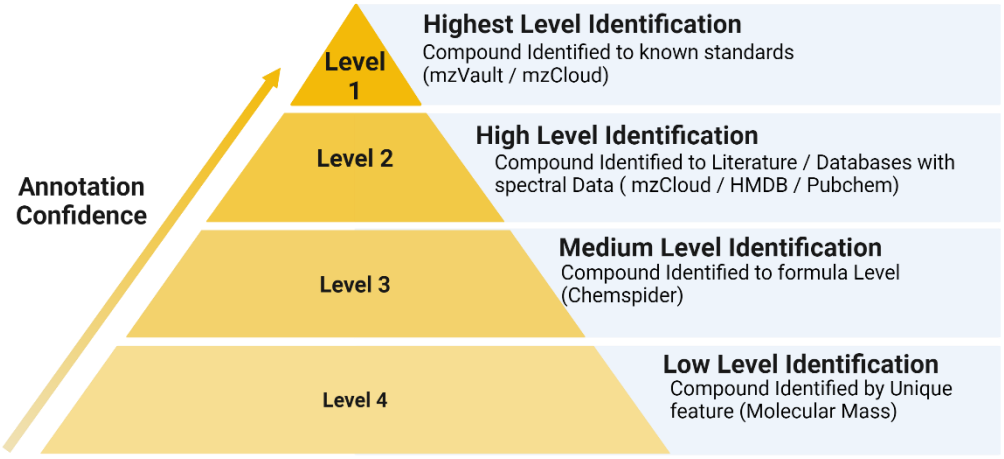

Figure 2-20 Hierarchy of compound annotation in metabolomic analysis.

Levels are in line with those presented by The Chemical Analysis Working Group (CAWG), Metabolomics Standards Initiative (MSI) (Sumner et al., 2007).

Confidence levels used within this thesis are as follows:

- **Level 1** represents the highest level of Identification with a validated ID against a reference standard.

- **Level 2** considered Putative ID using MS^2 spectral data.
- **Level 3** Tentative IDs Compounds identified to formula level Potentially with an accompanying database match.
- **Level 4** consists of compounds identified by a unique feature such as molecular mass.

To assess the annotation confidence that could be expected in datasets analysed using the workflow outlined in this chapter, Iteration 10 data generated from pooled faecal material evaluated in 2.7 LCMS Deep Scan Experiments: Determining Optimal Coverage was analysed for the highest level of compound annotation that could be assigned to a compound for both positive and negative ionisations (Figure 2-21 Compound annotations by highest level identification.). Faecal Sample data was chosen for the anticipated complexity and diversity of compounds, increasing the challenge of compound identification.

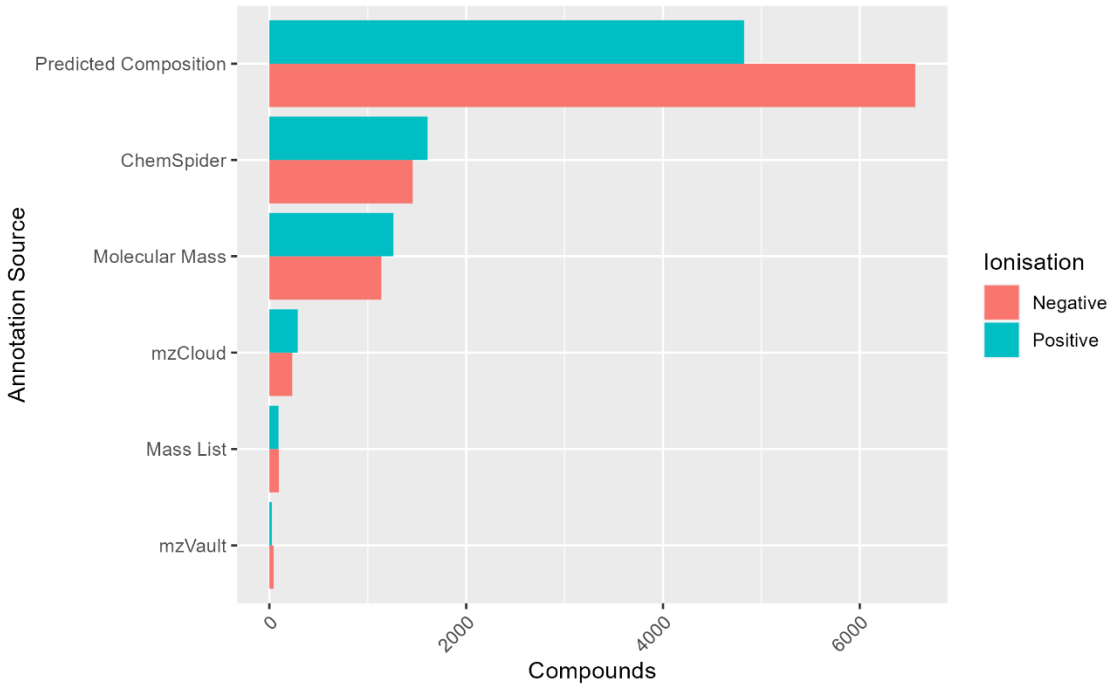

Figure 2-21 Compound annotations by highest level identification.

The workflow identified 8090 compounds in positive ionisation and 9528 compounds in negative ionisation. This resulted in 2237 named annotations for positive and 1881 for negative.

Of these, 313 positive IDs and 275 negative IDs belonged to the two highest confidence levels of MS2 spectra-based annotations: 25 positive and 43 negative in mzVault, and 288 positive and 232 negative in mzCloud.

Tentative compound IDs from MS^1 form the majority of the annotations present within the dataset.

Mass List matches based on molecular mass and retention times obtained running compounds locally, but lacking MS^2 spectra resulting in 93 Positive and 96 negative identifications.

ChemSpider database searches based on MS1 scans using molecular mass and predicted formula obtained 1607 positive IDs and 1457 negative IDs. Compounds which could be annotated to predicted compositions by molecular weight but lacked database matches were 4825 positive and 6563 ne negative.

Compounds which could not be identified to a level beyond a unique molecular mass were 1260 positive and 1137 negative.

2.10 Discussion

This chapter outlines an LCMS approach designed for the untargeted analysis of the intestinal metabolome (Figure 2-22 Final LCMS Metabolomic Workflow for Untargeted Analysis), utilising faecal samples and *in vitro* gut models. In developing this method, it has been important to strike a balance between cost, time, and breadth of data generation, necessitating certain compromises. Within the scope of this thesis, limiting analysis to a single column type was deemed an acceptable concession. This decision was made to ensure the feasibility of the study within the timeframe whilst maintaining a focus on depth and quality throughout the process.

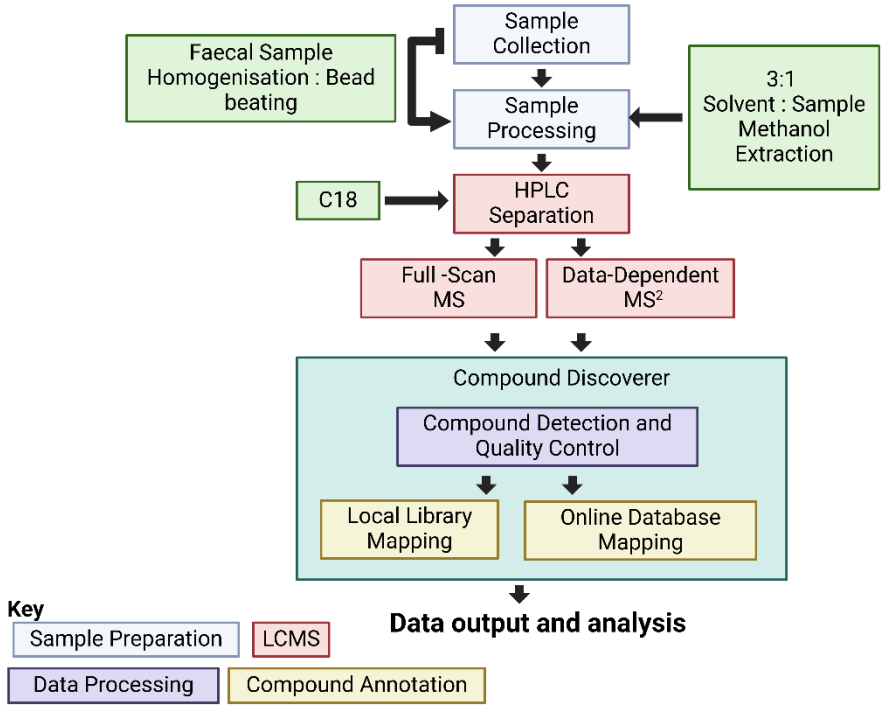

Figure 2-22 Final LCMS Metabolomic Workflow for Untargeted Analysis

Sample Processing

In the development of the metabolomics workflow outlined in this chapter, a core objective was to design a scalable and standardised processing and extraction method, which would facilitate the comparative analysis of large sample sets.

Manual disruption, whilst effective as a means of sample homogenisation, is time-consuming and introduces variability between samples, both of which are

undesirable for large-scale studies. Bead-beating was explored as a potential alternative to manual methods aiming to achieve consistent sample homogenisation. Bead beating is commonly employed as a Lysis and homogenisation technique in omics studies and often serves as an alternative to sonication (van Burik et al., 1998; Santaus et al., 2019). Sonication has been shown to impart higher energy into samples, potentially leading to increased microbial lysis, which can leak intracellular metabolites into the sample (Starke et al., 2019). Bead beating offers a gentler homogenisation approach, creating uniform slurries from faecal samples whilst minimising microbial lysis, which may alter the extracellular metabolome. Through investigation, bead beating demonstrated comparable efficiency in sample processing with less effort and within a significantly shorter timeframe.

Metabolomics, as an emerging field, presents a unique challenge for sample analysis, resulting in a lack of uniformity across current extraction protocols with variations in solvents utilisation and extract proportions (Theriot et al., 2014; Wang, Z. et al., 2018; Franzosa et al., 2019; Sun et al., 2019).

Previously, different solvents have been shown to preferentially extraction some compound classes (Erben et al., 2021). It was important to establish the effect various solvent extractions have on both faecal and *in vitro* models' samples.

As expected, due to reduced sample complexity, all extraction solvents performed more similarly in, *in vitro* model samples than they did in faecal samples. The solvents clustered together closely in both samples, but sample types did not overlap. The comparable nature of extraction potentially explains the heterogeneity of selection seen among extraction protocols in the literature (Theriot et al., 2014; Yu et al., 2017; Franzosa et al., 2019; Sun et al., 2019; Erben et al., 2021). In untargeted analysis, focused on an exploratory approach examining gross metabolome changes solvent choice may be less critical compared to targeted or quantitative studies valuing sensitivity. However, as potential biomarkers reveal themselves, assessing recovery efficiency for different solvents may become increasingly important (Lu et al., 2017).

Due to the emerging area of research LCMS, extraction methods are continually evolving, Kelly, P.E. et al. (2022) describe a valuable method for extractions of metabolites from freeze dried faecal samples. The extraction protocol within this thesis focuses on an initial emulsification step, followed by centrifugation and

filtration. The juxtaposition of sample protocols effectively highlights the importance of application within method development. The methods presented by Kelly, P.E. et al. are optimised for the study of chronic non-infectious gastrointestinal disease (Crohn's disease and coeliac disease) where variation in sample consistency may affect metabolite concentration and extraction efficiency. The method described within this thesis is presented as an effective means of metabolite extraction from *in vitro* models and acute diarrhoeal samples with an emphasis on the removal of human cells and possible pathogens for storage and biosafety concerns.

HPLC

C18 is a widely used reverse-phase HPLC chromatography chemistry valued for its versatility. An objective of this chapter was to assess the scope of compounds which it is suitable for resolving. Utilising metabolite libraries comprised of known standards, provided a defined landscape for testing the analytical boundaries.

Whilst it is unlikely that any single column chemistry would be able to effectively resolve all compounds within a complex and diverse sample such as faecal samples or *in vitro* gut model samples, the aim was to identify which compound classes may be underrepresented in results obtained with this method.

Here, specific compound classes (fatty acyls and steroids, steroid derivatives and organooxygen compounds) with XLogP values at the extreme ends of the polarity spectrum of compound standards were identified as being areas where C18 under-preformed compared to other compound classes.

Given the scope of the experimental question addressed in this thesis, the underrepresentation of certain compound classes was deemed acceptable, as it is unlikely to significantly affect the overall identification of putative biomarkers within the study. With approximately 20% of total compounds annotated, there is a substantial pool of thousands of identified compounds already available for assessment as potential biomarkers. Rather than attempting to further increase compound capture and annotation, it was decided that it may be more productive to named compounds with strong associations. However, it is important to identify and acknowledge these limitations, with the perspective

that the absence of these under-represented compounds from the data may be due to analytical limitations rather than biological phenomena.

Findings align with those seen previously (Babu et al., 2022). The level of interpretation required when reviewing chromatographic and spectral data may account for some differences in findings depending on what is deemed acceptable in chromatographic peaks by researcher assessment or algorithmic analysis. Additionally, variations in instrumental set-up apart from column type, such as chromatographic gradients and solvents, may play a role. Library generation is focused on the generation of quality MS2 spectra for compound identification, and variations in MS2 experimental parameters also plays a roll. These intrinsic factors of LCMS contribute to the remarkable versatility as an analytical platform, yet they may also pose limitations when comparing data across studies. This underscores the importance of local validation for effective compound identification.

In the future, utilising different column chemistries will allow for the expansion into additional compound recovery as research questions become focused on specific compound classes that may be beneficial to data generation (Wang, J. et al., 2015).

Data-Dependant Data Acquisition

For the data-dependent data acquisition experiment used to generate MS2 spectral data, three DDA iterations are considered sufficient for effective spectral library matches see 2.7 LCMS Deep Scan Experiments: Determining Optimal Coverage. While the percentage of MS2 coverage reliably increases up to 10+ iterations, the number of annotations does not follow the same trend. It is important to note that these observations are not definitive rules; they are subject to change based on database completion and would merit further review in the future.

Compound Annotation

The outlined workflow exploring compound annotation in pooled faecal samples yielded a high proportion of useable data, a significant factor contributing to this being the high resolution of the Orbitrap MS instrument used. This enabled precise differentiation of compound masses, enhanced by the implementation of post-processing steps.

High-level compound annotations contributed to a noteworthy proportion of data generated, allowing for the confident identification of many potential compounds of interest. Medium-level annotations were also a valuable addition, providing reliable identifications of compounds which could be selected for validation against standards in the future, particularly if they prove to be potential biomarkers of interest. The prominent contribution of level 1-3 annotations resulted in named compound IDs for 4118 compounds, approximately 20% of compounds across both ionisation modes (2238 Positive / 1881 Negative).

Metabolomics data analysis is constrained by low annotation levels (de Jonge et al., 2022), making data comparison and interpretation problematic. Given the extensive variety of compounds within the intestinal metabolome, the complexity of compounds and the formative stage of this field of study, compound identification will continue to be a challenge resulting in gaps in metabolomic datasets. In the realm of metabolomics research, the results of the workflow implemented align with anticipated outcomes.

Databases

Databases are an essential component in metabolomics. Several have been employed throughout the development of this pipeline and will continue to be instrumental to data curation in the future. Two external compound databases are integrated directly into the compound discoverer workflow for compound identification: mzVault and ChemSpider.

mzCloud is a spectral library containing over 32,000 compounds and continues to expand. Matches are established based on spectral information generation during the MS^2 experiment. Spectra within the database are experimentally generated using known standards in a reference lab. These factors collectively result in high-confidence annotations to the data, potentially aligning with level 1 compound annotations comparable to standards run locally. However, the additional dimension of retention time data is absent. Multiple fragmentation pathways for compound breakdowns are listed with ionisation matching for additional specificity. As with many proprietary products, mzCloud is not freely available for integration into all workflows, its use is predicated on software licences, which incur an additional cost.

ChemSpider is the second database integrated into the compound discoverer workflow. Annotations are derived from molecular mass and compound formula data, generated from MS^1 experiments. While reliance on MS^1 data may result in a lack of specificity, identifications potentially resulting from MS^1 data may lead to potential misattributions to isometric compounds possessing the same formula but differing structural compositions, it does offer the advantage of encompassing a broad scope of compounds.

Within the workflow, database matches typically return a table of ranked matches. In this context, annotations can be manually changed in instances where a compound ID is believed to be incorrectly attributed. For compounds of particular interest, it can be advantageous to subject them to additional processes of manual data curation and verification. However, applying this level of manual curation to the entire set of compounds would be unrealistic.

In addition to databases used directly for compound mapping The Human metabolome database, PubChem, and Chemical Entities of Biological Interest (ChEBI) ((ChEBI), 2024) have been integral to manual data curation.

The Human Metabolome Database with 220,945 metabolites and 4064 compounds with MS^2 reference spectra is not specific to faecal samples, however, encompasses human-associated metabolites and compounds associated with disease. This database profiled high quality information which can be use in manually curating LCMS data, hosting experimentally and computationally generated fragment data.

PubChem hosts information of over 109 million compounds and ChEBI over 60,000 fully annotated compounds are both extensive but non-specific databases. They contain a wealth of useful information which have been instrumental in method development.

Compound Ontology and Taxonomy

Compound ontology present a significant challenge in chemical taxonomy. The choice of which may be specific to the domain of study. For the purpose of this thesis, exploring the human-associated intestinal microbiome and it’s metabolomic interactions, the taxonomy provided by the Human Metabolome Database made the most intuitive sense for the articulation of the research objective and themes explored.

Alternative ontological frameworks were utilised to offer valuable perspectives from alternative viewpoints and have been used to provide useful information during development steps.

2.11 Summary

The work in this chapter provides an initial exploration of how LCMS can be applied to metabolomic studies of the intestinal microbiome.

An effective and scalable extraction protocol was identified using bead beating to maximise metabolite recovery and consistency from faecal samples.

Optimal DDA experimental parameters were explored to enhance compound annotation within the dataset, improving the reliability of data interpretation.

A local library was developed to aid in compound annotation, serving as a reference database and as experimental data set to assess the practical limitations of the outlined analytical pathway, evaluating the recovery capacity of the pipeline and identifying which compound classes may be underrepresented in the final experimental data, highlighting potential areas for further development in the future.

Chapter 3
Gut Model

3.1 Introduction

The human intestinal microbiome stands as an intricate ecosystem characterised by its dynamic complexity. Explorations into the interplay between constituents of the microbiota and host present a myriad of challenges. Navigating the exploration ecological landscape requires significant thought in the selection of an appropriate model.

In modelling the intestinal microbiome, the system must possess the capacity to sustain a diverse range of facultative and anaerobic organisms over periods ranging from weeks to months to accurately capture the shifts in microbial composition and functional capacity which may occur during the exposure to antibiotics and other therapeutic interventions. This requires the precise ability to maintain tight experimental parameters, appropriately reflecting the delicate balance present within the natural system. The model must also be repeatable over multiple replicates to ensure robustness in findings associated with experimental outcomes.

For this project, the Leeds Human Gut Model (HGM) (Freeman et al., 2005) was selected for its proven ability to meet the above criteria, alongside a newly developed model the MiGut created as an evolution to the Leeds HGM. The MiGut has been optimised for the application of molecular methods such as metagenomics and metabolomics. In this study, the aim was to produce a state of dysbiosis within the models using clinically reflective antibiotic dosing regimens, demonstrating a functionally dysbiotic state by challenging with pathogens known to be associated with the condition.

3.2 Experimental Design

The objective of this model was to investigate the potential for antimicrobials to induce significant long-term alterations to microbiome taxonomy and function, resulting in the opening of ecological niches which could be exploited by the exogenous pathogens introduced at week 9 **Figure 3-1 Dysbiosis Model Timeline**.

An initial 2-week steady-state period allows microbial populations within the system to reach self-regulated homeostasis. Following the establishment of a baseline healthy microbiome, antimicrobial instillation was initiated.

Antibiotics, chosen for their clinical relevance, were employed to simulate the repetitive exposure a patient might undergo during a chronic infection. Antibiotic instillation phases each consisted of a 5-day dosing regimen followed by 2 days with no antibiotics administered.

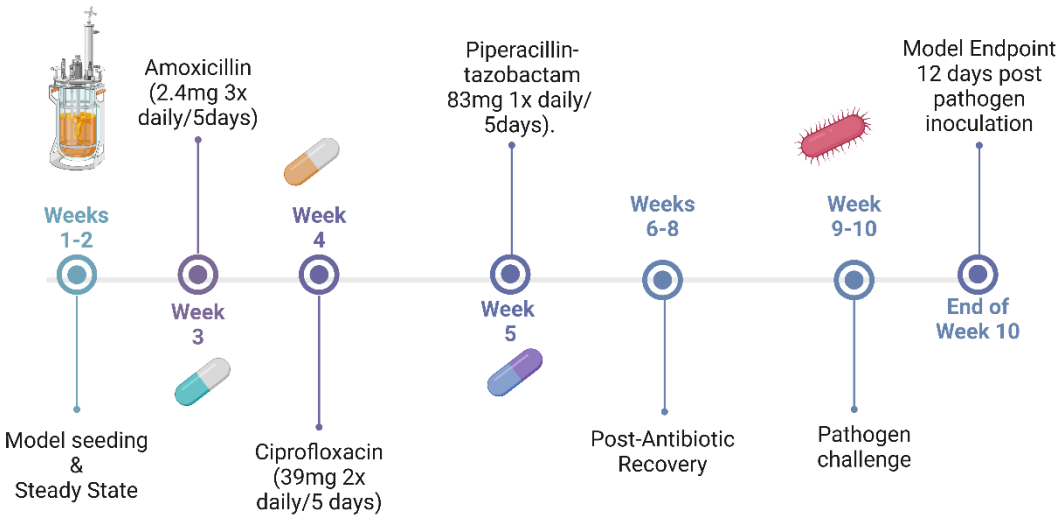

Figure 3-1 Dysbiosis Model Timeline. The timeline depicts key experimental time points for both MiGut and Leeds HGM models. Major experimental time points including model seeding, antibiotic instillation and pathogen challenge are highlighted. Following completion of the final phase of antibiotic therapy, the model was maintained free from intervention for 3 weeks to allow the microbiome to recover and reach a new homeostasis.

At week 9, the ability of the recovered microbiome to resist the colonisation and expansion of exogenous pathogens was challenged by the introduction of two clinically relevant strains, *C. difficile* (027 210) and *K. pneumoniae, a Klebsiella*

pneumoniae carbapenemase (KPC) producing strain. The colonisation and expansion of introduced pathogens were monitored as well as the production of *C. difficile* toxin.

Metagenomic and Metabolomic samples were collected from Vessels 1 & 3, with each sampling occurring on the final day immediately before the next therapeutic intervention or prior to the introduction of pathogens. Extracted metabolites and DNA were systematically stored for subsequent batch analysis subsequent date. Following the introduction of Pathogens to the model, sampling scope was increased to track the expansion and proliferation of introduced strains as well for the presence of *C. difficile* toxin within the system.

3.3 Methods

3.3.1 *In-vitro* models

Two model variations are employed in the study both of which have been validated for use in the study of the intestinal microbiome. The Leeds Human Gut Model and the MiGut platform model variations address specific requirements of culture-based and omics-based exploration of the intestinal microbiome. Models are seeded with ex-vivo populations derived from faecal material from volunteers.

The following sections outline the methods used for in vitro models. Unless stated otherwise, the described methods apply to both systems, with any deviations explicitly noted.

3.3.1.1 Ethical approval

The collection and use of human faeces in the model has been approved by the School of Medicine Research Ethics Committee, University of Leeds (MREC-15-070-Investigation of the interplay between Commensal Intestinal Organisms and Pathogenic Bacteria).

3.3.1.2 Leeds Human Gut Model

The Leeds Human Gut Model (HGM) is an *in vitro* model developed by the Healthcare-associated Infection (HCAI) Research group primarily for the study of *C. difficile* infection and adapted from the triple-stage model originally developed by MacFarlane *et al. (Macfarlane et al., 1998)*. This model consists of a triple-stage continuous chemostat arranged as a cascading weir system (Figure 3-2 The Leeds human gut model (HGM) and configuration overview). Each of the individual vessels within the model maintains distinct nutrient and environmental conditions to reflect the changing physiological conditions from the proximal to the distal colon (Crowther et al., 2016).

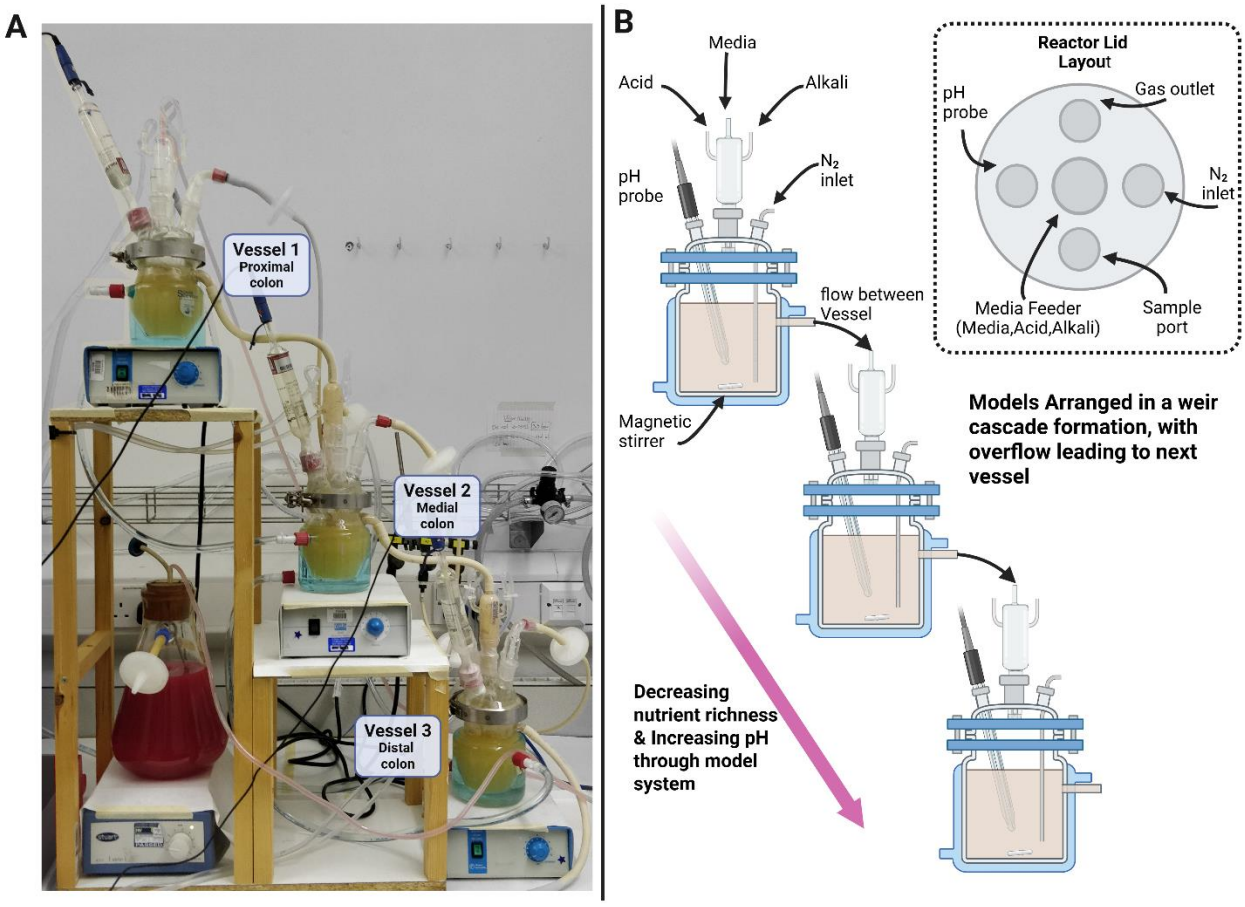

Figure 3-2 The Leeds human gut model (HGM) and configuration overview .(A) Picture of a single replicate of Leeds HGM. (B) Configuration of Leeds HGM, showing lay out of media, acid, alkali, nitrogen inputs and gas outlet and sample port. The system is a weir cascade with vessel overflow feeding the next in weird with outflow from Vessel 3 directed to waste

Model Parameters

The Leeds HGM maintains tight environmental controls within the system, which reflect the physiological conditions of the human colon.

Consistent pH ranges are maintained through the addition of acid (1M Hydrochloric acid) and base (1M sodium Hydroxide). The pH within the vessels is continually monitored via a probe within each of the vessels, and acid/base dosing is handled via automated control units set to dose when readings exceed the upper/lower bounds described in Table 3-1 Leeds HGM vessels and their respective conditions..

Vessel contents are continually purged with nitrogen to maintain an anaerobic atmosphere reflective of the colon. Nitrogen is supplied via a nitrogen generator. An anaerobicity indicator (resazurin) present in the media serves to provide a visual indication of whether vessels have been sufficiently purged of oxygen. In the absence of oxygen, the vessels maintain a “straw-like” colouration; when exposed to oxygen, they develop a pink-red hue.

Temperatures within the vessel are maintained at 37^{o}C by a water jacket. Water temperature is maintained by a heated water bath with a pump to circulate water throughout the system.

Vessels are arranged in a weir cascade; this allows retention to be controlled by media flowing into the system. Media is pumped into vessel one at a rate of 0.3mL/min resulting in a retention time within the model of ~48 hours.

Homogeneity in the contents of each vessel is achieved by constant agitation with a magnetic stirrer and flea to ensure uniformity of nutrients and antibiotics.

In some instances, excessive foaming within the reactors may occur. Foaming its ameliorated with the addition of an aqueous-silicone antifoam emulsion (Antifoam B Sigma) suitable for use in microbiological culture systems. Antifoam concentrate is used to create a 10% solution in vegetable oil as per supplier recommendations.

Table 3-1 Leeds HGM vessels and their respective conditions.

	Biological Analogue	Volume (mL)	pH	Temperature (ºC)
Vessel one	Proximal colon	280	5.5 ± 0.2	37
Vessel two	Medial colon	300	6.2 ± 0.2	37
Vessel Three	Distal colon	300	6.8 ± 0.2	37

In-vitro HGM growth Media

The HCAI research group has developed a complex defined growth media designed specifically for the sustainability of intestinal microbial populations over longitudinal experiments. The constituents described in Table 3-2 Leeds HGM Growth Media. Growth media is made in 2L volumes. Except for glucose, all constituents are added to 1985mL of distilled water in a 2L Buncher flask. A vent filter (Merck Millipore, USA) is added to the side arm. A bung with two ports; one for the addition of supplements post autoclaving and another for the drawing of media via peristaltic pump. A Magnetic stirrer bar is added for mixing of the media to prevent sedimentation. Following autoclaving, glucose (dissolved in 10mL distilled water) and anaerobicity indicator resazurin (5mL) are added via the addition port in the top of the flask through a 0.22µm syringe filter (Merck Millipore, USA).

Table 3-2 Leeds HGM Growth Media

Constituent	Amount	Supplier
Magnesium sulphate	0.01 g/L	Sigma
Calcium chloride	0.01 g/L	Sigma
Sodium chloride	0.1 g/L	Sigma
Di-potassium mono hydrogen phosphate	0.04 g/L	Fisher
Potassium dihydrogen phosphate	0.04 g/L	Fisher
Sodium hydrogen carbonate	2.0 g/L	Sigma
Haemin	10 ml/L	Sigma
Cysteine HCL	0.5	Sigma
Bile Salts	0.5	Sigma
Arabinogalactan	1.0	Chem Cruz
Tween 80	2 mL/L	Sigma
Pectin	2.0	Acros Organics
Starch	3.0	Oxoid
Vitamin K1	10 µL/L	Sigma
Peptone Water	2.0	Oxoid
Yeast extract	2.0	Oxoid
Chenodeoxycholic acid	0.25 mg/L	Sigma
Lithocholic acid	0.017 mg/L	Sigma
Mucin	1 g/L	Sigma

Donor Specimen Collection

Initial bacterial populations within the model are derived from a faecal slurry created from the pooled faeces of healthy donors (>18 years old), and mixed gender. Healthy donors are defined as having no current acute gastrointestinal distress and no history of antibiotic use within the previous 3 months. Participants self-collected specimens using the collection kits provided.

Samples were stored in sterile 180m: poly-propylene containers, which were then stored anaerobically using AnaeroGen W-Zip Compact generator system (Oxoid) to maintain an anaerobic environment prior to delivery to the laboratory. Upon delivery to the laboratory, the samples were transferred to an anaerobic workstation (Don Whitley Scientific A95).

Screening of Donor Faeces

To ensure the absence of experimental pathogens within donor faeces used in the seeding of the models, all donor specimens were screened for the presence of *C. difficile* GDH antigen as well as undergoing culture-based screening for *C. difficile* and KPC on selective media described in Table 3-4 Introduced Pathogens and Media for Culture

.

Preparation of the Faecal Slurry

The faecal slurry for seeding the model is made from equal parts taken from 3 donor specimens. Specimens are weighed on a laboratory balance into a stomacher bag with an integrated filter. The pooled faecal sample was diluted with sterile anoxic PBS. The diluted sample was homogenised using a laboratory blender (Stomacher). The integral strainer was closed to retained fibrous material, and the slurry was decanted into a sterile 2000mL bottle and placed in an anaerobic workstation for aliquoting. The Faecal slurry was separated into 160ml aliquots.

Seeding Model

Each vessel is inoculated with 160mL of faecal slurry after purging the empty system with nitrogen overnight.

3.3.1.3 MiGut

The MiGut platform is a scalable system developed from the Leeds HGM by William A. Davis Brich, which has been validated for the simulation of antibiotic-induced dysbiosis (Davis Birch et al., 2023).

The MiGut platform scales down the model volume from 300mL to 45mL and introduces remote system monitoring to improve efficiency, allowing for increased parallelisation. Temperature and pH remain the same as Leeds HGM. The flow rate of media is scaled to match vessel retention time within each stage of the model to that of the Leeds HGM.

The reduced scale of the MiGut platform allows for four parallel replicates within the spatial footprint of a single Leeds HGM system. Figure 3-3 (Details of the Migut platform and model.) show four replicates of the MiGut platform. Within the present study, samples for metagenomic (Migut replicates sampled, n=1) and metabolomic (Migut replicates sample, n=3) analysis were sampled from vessel 3 representing the distal colon (identified in Figure 3-3 (G) Vessel 3).

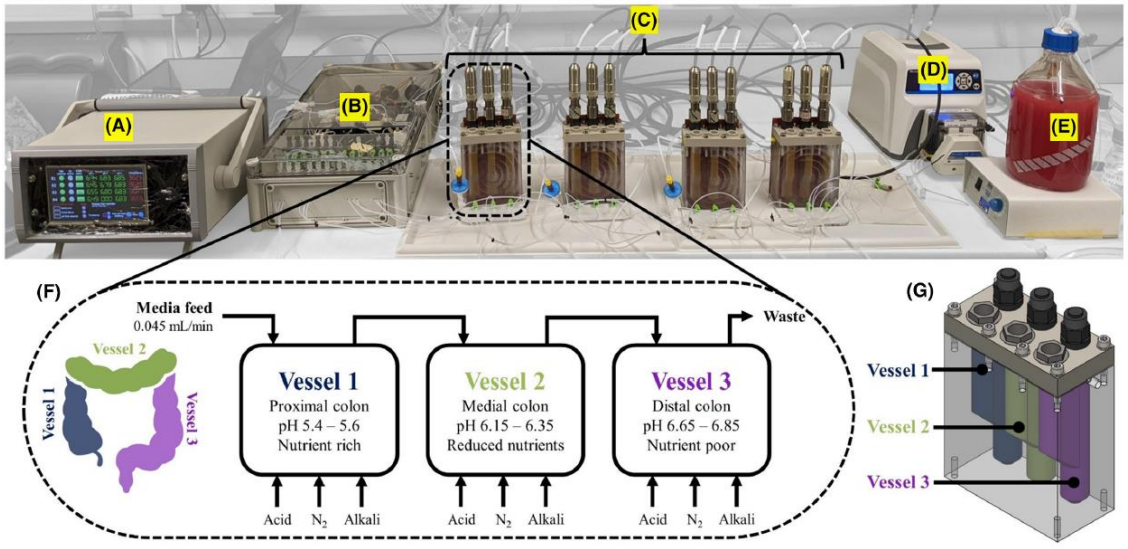

Figure 3-3 Details of the Migut platform and model. Central controller (B) pumping unit (C) four triple-stage gut models each capable of running a unique setoff conditions (D) peristaltic pump for media addition (E) nutrient-rich media (F) details the flow within the reactor (G) 3D render of reactor unit, with colour coded vessels. Image reproduced from MiGut: A scalable in vitro platform for simulating the human gut microbiome-Development, validation, and simulation of antibiotic-induced dysbiosis (Davis Birch et al., 2023)

As the over-arching purpose of this PhD is to explore the Intestinal microbiome and the changes induced beyond those seen at a taxonomic level, this model

provides an ideal platform providing biological replicates for analysis *via* Metagenomics/Metabolomics/rtPCR.

Seeding Model

Due to the reduced size of the MiGut platform, vessels are inoculated with 20mL of faecal slurry after purging the empty system with nitrogen overnight.

3.3.2 Bacterial Culture & Enumeration.

Samples were taken from the Leeds HGM for the enumeration of key culturable populations listed in Table 3-3 Key bacterial Strains and Culture Media. Sample frequency was twice weekly during steady state, increasing to three times a week with the beginning of antimicrobial dosing. Approximately 10mL was taken from vessels 1 and 3. Where required, samples were sub-aliquoted for metabolomic, metagenomics, and toxin. Samples were immediately transferred to an anaerobic workstation. They were serially diluted down to 10^{-7} into anoxic peptone waters and plated out in triplicate on selective and non-selective agars described in Table 3-3 Key bacterial Strains and Culture Media and Table 3-4 Introduced Pathogens and Media for Culture. Spore counts were achieved by diluting neat samples with equal parts ethanol (96% v/v) and incubating at room temperature for one hour to kill vegetative cells present and serial diluting. Colony morphology /characteristics were used to attribute genus-level identification to cultured organisms. A Matrix-assisted laser desorption/ionisation – Time of Flight (MALDI-TOF) biotyper platform (Bruker Daltonik Biotyper) was used on unique colonies to confirm identifications and give a species-level diversity estimate MALDI-TOF ID was performed at the end of each experimental phase with the aim of monitoring how the composition of target communities were evolving throughout the experiment.

Table 3-3 Key bacterial Strains and Culture Media.

Target Organism	Culture Media	Growth conditions
Total Facultative anaerobes	Nutrient Agar (NUT), Pre-poured (Oxoid)	24 hours at 37ºC Aerobically
Enterococcus	Kanamycin Aesculin Azide Agar (KAA)	48 hours at 37ºC Aerobically
Lactose Fermenting *Enterobacteriaceae*	MacConkey (MAC), Pre-poured (Oxoid)	24 hours at 37ºC Aerobically
Total Anaerobes	Fastidious Anaerobe agar (FAA) with 5% horse blood, Pre-poured (Oxoid)	48 hours at 37ºC Anaerobically
Clostridia	Fastidious Anaerobe agar (FAA) with 5% horse blood, Pre-poured (Oxoid)	48 hours at 37ºC Anaerobically
Lactobacillus	LAMVAB Agar (LVAB)	48 hours at 37ºC Anaerobically
Bifidobacterium	Beerens Agar (BEER)	48 hours at 37ºC Anaerobically
Bacteroides	Bacteroides Bile Esculin (BBE) agar	48 hours at 37ºC Anaerobically
Total spores	Fastidious Anaerobe agar (FAA) with 5% horse blood, Pre-poured (Oxoid)	48 hours at 37ºC Anaerobically

Table 3-4 Introduced Pathogens and Media for Culture.

Target Organism	Culture Media	Growth conditions
C.difficile (Ribotype 027), Total viable counts (TVC)	Brazier's CCEY agar + Moxifloxacin (MXF)	48 hours at 37ºC Anaerobically
C.difficile (Ribotype 027), spores	Braziers CCEY agar + Moxifloxacin (MXF)	48 hours at 37ºC Anaerobically
Carbapenemase-producing *Enterobacteriaceae* (CPE)	CHROMID CARBA SMART Agar (CARBA),Pre-poured (Biomerieux) Bi-plate - OXA-48 screening media + KPC and NDM-1 screening	24 hours at 37ºC Aerobically

3.3.2.1 MALDI-TOF Identification of Cultured Strains

Unique colonies were identified by eye by morphological variation on each agar type. Representative colonies for each subpopulation were selected for MALDI-TOF identification. Colonies were picked and transferred to a steel MALDI-TOF target plate, and 1µL of α-Cyano-4-hydroxycinnamic acid matrix was added and allowed to air-dry. The processing of prepared target plates was performed by the Microbiology Department of Leeds General Infirmary, feeding into pipelines for the processing of routine diagnostic samples. Samples were screened against MALDI Biotyper Reference Library (MBT Compass and MBT Explorer Software plus MBT Compass Library), which was used to confirm (genus) and attribute (species) level identifications.

Table 3-5 Quality Scores for MALDI-TOF Identification.

Score Value	Identification Level
2.3 – 3.0	Genus: Reliable / Species: Reliable
2.0 – 2.29	Genus: Reliable / Species: Probable
1.7 – 1.99	Genus: Probable

Identifications are scored numerically on a scale of 1.0 – 3.0 using Brunker's quality scoring algorithm. Identifications <1.7 were deemed unreliable and repeated. Genus and species IDs were recorded based on quality scores described in Table 3-5 Quality Scores for MALDI-TOF Identification.

Where required colonies could be sub-cultured onto non-selective media; Colombia blood agar (CBA) or Fastidious anaerobe agar (FAA) and identification reattempted on fresh growth following a further 24–48-hour incubation.

3.3.3 Antibiotic Instillation

Antibiotics were selected based on data from a previous study at Leeds Teaching Hospitals. This indicated that piperacillin-tazobactam, amoxicillin and ciprofloxacin were the most commonly prescribed broad-spectrum antibiotics received by patients with antibiotic-associated diarrhoea without CDI (n=50), or with CDI (n=15) in the preceding 12months. Together they accounted for >50% of previously prescribed antibiotics (Freeman J, 2021)

3.3.3.1 Antibiotic Dosage & Schedule

Antibiotic dosing was designed to reflect clinical regimes and achieve concentrations in Vessel 1 equivalent to those typically seen in the colon described Table 3-6 Antibiotics, Dosing Frequency and Leeds HGM Vessel 1 target concentration. Antibiotics were prepared for the week in advance. 50mL of each antibiotic was made in the required solvent, Table 3-6 Antibiotics, Dosing Frequency and Leeds HGM Vessel 1 target concentration. Antibiotics prepared in sterile distilled water were filter sterilised through 0.22µM syringe filters. 1mL aliquots were taken, transferred to sterile microtubes, and stored at -20^{o}C until needed.

Dosage concentrations were designed so that when administered to vessel one, they would achieve the desired concentration in the 280mL vessel volume. Daily dosing schedules reflect those typically seen prescribed to patients.

Table 3-6 Antibiotics, Dosing Frequency and Leeds HGM Vessel 1 target concentration.

Antibiotic	Dosing Frequency	Vessel 1 Concentration	Solvent
Amoxicillin (AMX)	3 x Daily	8mg/L	Sterile Distilled water
Ciprofloxacin (CIP)	2 x Daily	132mg/L	Acetic Acid
Piperacillin – Tazobactam (TZP)	3 x Daily	276mg/L	Sterile Distilled water

3.3.4 Pathogen Challenge

Pathogens were introduced on day 55 of the model, 3 weeks after the cessation of antimicrobial therapy. Vessel 1 was inoculated with $1x10^4$ cfu of *Klebsiella pneumoniae* (KPC) and $1x10^7$ cfu *C. difficile* spores (PCR ribotype 027 210, Maine Medical Centre (Portland) isolate) on day 55.

3.3.5 Cell Cytotoxicity Neutralisation Assay

Cell Cytotoxicity Neutralization Assay (CCNA) is often regarded as the golden standard for diagnosis of *C. difficile* infection (CDI) (McDonald, L.C. et al., 2018). CCNA tests for the presence of *C. difficile* toxin within a stool sample or gut model fluid with assays being predominantly sensitive to the presence of toxin B. This allows for the identification of active CDI rather than quiescent colonisation.

The CCNA assays faeces/HMG fluid against Vero cells (African Green Monkey Kidney Cells, ECACC 84113001). *C.difficile* toxin-induced cell death is confirmed by the neutralisation of cytotoxicity by the Presence of *Clostridium sordelli* anti-toxin.

Two CCNA assay variations were used throughout the experiment, testing for C. difficile toxin production within the model.

Samples were taken and tested for toxin the day before pathogen introduction to confirm the absence within the model before inoculation with *C. difficile*, and every day following.

An initial screening assay was preformed daily testing for the presence/absence of toxin. Following detection, a titration assay was used to ascertain the abundance of toxin within each of the vessels.

Toxin levels in each of the vessels were monitored daily post-inoculation with *C. difficile*. Following the detection of toxin within the model, concentrations were monitored daily.

3.3.5.1 Vero Cell Preparation

Vero Cell Culture

Vero cells are grown in flat bottom tissue culture flasks containing 20mL of Dulbecco's Modified Eagles medium described in Table 3-7 Vero Cell Growth Media, and incubated with 5% CO_2 at 37°C until confluent monolayers are formed after 48 hours. Monolayers are examined for confluency by microscopy (Olympus UK Ltd) prior to harvesting.

Table 3-7 Vero Cell Growth Media

Constituent	Amount	Supplier
Dulbecco's Modified Eagles Medium	20mL	Sigma
newborn calf serum	10 % (v/v)	Gibco
L-glutamine	1 % (v/v)	Sigma
antibiotic/antimycotic solution	1 % (v/v)	Sigma

1mL of trypsin-Ethylenediaminetetraacetic acid (EDTA), is added to 9mL Hanks Balanced Salt solution (HBSS) per flask passaged. Spent media from the cells to be passaged is discarded, the flask is washed with 5mL of trypsin-EDTA + HBSS to remove residual media. Washed cells are incubated at room temperature for 1 minute with the remaining 5mL trypsin-EDTA + HBSS, which is then discarded. The flask is recapped, and cells incubated at 37°C for 5-10 minutes, after which the cells will be detached from the bottom surface of the flask. 6mL of Vero cell Growth media is added and gently mixed with a serological pipette to create a uniform cell suspension. This volume can be altered if confluency is less than 100%; for 80% confluence, 4.8mL should be used (6mL x 0.8 = 4.8mL).

Flask Seeding

2mL of Vero cell suspension is added to each flask being prepared. 18mL of Vero Cell Growth Media is added to each flask. The passage date is recorded, and the passage number is incremented by 1 and recorded on the flask. Flasks are returned for incubation at 5% CO_2 at 37°C for 48 hours.

CCNA Tray Seeding

For each tray a 2:18mL of Vero cell suspension: Growth media is made in a sterile universal. 180µL of cell suspension is added to sample wells, and 160 µL added to negative wells shown in Figure 3-4 CCNA Vero Cell Seeding Layout.

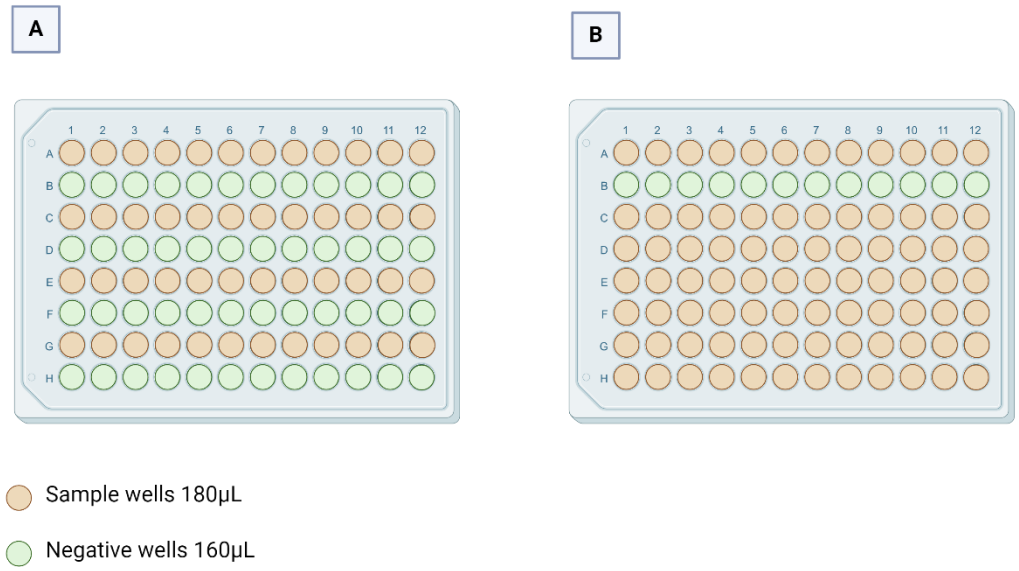

Figure 3-4 CCNA Vero Cell Seeding Layout. (A) Screening assay layout (B) Titration assay layout

After seeding, trays are labelled with date of production and are returned for incubation at 5% CO_2 at 37ºC for 48 hours.

3.3.5.2 CCNA Assay Protocol

Model Sample Processing

1mL of sample is aliquoted into a sterile microcentrifuge tube and centrifuged at 10,000g for 10 minutes to pellet out cells. Supernatant filtered sterilised with a 0.22µM syringe filter into a new sterile 1mL microcentrifuge tube. Filtered toxin samples can then be stored at 4ºC until required.

CCNA Tray inoculation

20µL of *C. sordellii* antitoxin (Prolab Diagnostic) is added to all Negative wells in the tray using a multi-dispenser pipette (Multipette M4, Eppendorf) , rows B,D,F,H in screening trays, row B in titration assay Figure 3-4 CCNA Vero Cell Seeding Layout.

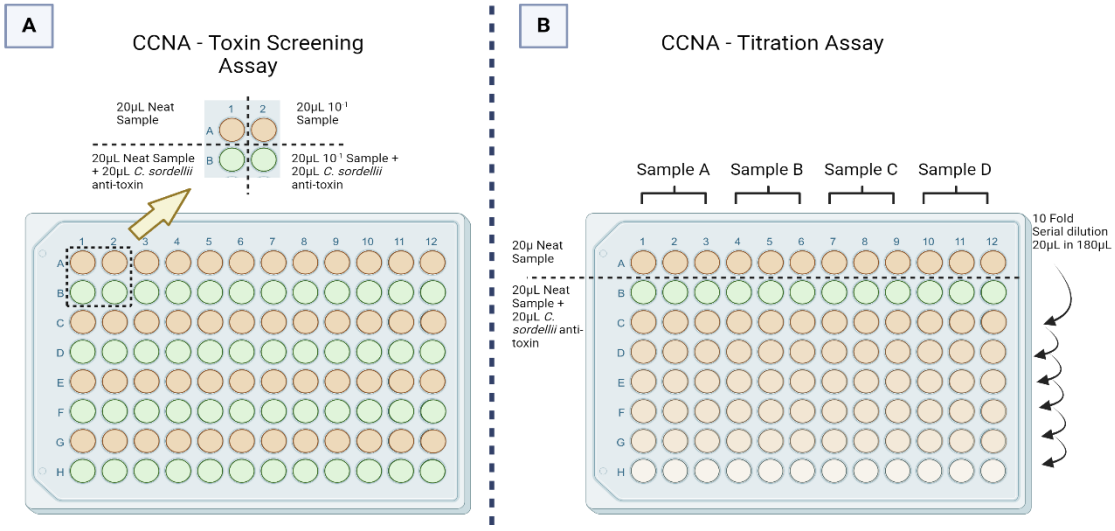

Figure 3-5 CCNA Assay variations for Detection and Quantification of *C. difficile* Toxin.

(A) Plate format used in the screening for *C. difficile* Toxin (B) Titration plate format used for the quantification of *C. difficile* Toxin.

For Screening assays, 20µL of filtered toxin is inoculated into a 180µL Sample well and gently pipette mixed to prevent disturbance of cell monolayers adhered to the bottom of the well. 20µL from the well can be carried over to the next well creating a 10-fold dilution. This is repeated in the row below containing 160µL + 20µL Antitoxin. Previous steps are repeated for sample replicates and positive control.

Titration assays 20µL of neat sample is added for replicates across row A and B. A 12-channel pipette is then used to create a 10-fold serial dilution into rows C, D, E, F, G and H. After inoculation, trays are incubated at 5% CO_2 at 37^{o}C.

Reading CCNA Results

Trays are read after 24 and 48 hours, cell appearance is assed for the impact of toxin within the sample, Healthy cells remaining adhered to the bottom of the well are deemed to be negative, some rounding may be seen within healthy populations but should be present in <50% of cells. Wells with cell rounding (>50% cells) and detachment from the well surface, which is also neutralised in the corresponding negative well are considered positive Figure 3-6 Vero cells displaying Positive and Negative reactions. Wells with >50% cells rounding that fail to be neutralised by anti-toxin are considered to display non-specific cell rounding and recorded as negative for *C. difficile t*oxin. If all wells for a sample including those inoculated with anti-toxin display >50% rounding, no result is recorded and test for that day is repeated.

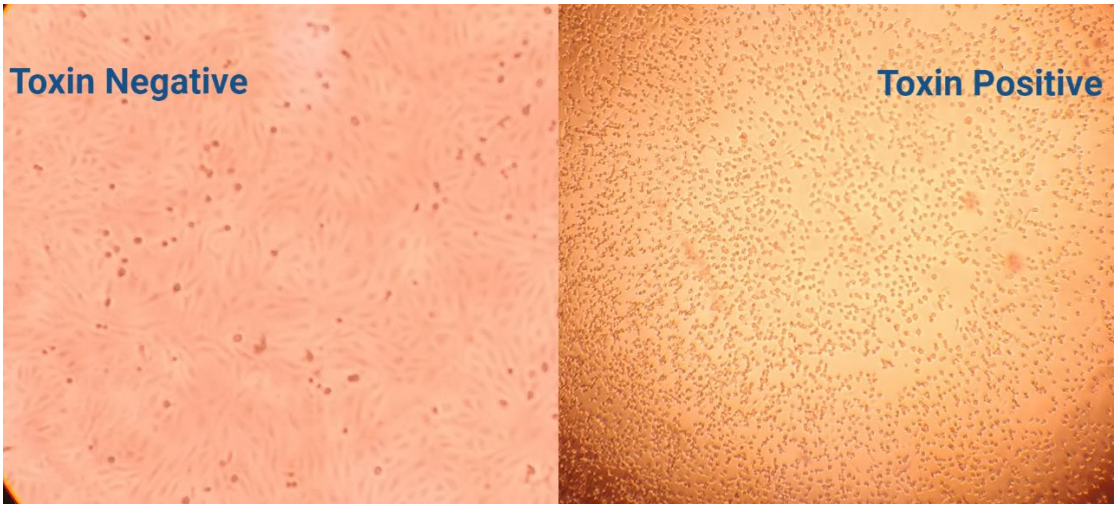

Figure 3-6 Vero cells displaying Positive and Negative reactions.

The results of screening assays are reported as Positive/Negative for the day. Titration results are reported at relative units (RU). RU are reported based on the highest dilution in which a positive result was recorded. A positive result seen in a 10^{-1} would be given a RU of 1.

3.3.6 Metagenomic and Metabolomic Sample Collection

Samples for metagenomics and metabolomics were taken on the final day of all experimental phases; Steady State, AMX, CIP, TZP, Post Antibiotic recovery, and Pathogen Challenge.

Samples for metagenomics and metabolomics were processed as stated in 2.4 Sample Processing and 4.4.1 DNA Extraction and stored for batch analysis later.

3.4 Results

3.4.1 Impacts of Antimicrobials on the Gut Microbiome

Data described in the following sections are Log_{10} colony-forming unit(CFU)/mL and Log_{10} transformation of mean CFU/mL. Mean CFU/mL is calculated from 20µL of inoculum across 3 plates. This sampling method results in a limit of detection of 16.66 CFU/mL (1.22 Log_{10} CFU/mL).

Each week's antibiotic period outlined consists of a 5-day active antibiotic instillation period followed by a 2-day period without dosing. Peak disruption values are taken from the final day of dosing. This is assumed to be where concentrations of therapeutics were at their highest.

3.4.1.1 Facultative Anaerobes

Facultative Anaerobes – Vessel 1

The initial phase of the experiment involved running the model without any interventions, allowing for the colonization and stabilization of monitored populations within the vessel. During this period, lactose-fermenting (LF) *Enterobacteriaceae* exhibited higher abundance (6.42 Log_{10} CFU/mL) compared to Enterococcus spp. (6.38 Log_{10} CFU/mL) Figure 3-7 Facultative Anaerobes: Leeds HGM Vessel 1 (Proximal).

The introduction of amoxicillin resulted in an initial decrease in the overall abundance of facultative anaerobes within the proximal region, decreasing from 7.20 Log_{10} CFU/mL to 6.35 Log_{10} CFU/mL. Populations began to recover during active antibiotic instillation and continued even after cessation of amoxicillin dosing, with facultative anaerobes reaching 7.76 Log_{10} CFU/mL prior to ciprofloxacin instillation. This increase can be largely attributed to an expansion of LF *Enterobacteriaceae*, which increased by 1.26 Log_{10} CFU/mL, while

Enterococcus spp. declined throughout the entire period, falling by 0.54 Log_{10} CFU/mL.

The instillation of ciprofloxacin coincided with a decrease in the overall abundance of facultative anaerobes, both during active instillation and the subsequent cessation of dosing period, dropping from 7.76 Log_{10} CFU/mL to 7.15 Log_{10} CFU/mL. LF *Enterobacteriaceae* also decreased, falling by 0.78 Log_{10} CFU/mL, and *Enterococcus spp*. declined by 1.05 Log_{10} CFU/mL.

During TZP installation, Enterococcus spp. increased in abundance by 1.51 Log_{10} CFU/mL, while *LF Enterobacteriaceae* decreased by 0.79 Log_{10} CFU/mL. Overall, total facultative anaerobes displayed a reduction of 0.42 Log_{10} CFU/mL

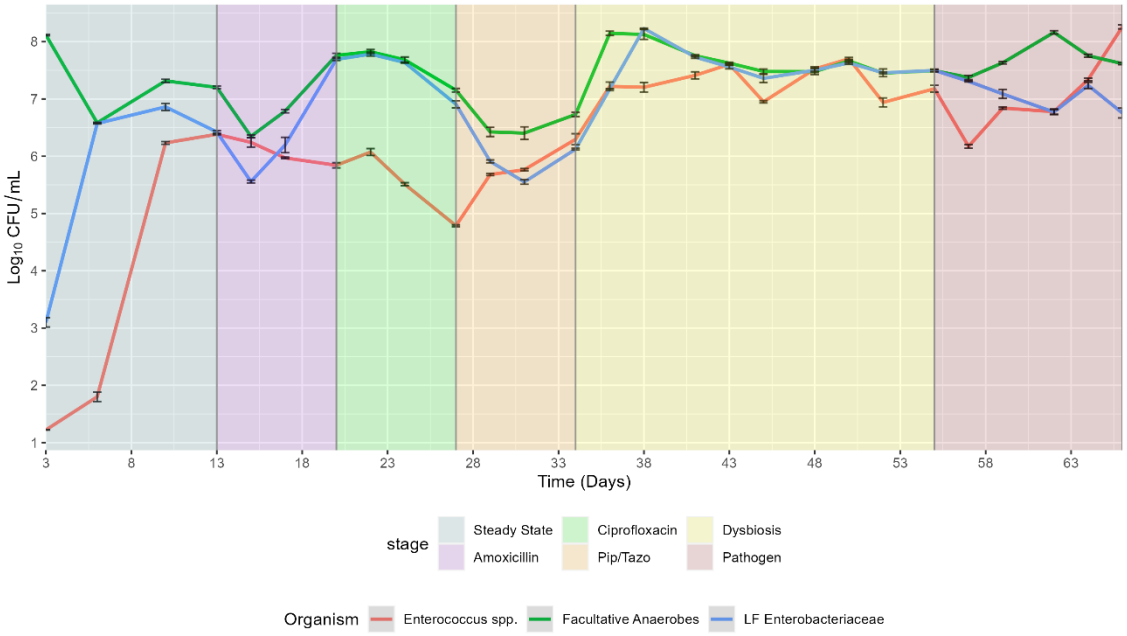

Figure 3-7 Facultative Anaerobes: Leeds HGM Vessel 1 (Proximal). Facultative Anaerobes (Green) *Enterococcus spp*. (Red) and Lactose Fermenting *Enterobacteriaceae* (Blue)

During the post-antibiotic recovery period, all populations swiftly rebounded, ultimately reaching abundances comparable to or exceeding their baseline levels by day 38 (7 days following the final antibiotic dose): Enterococcus spp. reached 7.20 Log_{10} CFU/mL, LF Enterobacteriaceae reached 8.23 Log_{10} CFU/mL, and facultative anaerobes reached 8.13 Log_{10} CFU/mL.

Facultative Anaerobes – Vessel 3

During steady state, the microbial populations within the system achieved stable abundances. Facultative anaerobes were present at 7.38 Log_{10} CFU/mL, while Enterococcus spp. and *LF Enterobacteriaceae* established at levels of 6.13 Log_{10} CFU/mL and 6.17 Log_{10} CFU/mL, respectively Figure 3-8 Facultative Anaerobes: Leeds HGM Vessel 3 (Distal).

Upon introduction of amoxicillin, there was an initial decrease in the abundance of Facultative Anaerobes, which later stabilized, resulting in an overall decrease of 0.77 Log_{10} CFU/mL during the period. *Enterococcus* spp. exhibited a steady reduction in abundance initially, followed by a sharp decline during the 2-day cessation of the dosing phase, resulting in a total decrease of 1.23 Log_{10} CFU/mL over the period. In contrast, LF *Enterobacteriaceae* displayed consistent growth in abundance throughout amoxicillin instillation, resulting in an increase of 0.39 Log_{10} CFU/mL.

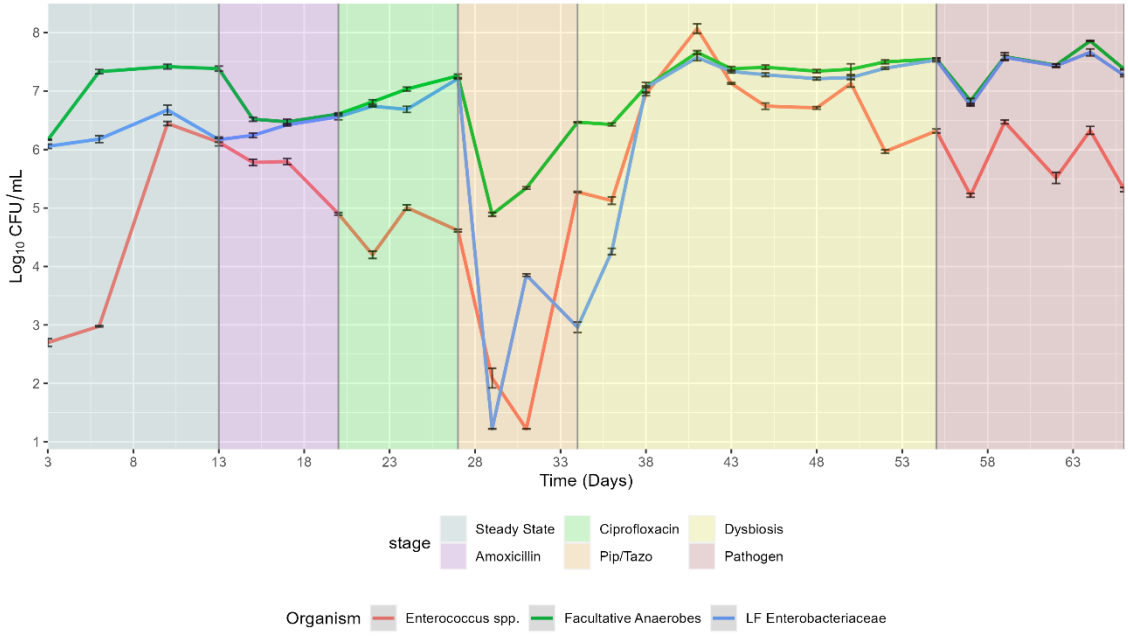

Figure 3-8 Facultative Anaerobes: Leeds HGM Vessel 3 (Distal). Facultative Anaerobes (Green) *Enterococcus spp*. (Red) and Lactose Fermenting (LF) *Enterobacteriaceae* (Blue)

Ciprofloxacin instillation corresponds with a continuous rise in total facultative anaerobes, increasing by 0.65 Log_{10} CFU/mL consistently throughout both the active dosing and drug-free interval. Similarly, LF *Enterobacteriaceae* experienced a concurrent increase of 0.65 Log_{10} CFU/mL over the same

duration. While *Enterococcus spp*. displayed some recovery following an initial decline, ultimately, they exhibited a net decline of 0.29 Log_{10} CFU/mL.

With TZP dosing, all populations experienced steep initial decreases in abundance; Facultative Anaerobes fell by 2.37 Log_{10} CFU/mL, LF Enterobacteriaceae decreased by 5.99 Log_{10} CFU/mL, falling below the limit of detection (LOD), and Enterococcus spp. declined by 2.52 Log_{10} CFU/mL. Despite some recovery during active antibiotic instillation, both LF Enterobacteriaceae and Facultative Anaerobes failed to reach abundances seen prior to TZP, resulting in net reductions of 4.25 Log_{10} CFU/mL and 0.79 Log_{10} CFU/mL, respectively. Whilst *Enterococcus spp*. also fell below the LOD during active installation, the population flourished following the cessation of antibiotic dosing. *Enterococcus spp*. quickly recovered to levels exceeding TZP introduction, surpassing *LF Enterobacteriaceae* in abundance.

Populations in the distal portion of the model returned to abundances seen prior to the administration of any antibiotics at day 41, 10 days after the final dose of antibiotics.

3.4.1.2 . Obligate Anaerobes

During the steady state within the proximal stage of the model, Bifidobacterium spp. reached 6.80 Log_{10} CFU/mL, Lactobacillus spp. stabilised at 7.30 Log_{10} CFU/mL, Bacteroides spp., despite being present in the faecal slurry used to inoculate all vessels, failed to colonise at a level sufficient for detection via culture (less than 1.22 log_{10} CFU/mL) remaining below detectable levels for the remained for the experiment. Total anaerobes stabilised at 7.59 Log_{10} CFU/mL.

Obligate Anaerobes – Vessel 1

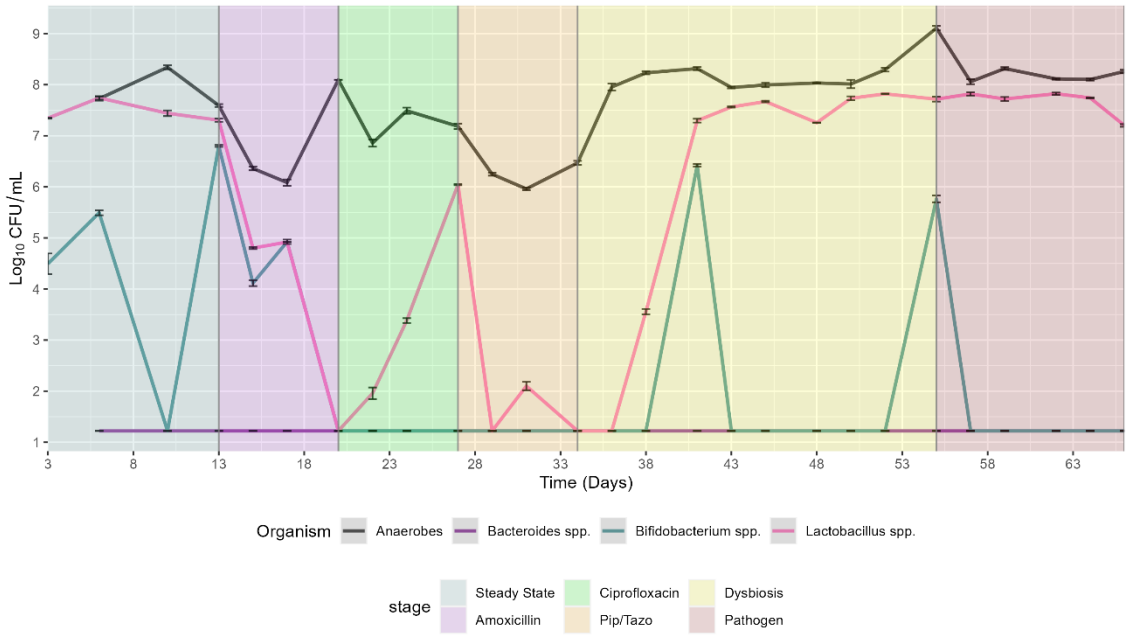

Figure 3-9 Obligate Anaerobes: Leeds HGM Vessel 1 (Proximal). Total Anaerobes (Blank) *Bacteroides spp*. (Purple) and *Bifidobacterium spp. (Green) Lactobacillus spp.* (Pink)

Amoxicillin instillation sees an initial decrease in total anaerobes in the proximal vessel decreasing to 6.08 Log_{10} CFU/mL. Bacterial abundances increase rapidly following the cessation of dosing, increasing to 8.10 Log_{10} CFU/mL, a net increase of 0.51 Log_{10} CFU/mL. The overall increase in abundance potentially being attributed to *LF Enterobacteriaceae*. *Bifidobacterium spp.* and *Lactobacillus spp.* both declined during amoxicillin instillation, falling below the limit of detection by the end of the active dosing and following the cessation of dosing.

Lactobacillus spp. recovered during ciprofloxacin instillation, reaching 6.04 Log_{10} CFU/mL by the end of the phase. *Bifidobacterium spp.* remained below

detectable levels. Despite the recovery in Lactobacillus spp., total anaerobes levels fell during the ciprofloxacin trend.

TZP saw reductions in total anaerobes, decreasing by 0.72 Log_{10} CFU/mL. *Lactobacillus spp.* fell below levels detectable via culture and both *Bifidobacterium spp.* and *Bacteroides spp.* remained undetectable.

After the cessation of all antibiotics, anaerobic spp. failed to reestablish to the same extent as facultative spp. *Lactobacillus spp.* returned to pre-exposure levels by day 41, 10 days after the final antibiotic dose. Transient detection of Bifidobacterium spp. was noted at days 41 and 55, however they failed to stabilise.

Obligate Anaerobes – Vessel 3

In the distal portion of the model, all monitored anaerobic populations were successfully established Figure 3.3-10 Obligate Anaerobes: Leeds HGM Vessel 3 (Distal).. The total anaerobic count reached 8.58 Log_{10} CFU/mL, with *Bacteroides spp*. at 5.51 Log_{10} CFU/mL, Bifidobacterium spp. at 7.02 Log_{10} CFU/mL, and *Lactobacillus spp*. at 7.59 Log_{10} CFU/mL.

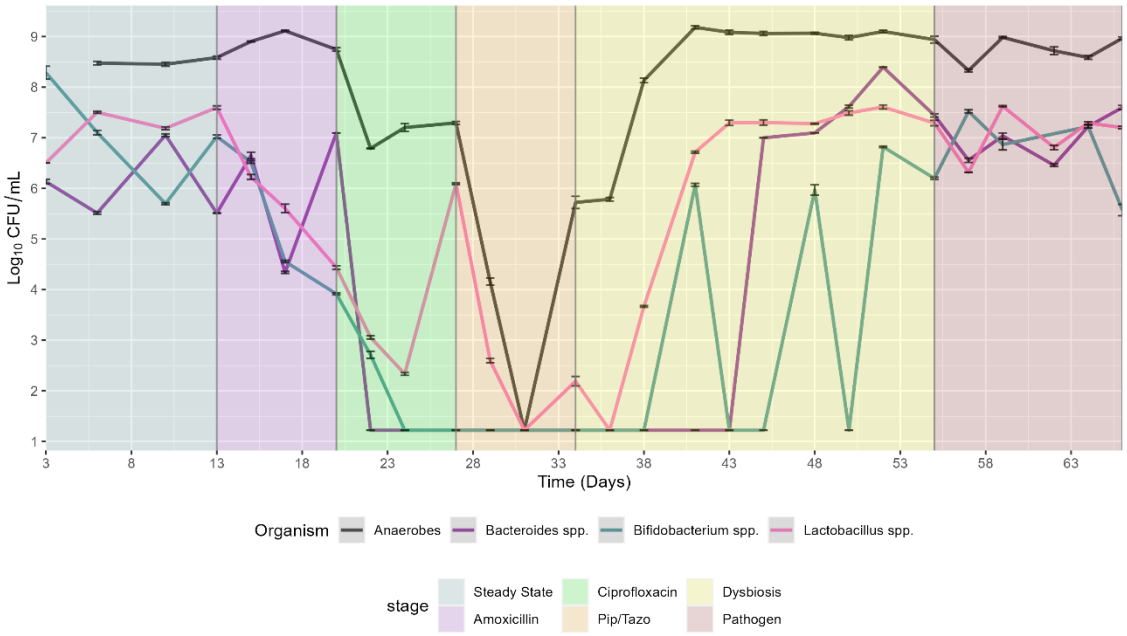

Figure 3.3-10 Obligate Anaerobes: Leeds HGM Vessel 3 (Distal). Total Anaerobes (Blank) *Bacteroides spp*. (Purple) and *Bifidobacterium spp. (Green) Lactobacillus spp.* (Pink)

Throughout the course of amoxicillin exposure, total anaerobe levels remained relatively stable, concluding at 8.74 Log_{10} CFU/mL. Both *Lactobacillus spp.* and *Bifidobacterium spp*. experienced consistent declines during the amoxicillin treatment, decreasing by 3.16 Log_{10} CFU/mL and 3.11 Log_{10} CFU/mL, respectively. Despite an initial drop to 4.34 Log_{10} CFU/mL during active dosing, *Bacteroides spp.* rebounded during the drug-free interval, reaching 7.10 Log_{10} CFU/mL before the commencement of the next treatment.

Ciprofloxacin dosing coincided with a net drop in the total abundance of anaerobes, falling 1.45 Log_{10} CFU/mL. *Bacteroides spp.* and *Bifidobacterium spp*. both drop below the limit of detection during active antibiotic dosing and remain at undetectable levels throughout the drug-free interval. *Lactobacillus spp.* initially fell in abundance by 2.1 Log_{10} CFU/mL however recovered during the drug-free interval resulting in a net increase of 1.66 Log_{10} CFU/mL over the period.

Bifidobacterium spp. and *Bacteroides spp.* remained below the limit of detection throughout TZP installation. *Lactobacillus spp. fell to undetectable levels during active dosing, recovering to 2.19* Log_{10} CFU/mL during cessation of dosing.

The Proximal microbiome largely recovers by day 45 with Total abundances of anaerobes, *Lactobacillus spp., and Bacteroides spp.* recovering to abundances like those seen prior to antibiotic instillation. Transient detection of *Bifidobacterium spp*. was recorded on days 41 and 48; however, they failed to stabilise until day 52, 21 days after the final antibiotic dose.

3.4.1.3 Predominant Species Diversity Estimates (MALDI-TOF)

At the conclusion of each experimental stage, species diversity estimates were made. Predominant visually distinct colonies were selected and analysed via MALDI-TOF biotyping Table 3-8 Predominant Bacterial species recovered by bacterial culture by agar media type Leeds HGM Vessel 1 + Vessel 3 (MALDI-TOF). While some populations, such as *Enterococcus spp.* and *LF Enterobacteriaceae*, remained consistent throughout the experiment, others experienced restructuring of predominant strains or reductions in species diversity. Notably, total anaerobes exhibited a decline in diversity following the instillation of Amoxicillin and Ciprofloxacin after which *Escherichia coli*

dominated the microbial community. This change persisted through the remained of the experiment.

Lactobacillus spp., species diversity decreased. During steady-state, *Lactobacillus* was primarily dominated by the species *gasseri plantarum*, *fermentum*, and *oris*. Following exposure to a*moxicillin* and c*iprofloxacin*, *Lactobacillus plantarum was the most common species detected by MALDI-TOF.*

Bifidobacterium spp. fell below the limit of detection following CIP dosing and, only one of the two strains identified by MALDI-TOF prior to antibiotic treatment recovered to detectable levels through culture.

Table 3-8 Predominant Bacterial species recovered by bacterial culture by agar media type Leeds HGM Vessel 1 + Vessel 3 (MALDI-TOF)

Target Organisms	Steady State	Day 5 Amoxicillin	Day 5 Ciprofloxacin	Day 5 Piperacillin/Tazobactam	Post Antibiotic Recovery
Total Facultative anaerobes	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*	*Escherichia coli* *Klebsiella pneumoniae*
Enterococcus spp.	*Enterococcus faecalis*	*Enterococcus faecalis*	*Enterococcus faecalis*	*Enterococcus faecalis*	*Enterococcus faecalis*
Lactose Fermenting Enterobacteriaceae	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*
Total Anaerobes	*Escherichia coli* *Bifidobacterium longum* *Lactobacillus gasseri*	*Escherichia coli* *Enterococcus faecalis* *Hungatella hathewayi* *Bacteroides ovatus*	*Escherichia coli*	*Escherichia coli*	*Escherichia coli*
Lactobacillus spp.	*Lactobacillus gasseri* *Lactobacillus plantarum* *Lactobacillus fermentum* *Lactobacillus oris*	*Lactobacillus oris* *Lactobacillus plantarum* *Lactobacillus rhamnosus*	*Lactobacillus plantarum*	*Lactobacillus plantarum*	*Lactobacillus plantarum*
Bifidobacterium spp.	*Bifidobacterium longum* *Bifidobacterium pseudocatenulatum*	*Bifidobacterium longum*	None recovered	None recovered	*Bifidobacterium longum*
Bacteroides spp.	*Bacteroides ovatus*	*Bacteroides ovatus* *Bacteroides faecis*	None recovered	None recovered	*Bacteroide faecis*
Total spores	*Hungatella hathewayi / effluvii*	*Hungatella hathewayi / effluvii*	*Clostridium tertium* *Hungatella hathewayi / effluvii*		*C.difficile*

3.4.2 Antimicrobials and Colonisation Resistance

3.4.2.1 *Clostridioides difficile*

C. difficile spore preparation was inoculated into the model on day 55; the model was sampled prior to introduction, and levels of *C. difficile* spores and vegetative cells were below detectable levels in the model less than 1.22 Log_{10} CFU. Detectable levels of the bacterium were observed in both the proximal and distal regions by day 56. Initially, the total viable counts (TVC) were 4.64 Log_{10} CFU/mL and 4.56 Log_{10} CFU/mL respectively (Presented in Figure 3-11).

In the proximal region of the model, *C. difficile* populations remain consistent until day 63. The consistent presence of *C. difficile* spores alongside stable total viable counts (TVC)

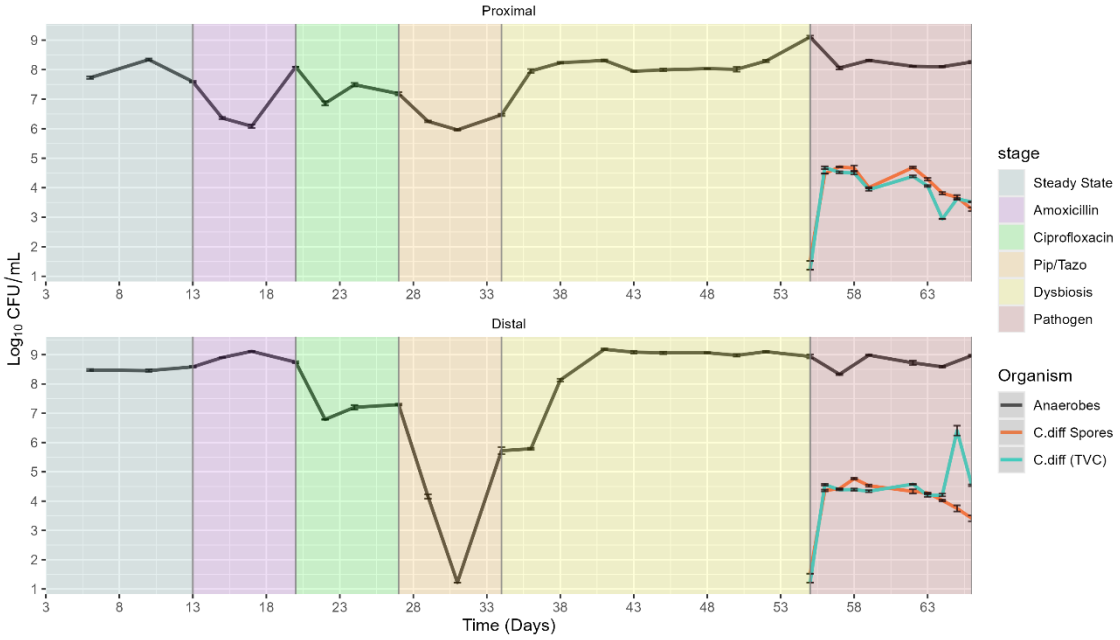

Figure 3-11 Leeds HGM - Clostridioides difficile Spore (Orange) and Total Viable Counts (TVC) (Blue) for Proximal (Top) and Distal (Bottom)

Distally there is differentiation of Spore and TVC counts on day 64, By day 66 spore counts were 3.27 Log_{10} CFU/mL and TVC counts 4.54 Log_{10} CFU/mL.

CCNA Toxin Assay

Active *C. difficile* infection of the model was confirmed with the presence of *C. difficile* toxin being detected on days 65 and 66.

3.4.2.2 Klebsiella pneumoniae

K. pneumoniae KPC– overnight culture added on day 55; prior to these no detectable levels of KPC-positive *K. pneumoniae* were detectable within the model less than 1.22 Log_{10} CFU/mL. The presence of KPC-positive *E. coli* was also undetectable prior to inoculation (presented in Figure 3-12).

In the proximal region, *K. pneumoniae* (KPC) levels stabilised at approximately 3 Log10 CFU/mL before expanding; by day 66, levels had reached 6.48 Log10 CFU/mL. On days 56 and 59, populations of KPC positive *E. coli* were recorded.

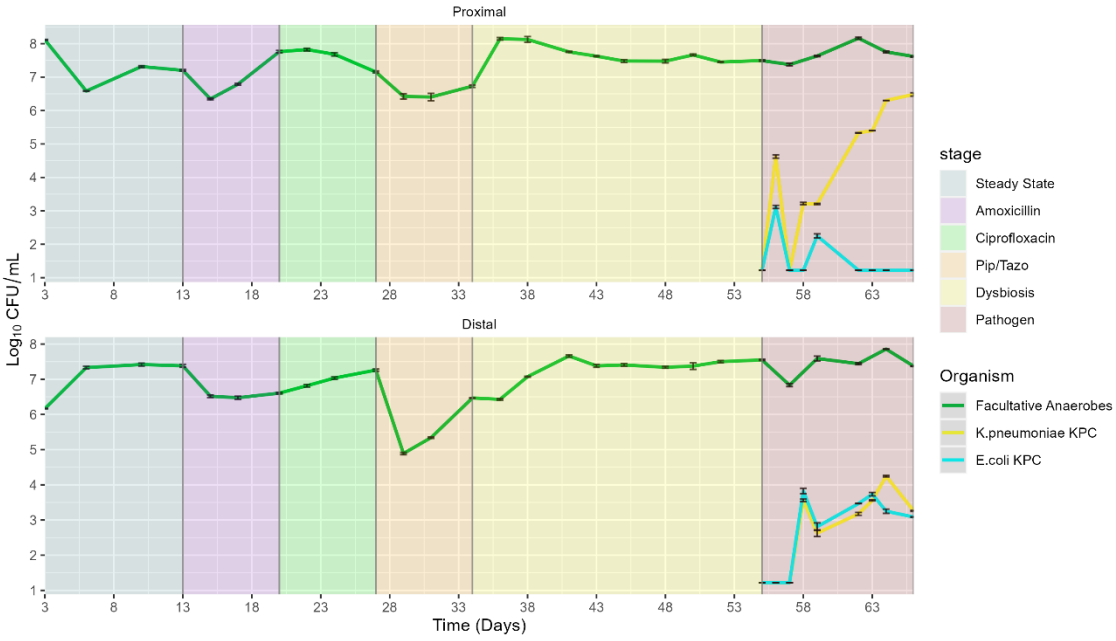

Figure 3-12 Leeds HGM - Detection of KPC positive *K. pneumoniae* (yellow) and *E. coli* (blue)

Distally, KPC-positive strains remained undetectable until day 58

3.1 Discussion

The purpose of the models described in this chapter is to establish a healthy *in vitro* microbiome derived from human populations, which could be utilised to simulate the long-term impacts of repeated antibiotic exposure, which are commonly seen in patients with persistent or recurrent infections. Such models offer significant advantages, including the ability to tightly control experimental conditions, isolate specific variables and conduct experiments without the complex ethical considerations associated with human or animal studies.

These models are crucial as they underlie the work in the following chapters as they serve to provide samples for metabolomic and metagenomic profiling, enabling exploration of the intestinal microbiome from both taxonomic and functional perspectives. Additional samples from this model as serve as material for method development purposes described in the preceding chapter.

A challenge inherent to the Leeds HGM is the resource and labour intensity required, which limits the number of replicates that can be performed at once. Running a single replicate in the context of the experiment outlined in this chapter provided several benefits:

- The vessel volume of the Leeds HGM means that additional aliquots can be taken without significantly impacting the bacterial populations.
- Culture-based sampling of the populations, whilst limited in its scope, can provide useful insights into the overall health of the model, as well as insight into the compositional changes occurring within the microbiome in a contemporaneous manner.

The MiGut system's scalability has been optimised for omics approaches, and as a consequence, concessions are made which allow for limited culture to be performed before potentially seeing adverse effects on the stability of the system.

Weekly quantitative RT-PCR was performed throughout the study on both the Leeds HGM and MiGut platforms as part of system validation. The results confirmed consistency between the two platforms (Davis Birch et al., 2023).

On this basis, culture results can be used to draw some interesting initial insights into what is potentially occurring within the microbiome, as well as being confident in the success of goals before investing in next generation sequencing and metabolomic profiling. Utilising both models together allowed for culture and omics approaches to be effectively combined leveraging the strengths of both approaches. The Additional replicates possible with the MiGut platform further increase experimental robustness in further analysis. In the following sections, experimental goals will be outlined, with an explanation of how the data suggests their achievement and the rationale for progressing confidently.

3.1.1 Establishing Baseline Microbiome

Based on previous experience working with the Leeds HGM and MiGut platforms, it is understood that an initial 2-week steady-state period is sufficient for bacterial populations to stabilise. During the steady-state period, stable detection of monitored populations is expected. A collapse of any key species or the emergence of species typically not associated with the intestinal microbiome could indicate a failure to establish a healthy baseline or contamination. These issues would prompt a discontinuation and restart.

Culture data points towards the establishment of a healthy microbiome. Culturable Bacteroides spp. failed to establish at detectable levels within Vessel 1. This is possible because the pH range of the proximal region is sub-optimal for the proliferation of *Bacteroides spp*. *Bacteroides spp.* have been shown to establish at lower abundances in co-cultures at pH levels less than 6; even when present as a dominant strain within the starting inoculum, they are outcompeted by other minority strains. Bacteroides begin to dominate co-cultures towards neutral pH (Ilhan et al., 2017). This is a probable cause for the difference in abundance of *Bacteroides spp.* recovered from proximal and distal regions.

Medial and distal regions (V2 and 3) are potentially more relevant to clinical symptoms of CDI, displaying the greatest colonisation and germination by *C. difficil*e and highest toxin production. Researchers are increasingly appreciating

that taxonomic delineations of the microbiome are insufficient to explain dysbiotic states. Thus, spatial differentiation of the intestinal microbiome in composition and functional capacity may be more important. Upstream metabolic outputs in the proximal regions could potentially compound alterations seen along the colon.

3.1.2 Impacts of Antimicrobials on the Gut Microbiome

The Leeds HGM has an extensive history of investigating the propensity of antimicrobials to induce antibiotic-associated dysbiosis, particularly concerning its potential to lead to *CDI (Freeman et al., 2003; Baines, S. D. et al., 2005; Baines, Simon D. et al., 2009; Chilton, Caroline H. et al., 2012; Buckley et al., 2021)*

AMX is a broad-spectrum antibiotic in the penicillin class, primarily used to target Gram-positive bacteria. AMX has been studied in the Leeds HGM for its propensity to illicit CDI following a 7-day dosing regimen of amoxicillin clavulanic acid (co-amoxyclav) (AMC) (Chilton, Caroline H. et al., 2012). AMC was shown to have a less deleterious effect than other antimicrobials associated with CDI whilst still being able to induce CDI. Chilton et al demonstrated that AMC impacted the microbiome, reducing the abundance of *Bacteroides spp*. and Bifidobacterium spp. and resulting in increased abundance of facultative anaerobes. Interestingly, trends observed in *Enterococcus spp*. differ, previously, AMC was shown to be associated with an increase in *Enterococcus spp*. within the model (Chilton, Caroline H. et al., 2012). In the present study, AMX instillation coincided with reductions in Enterococcus spp. in vessel 1 and vessel 3 of the model. This may be due to differences in microbiome composition between pooled faeces inoculum. The two most common *Enterococcus* strains in the human intestinal microbiome are *E. faecalis* and *E. faecium,* both strains are regularly isolated from the Leeds HGM through culture. These strains show differing susceptibility profiles to AMX, with *E. Faecalis* being susceptible and *faecium* displaying increased resistance (Schouten et al., 1999). Throughout this run, faecalis was identified as the predominant *Enterococcus spp*. across all time points via MALDI-TOF

Table 3-8 Predominant Bacterial species recovered by bacterial culture by agar media type Leeds HGM Vessel 1 + Vessel 3 (MALDI-TOF). The pooling of faecal samples used in seeding the models is done with the purpose of mitigating variations between runs, but there may still be variations in dominant species between runs. The previous observations of AMX in the Leeds HGM do not have published species-level identification available; this reinforces the need for greater resolution than genus-level identification for characterising the microbiome particularly when de-risking antibiotic selection.

Ciprofloxacin (CIP) is a board spectrum fluoroquinolone with activity against Gram-positive and Gram-negative bacteria. Observations from the present study are comparable to previous observations in the Leeds HGM, with marked reductions in Bacteroides spp. and *Bifidobacterium spp.* falling below the limit of detection during CIP instillation and Lactobacillus spp. and Enterococcus spp. undergoing smaller decreases in abundance (Saxton et al., 2009). Proximally, in vessel 1, some recovery of *Lactobacillus spp*. is seen. However overall populations are still suppressed compared to levels during steady state. CIP instillation coincides with reductions in culturable species diversity in total anaerobes and *Lactobacillus spp.,* with both groups being dominated by a single species.

Piperacillin-Tazobactam (TZP) is a combination antibiotic consisting of piperacillin, a broad-spectrum penicillin, and tazobactam, a beta-lactamase inhibitor. This combination provides a broad range of activity, targeting both Gram-positive and Gram-negative bacteria. TZP's impact on the microbiome and its potential to act as a driver for the spread of Carbapenemase-producing Enterobacteriaceae (CPE) within the microbiome has been previously studied in the Leeds-HGM. Rooney *et al.* (Rooney et al., 2019) demonstrated that the selective antibiotic selective pressure of TZP was associated with the transfer of antibiotic-resistance genes from *K. pneumoniae* to *E. coli*. In the present study, TZP induced marked reductions in all monitored populations mirroring what had been observed previously (Harris et al., 2021). Species which had previously fallen below the LOD, such as *Bifidobacterium spp.* and *Bacteroides spp*. remained suppressed throughout TZP installation.

All three antibiotics in the present study were able to induce marked changes in the intestinal microbiome, in some instances compounding the effect of the previous, reducing overall bacterial abundance and species diversity.

3.1.3 Recovery

In each antibiotic phase, there is a two-day drug-free interval. During this interval, it is anticipated that the concentration of the antibiotic within the model may decline to levels potentially below the therapeutic threshold. This pause allows for a brief recovery period before the initiation of the subsequent antimicrobial treatment. Consequently, it may be expected to observe an increase in certain bacterial populations at sampling points immediately prior to the next treatment. The duration of this window depends on factors such as the retention concentration, and half-life of the antibiotic in question and the susceptibility of the bacterial populations.

Following the cessation of AMX treatment, *Bacteroides* populations recovered quickly to pre-dosing levels, with the emergence of *Bacteroides faecis* as a prominent strain identified via MALDI-TOF.

The period following cessation of CIP instillation was characterised by an increase of Lactobacillus populations but with a marked reduction in species diversity. Following CIP instillation, culturable strains had been reduced to a single species, *Lactobacillus plantarum, recoverable by culture*.

There was some recovery of *LF Enterobacteriaceae* during TZP instillation. This may indicate the selection of resistant subpopulations within the model, although this was not specifically measured here. Previous work demonstrated that TZP dosing alone is sufficient to drive the emergence of resistance genes within the system following six days of instillation (Rooney et al., 2019). In the present study, there were no detectable CPE colonies recovered at screening prior to KPC introduction on day 55, so it is unclear whether this recovery could be attributed to the emergence of resistance within the model.

Following TZP instillation, there was a 3-week period without intervention. This was implemented to allow for the microbiome to reestablish and reach a new homeostatic equilibrium. The purpose of this was to assess the degree to which

the microbiome was able to recover, and its potential to resist colonisation and expansion of exogenous pathogens following a lengthy recovery period. Previous *in-vitro* studies with the Leeds-HGM exposed to *C. difficile* in the absence of antimicrobials have shown that in instances of intact colonisation resistance, spores remain quiescent, failing to establish within the model (Baines, S. D. et al., 2005; Chilton, Caroline H. et al., 2012) reaffirming that the expansion of *C. difficile* within the model can be attributed to functional dysbiosis and a lack of colonisation resistance.

Facultative aerobe populations returned to pre-antibiotic exposure levels by days 7 and 10 post-antibiotic treatment for the proximal (Vessel 1) and distal regions, respectively. The prolonged recovery time for distal populations may be attributed to increased retention time within the system and the gradual clearance of antibiotics following the final dose.

Anaerobic populations recovered by days 10 and 14 post-final antibiotic dose. Similar to facultative populations, the delayed recovery of distal anaerobic populations could be linked to the model's retention time and the gradual clearance of antibiotics within the system.

The recovery of proximal microbiota (in vessel 1) was not as robust as that observed in the distal vessel (vessel 3). During the dysbiosis and subsequent recovery period, transient detection of *Bifidobacterium spp*. was noted. However, these populations failed to stabilize within the experimental timeframe, leaving uncertainty regarding their potential for resolution given additional time. The transient presence of *Bifidobacterium spp*. may be attributed to biofilm detachment and subsequent seeding within the model, a phenomenon previously studied in the context of the Leeds-HGM system (Normington et al., 2021). Although this experiment did not explicitly incorporate biofilm structures, microbial biofilms are known to form on the surfaces of fermentation vessels. Transient detection of *Bifidobacterium spp*. was also observed in the distal vessel (vessel 3) 10 and 17 days post-final antibiotic dose, with consistent detection occurring by day 21 post-antibiotic, persisting until the end of the experiment indicating more established recovery.

Despite the microbiome largely returning to pre-antimicrobial abundances and composition, a notable reduction in species diversity persisted throughout the experiment. Nearly all targeted groups assayed by culture became dominated by a single strain.

3.1.4 Pathogen Expansion

C.difficile was detected in both proximal (vessel1) and the distal (vessel 3) on day 56, one day after inoculation.

In the proximal region (vessel 1), *C. difficile* remained predominantly in spore form, exhibiting a gradual decline towards the end of the two-week monitoring period. The consistent presence of C. difficile spores alongside stable total viable counts (TVC) implies a potential absence of active germination and passive colonization processes.

The distal region (vessel 3) showed a steady detection of *C. difficile*, predominantly as spores, until day 64, with spore and vegetative cell counts remaining within 1 $\log_{10}$ CFU/mL. From day 64, there was a notable reduction in the recovery of spores and an increase in vegetative cells. This observed increase in vegetative $\log_{10}$ CFU/mL, accompanied by a decrease in spore $\log_{10}$ CFU/mL, is indicative of germination and active growth. Differentiation of *C. difficile* spores and vegetative cells coincide with the detection of *C. difficile* toxin within the model assayed via semi-quantitative CCNA. The divergence of Spore and TVC counts starting day 64 coupled with the present of *C. difficile* toxin suggests germination of spores and active colonisation of the model trending towards simulated CDI. Germination of *C. difficile* spores and active production within the model indicated simulated *C. difficile* infection within the system.

Klebsiella pneumoniae (KPC) was detected on day 56 in the proximal vessel (Vessel 1) and on day 58 in the distal vessel (vessel 3) following it’s addition on day 55.

The potential spread of KPC genes to *Escherichia coli* (*E. coli*) was indicated by transient detection in the vessel 1 on days 56 and 59, and consistent presence in the distal vessel (vessel 3) from day 58 onwards.

Previous studies (Harris et al., 2021) (Rooney et al., 2019) have investigated the introduction of KPC to similar models and its potential to disseminate resistance to other Enterobacteriaceae, with results aligning closely with the observations made in this study. While antibiotic resistance genes were not explicitly screened for in the present study, the emergences of KPC-positive E.coli stains on antibiotic resistance screening plates suggest the transfer of resistance from inoculated KPC-positive *K. pneumoniae* to the endogenous microbiota. The detection of KPC-positive *K. pneumoniae* and *E. coli* remained steady within the distal region of the model until the end of the experiment. In contrast, in the proximal vessel (vessel 1), KPC rapidly expanded and became a major constituent of the total facultative anaerobes recovered by culture.

3.2 Summary

This chapter demonstrates that dysbiosis can be achieved in an *in-vitro* model using a clinically relevant antibiotic dosing regimen, leading to a persistent disturbance to the gut microbiota. This is evidenced by the disruption to key microbial genera, reduction in dominant species diversity and the expansion of introduced pathogens explored by culture-based methods.

Chapter 4
Metagenomic investigation of the *In-vitro* Intestinal microbiome.

4.1 Introduction

Antibiotic therapy induced significant disruption in microbial populations within the in-vitro gut model. While isolation and enumeration of key taxa by culture revealed the recovery of the bacterial composition of selected genera to pre-antibiotic levels, the colonisation and proliferation of *C. difficile* and *K. pneumoniae* within the model following the three-week recovery period indicated a sustained loss of colonisation resistance. The lack of colonisation resistance, despite the return of monitored taxa to their original composition, suggests that culture-based methods may not effectively capture these changes.

This chapter employs shotgun metagenomics for taxonomic and functional analysis to provide a deeper understanding of the microbiome beyond what can be achieved with culture-based methods alone. By utilising metagenomic in this context, the aim is to offer deeper insights into the global taxonomic shift and inferred metabolic pathways of the microbiome. Metabolic pathway analysis will provide a basis for subsequent metabolomic investigation.

4.2 Chapter Aims

This chapter seeks to provide a deeper understanding of the microbiome's compositional and functional dynamics and their response to antimicrobial treatment. Specifically, the aims are:

1. **Explore the Compositional Dynamics of the Microbiome**
    - Investigate the microbiome with greater resolution afforded by metagenomics, encompassing a broader range of bacterial genera, including non-culturable species.
    - Examine species diversity with bacterial genera to understand how it may be altered by antimicrobial treatment.

2. **Perform Functional Analysis Beyond Population Dynamics**
    - Determine whether dysbiosis characterised in this instance by the failure to resist the colonisation and expansion of pathogens can be partially attributed to a loss of functional potential or if it is more likely due to downstream effect on the transcriptome of the proteome.

4.3 Experimental Outline

Metagenomics analysis utilises samples taken from the in-vitro model previously described. Samples were taken from the Leeds Human Gut model and MiGut platforms. Sample time points are outlined in Table 4-1 Metagenomic Sample Timepoints for Leeds HGM and MiGUT. If possible, sample days were chosen on the final day of each experimental phase, immediately before the commencement of the next antimicrobial installation or pathogen inoculation. In instances where alternative sample days were chosen, shifts are indicated, for example, from Day 34 to Day 36, denoted by a change of +2.

Table 4-1 Metagenomic Sample Timepoints for Leeds HGM and MiGUT. Parentheses indicate alternative sample day shifts.

Stage	Leeds HGM		MiGut	
	Vessel 1	Vessel 3	Vessel 1	Vessel 3
Steady - State	Day 10	Day 10	Day 10	Day 10
Amoxicillin	Day 20	Day 20	Day 20	Day 20
Ciprofloxacin	Day 27	Day 27	Day 27	Library Failure
Piperacillin / Tazobactam	Day 36 (+2)	Day 36 (+2)	Library Failure	Library Failure
Post–Abx Recovery	Day 55	Day 55	Day 55	Library Failure
Post - Pathogen	Day 64	Day 64	Day 64	Day 64

Samples for metagenomics were taken from the Leeds HGM and a single iteration of the MiGUT platform. This decision was made to ensure sufficient depth within the metagenomic analysis. Reviewing 16s data from Leeds HGM and all MiGut iterations confirmed microbiome structures across the models were following similar trajectories (Figure 4-1 NMDS Ordination of 16s sequencing from Leeds HGM and MiGut platforms.)

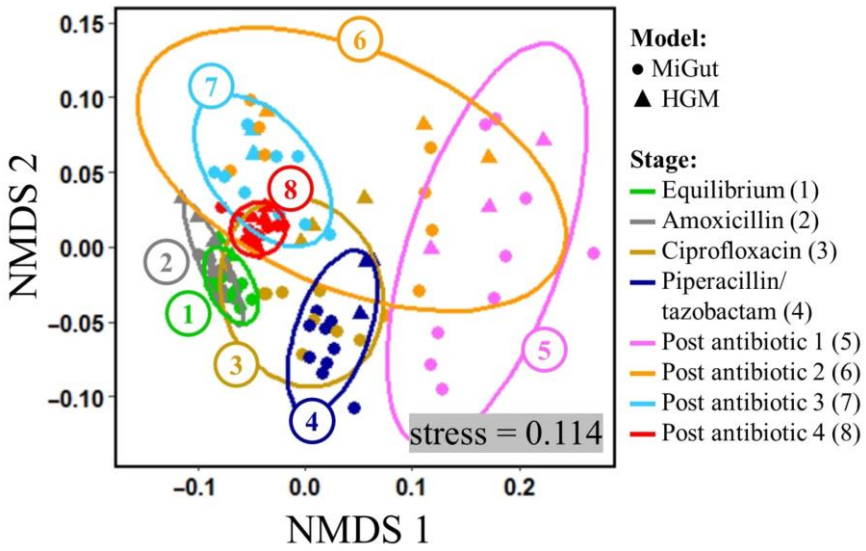

Figure 4-1 NMDS Ordination of 16s sequencing from Leeds HGM and MiGut platforms. Image reproduced from MiGut: A scalable in vitro platform for simulating the human gut microbiome-Development, validation, and simulation of antibiotic-induced dysbiosis (Davis Birch et al., 2023).

With these observations from the 16s DNA analysis in mind, confidence was gained that metagenomic observations would be representative of the model cohort and be sufficient to support and contextualise future metabolomic insights.

4.4 Methods

4.4.1 DNA Extraction

Samples for DNA extraction were immediately frozen at -80°C upon collection for batch processing at the conclusion of the experiment. DNA extraction was conducted from pellets obtained by centrifuging 1 mL of Gut Model fluid at 16,000 × g for 15 minutes. The pellets were processed using the FastDNA Spin Kit for Soil (MP Biomedicals) according to the manufacturer's instructions. Extracted DNA preparations were subsequently stored at -80°C until being dispatched for off-site analysis.

For MiGut samples, the supernatant was reserved for metabolite extraction as outlined in 2.4 Sample Processing

4.4.2 Bioinformatics software

Quality Control

FastQC (v0.11.9) performs quality control analysis on raw sequencing reads, evaluating key metrics such as quality scores, sequence length distribution, and adapter content.

MultiQC (v1.12) (Ewels et al., 2016) Aggregates multiple FastQC reports into a comprehensive summary, facilitating quick assessment and comparison of quality control metrics across various samples.

Read Trimming

Trimmomatic (v0.39) (Bolger et al., 2014) Removes adapter sequences and low-quality bases from raw sequencing reads, enhancing data quality and reliability for downstream analysis. Trimmomatic was used in Single End (SE) and Paired End (PE) palindromic modes. Trimmomatic was used with the Default parameters for processing Illumina reads with the dropping of reverse reads off

Read Merging

PEAR (v0.96)(Zhang et al., 2014): Paired-end read merger, used to merge overlapping paired-end reads generated from next-generation sequencing

platforms, improving the accuracy and length of assembled sequences. PEAR was used with default parameters

Alignment

DIAMOND (2.0.8) (Buchfink et al., 2015; Bağcı et al., 2021) is a high-speed sequence alignment tool optimised for use on both workstations and high-performance computing environments. Alignment was performed against the NCBI non-redundant protein sequence database (NCBI-nr database), accessed on May 3, 2023. Using default Parameters.

Meta-Genome analyser

MEGAN6 (6.22.2) (Bağcı et al., 2021)

R Packages used

- DESeq2 (v1.44.0): Utilised for differential analysis (Anders and Huber, 2010).
- dplyr (v1.1.4): Tools for data manipulation, focusing on data frames.
- ggplot2 (v3.5.1): Used for creating visualisations.
- Patchwork (v1.2.0): Facilitates the combination of multiple ggplot2 into a single figure.
- Tidyverse (v2.0.0): A collection of R packages utilised in data science, including ggplot2, dplyr.
- Vegan (v2.6-6.1): Employed for community ecology analyses; in this chapter, Vegan's distance matrix function is used for 2D ordination of the microbiome.

Metagenomic pathway analysis

Pathway analysis was done with Kyoto Encyclopaedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000; Kanehisa et al., 2023) 111.0, July 1, 2024

Additional information for metabolic pathways, including reaction direction, obtained from MetaCyc (Caspi et al., 2014)

4.4.3 Sequencing

The library preparation and Next Generation Sequencing (NGS) were carried out by the Next Generation Sequencing Facility at the University of Leeds. The sequencing process was executed on a single lane, using a 300-cycle kit to generate 150 base pair paired-end reads on a NextSeq 2k platform. Sequencing resulted in an average read depth of 27 million reads per sample.

4.4.3.1 Metagenomic Controls

Negative Control: Nuclease-free water taken from the sample plate was processed along experimental samples to monitor for contamination.

Quality control: Illumina PhiX was included as a quality control. PhiX is a 500bp library without indexes and balanced GC:AT ratio, included to identify over/under clustering and flow cell cluster efficiency.

4.4.3.2 Library preparation

After shipment for processing, some of the libraries failed to create libraries of sufficient quality to proceed to sequencing. Eight of these libraries, covering phases of anti-microbial instillation, experienced inefficient sheering during library creation. This issue was believed to be caused by the presence of antifoam added to the model to prevent foaming within the vessels, which interfered with DNA extraction and downstream processing. The issue coinciding with antimicrobial instillation is likely due to antimicrobial-induced cell death resulting in the release of intracellular protein, which in turn increased foaming and necessitated additional antifoam. Re-cleaning samples with ethanol precipitation overnight with 3M sodium acetate was sufficient for effective sheering.

Some libraries contain significant portions of adapter dimers, which are self-ligations of adapter sequences without library inserts. These preferentially bind to the flow cell with high clustering efficiency (Head et al., 2014). Libraries following TZP dosing failed to make libraries due to the presence of an adapter dimer. Consequently, the initial sample point Day 34 had to be dropped for both MiGUT and Leeds HGM platforms. Additional samples had been taken throughout the experiment for Leeds HGM which was not possible with the MiGUT as it would deplete the system. One of the additional Leeds HGM

samples was chosen to cover the sample point. The next available sample point for TZP was day 36. Since no further antimicrobial instillation followed TZP, this was deemed an appropriate concession to have a point covering this model phase with the caveat that the microbiome had an additional two days to recover.

Ciprofloxacin (day 27) and Post–Antibiotic Recovery (day 55) failed to make Libraries due to the presence of Adapter dimers, believed again to be an issue of antifoam accumulation leading to ineffective DNA extraction resulting in low starting material for Library generation.

With these issues addressed, to the best extent, the 20 samples that generated libraries cover all experimental phases for the Leeds HGM, with the majority of time points having supporting observations in the MiGUT platform.

4.4.4 Preparing Paired End Reads for MEGAN

For the analysis of sequencing data in this study, DIAMOND was used for alignments due to its integration with MEGAN6 and optimisations for workstation use. It is recommended that single reads be used when working with DIAMOND. It is possible that in cases where paired-end data is available to utilise both, each read pair is treated as a single read to maintain consistency in the analysis process. Given access to paired-end data, the preference is to utilise both reads to maximise the accuracy of sequence assignments whenever feasible.

In determining the appropriate processing pipeline to employ, three approaches were evaluated (Figure.4-2 Pre-processing pipelines considered for Metagenomic data preparation. Steady-state samples were the focus of this comparison. These pre-exposure samples to antibiotics are presumed to exhibit the highest degree of diversity and complexity.

Pipeline A involves a quality control procedure for the forward reads, entailing the trimming of adapter sequences and removing low-quality reads using Trimmomatic SE in conjunction with supplied Illumina adapter sequences.

Pipeline B adopts a simple strategy by concatenating the forward and reverse reads into a single FASTA file, leveraging all generated sequences. The quality

control steps employed in this pipeline mirror those utilised in Pipeline A to ensure data integrity and reliability.

Pipeline C incorporates Trimmomatic PE for palindromic adapter sequence removal, which offers increased specificity compared to SE trimming. The QC (Quality Control) for this step generates four output files: paired sequences from Read 1 and Read 2 and unpaired sequences from both reads. Paired sequences are merged using PEAR to create contigs, aiming to produce longer sequences for more precise alignments. Unpaired sequences from Read 1 and Unassembled sequences from Read 1 are concatenated with the file to utilise as many sequences as possible. However, corresponding sequences from Read 2 are omitted to avoid over-representing assignments from unpaired reads, ensuring the accuracy and validity of the analysis results

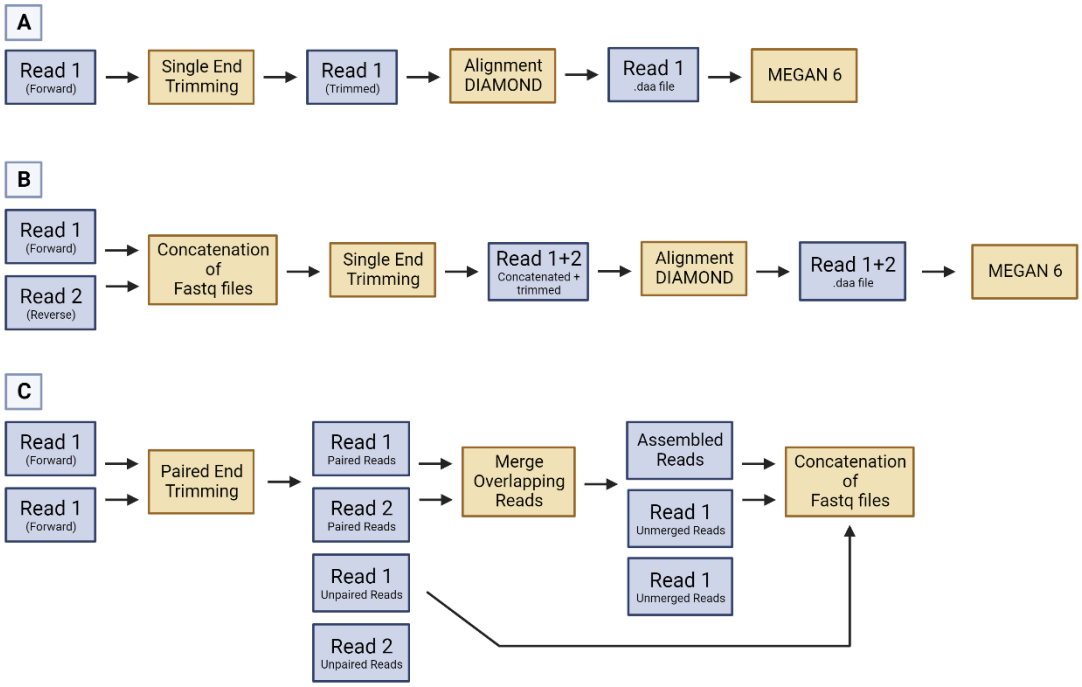

Figure.4-2 Pre-processing pipelines considered for Metagenomic data preparation. (A) Processing workflow utilising only forward Reads (B) Simple concatenation of forward and reverse reads (C) Merging overlapping segments between forward and reverse reads to increase overall sequence length.

.

4.4.4.1 DIAMOND alignment comparison

Pipeline A, using sequences from read 1, was used as a baseline for comparisons for aligned queries and the total time for DIAMOND to complete. Utilising read 1 only resulted in a processing time of 9.2 and 7.6 hours and aligned 24511880 and 20756956 queries for MiGUT and Leeds HGM samples, respectively.

Table 4-2 Queries Aligned in DIAMOND with Processing Pipelines

Reactor	Processing Pipeline	Alignment Time (hours)	Queries aligned
MiGut Distal	Read 1Only	9.2	24711880
	Read 1 + Read 2 Concatenated	22.4	49512423
	Read 1 + Read 2 Merged	12.8	25084436
Leeds HGM Distal	Read 1Only	7.6	20756956
	Read 1 + Read 2 Concatenated	18.56	49512423
	Read 1 + Read 2 Merged	10.5	21127174

Pipeline B, which utilises all sequences from Read 1 and Read 2 files, expectedly took the longest of the pipelines reviewed and resulted in the most queries aligned. For both platforms, the total time increased by approximately 144%, resulting in a 100% and 138% increase in alignments for MiGUT and Leeds HGM, respectively.

Pipeline C, which merged sequences from Read 1 and Read 2, showed an increased alignment time of 37% and 39%, resulting in an increase of queriers aligned by alignment by 1.5% and 1.7% for MiGUT and Leeds HGM, respectively.

The increase in alignments from pipeline A is unlikely to result in a meaningful change to the data, as concatenating Reads 1 and 2 essentially doubles the amount of data, resulting in each sequence being aligned twice. Pipeline C shows a minor increase in reads for both model platforms. By increasing read length, the specificity for alignment is anticipated to improve. With this in mind, NGS data will be processed using this method.

4.4.5 Microbiome Data Processing and Statistical Analysis

4.4.5.1 Distance Matrix Computation Methods

The dissimilarity between samples was calculated through a Bray-Curtis similarity matrix. This measure, in contrast to Shannon diversity, considers both species present and abundance within samples. This matrix is used to quantify differences between different samples based on genus composition. Distance matrixes were calculated with the `vegdist()` function within the R package Vegan.

4.4.5.2 Genus Ordination

Non-metric multidimensional scaling (NMDS) was employed to visualise the relationships between samples. This analysis utilised the Bray-Curtis dissimilarity matrix previously outlined. NMDS was performed using the `metaMDS()` within Vegan, which reduces the dimensionality of complex data sets to represent them in a two-dimensional space.

4.4.5.3 Differential abundance analysis

Differential abundance analysis was conducted to compare genus and functional data across experimental phases against the steady state baseline. Taxonomic and functional exported from MEGAN were prepared and analysed using DESeq using the `DESeqDataSetFromMatrix()` function.

4.5 Results

4.5.1 Taxonomic analysis

4.5.1.1 Microbiome Compositional Shifts as a Result of Antimicrobial Instillation

A pairwise distance matrix was computed between all Leeds HGM and MiGUT samples, utilising the Bray-Curtis dissimilarity index. Non-metric Multidimensional Scaling (NMDS) was then performed, resulting in a stress value of 0.1004; this indicates a reliable representation of the data in reduced dimension. Plotting microbiome data in this manner provides an intuitive assessment of beta diversity, illustrating how the composition of the intestinal microbiome responds to antimicrobial intervention.

Vessel 1 - Compositional Changes

Significant shifts in the microbiome composition within vessel 1 were observed for both Leeds HGM and MiGut Platforms (presented in Figure 4-3 Genus Level Compositional Changes in Vessel 1)

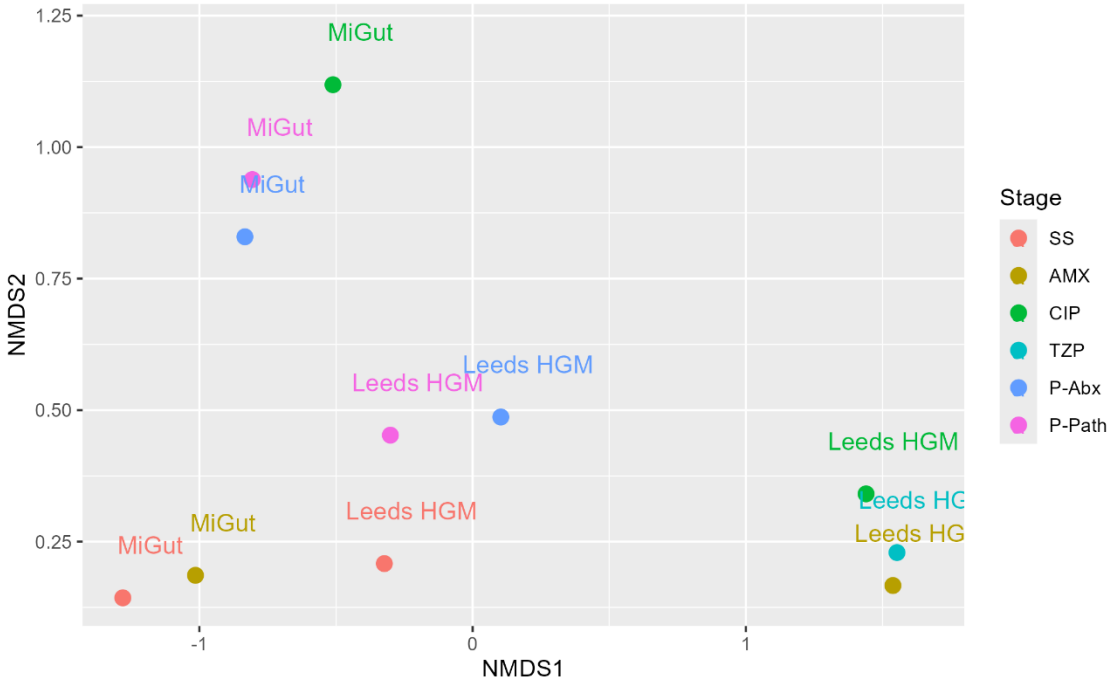

Figure 4-3 Genus Level Compositional Changes in Vessel 1 . NMDS plot generated with Bray-Curtis dissimilarity. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

In the Leeds HGM, TZP clusters with AMX and CIP, the more minor changes evident at this experiment phase may be attributed to the microbiome already being severely depleted.

Both models recovered towards steady state compositions following the three-week post-antibiotic recovery period. In this instance, recovery in the Leeds HGM appears to be more complete than in the MiGut.

Further changes were induced in both models following pathogen introduction and potentially attributed to colonisation and proliferation of *C. difficile and K. pneumoniae* within the model.

Vessel 3 - Compositional Changes

The distal portion of the models also displayed a compositional shift to the microbiome in response to antimicrobial instillation present in Figure 4-4 Genus Level Compositional Changes in Distal Colon.

AMX induced less pronounced changes in the Leeds HGM in Vessel 1 of the model than in Vessel 3, exhibiting compositional shifts more closely aligned to those seen in the MiGut.

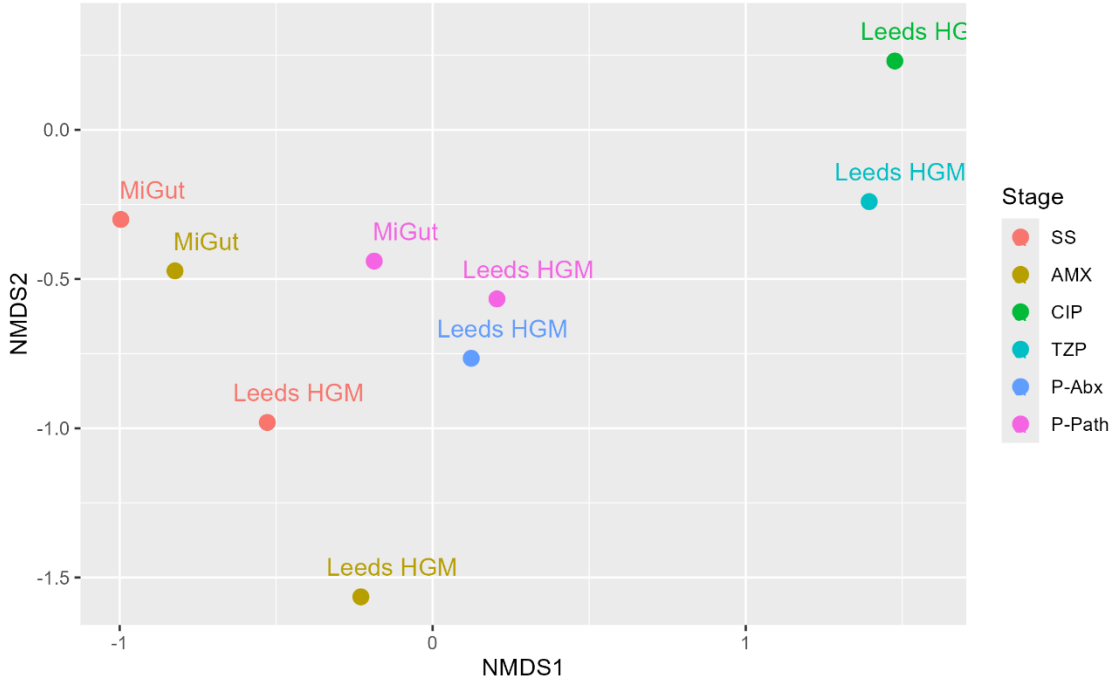

Figure 4-4 Genus Level Compositional Changes in Distal Colon. NMDS plot generated with Bray-Curtis dissimilarity. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Restoration of Disrupted Microbiomes Trend towards Original spatial Composition

Repeated exposure to antibiotics in both Vessel 1 and Vessel 3 of the in-vitro model drives the microbiome towards a distinct dysbiotic taxonomic profile (Figure 4-5 Genus Level Composition alterations, Vessel 1 vs). This profile appears to be shared between proximal (vessel 1) and distal (vessel 3) regions, which, prior to antibiotic instillation, appear to have unique ecologies distinct to their spatial location.

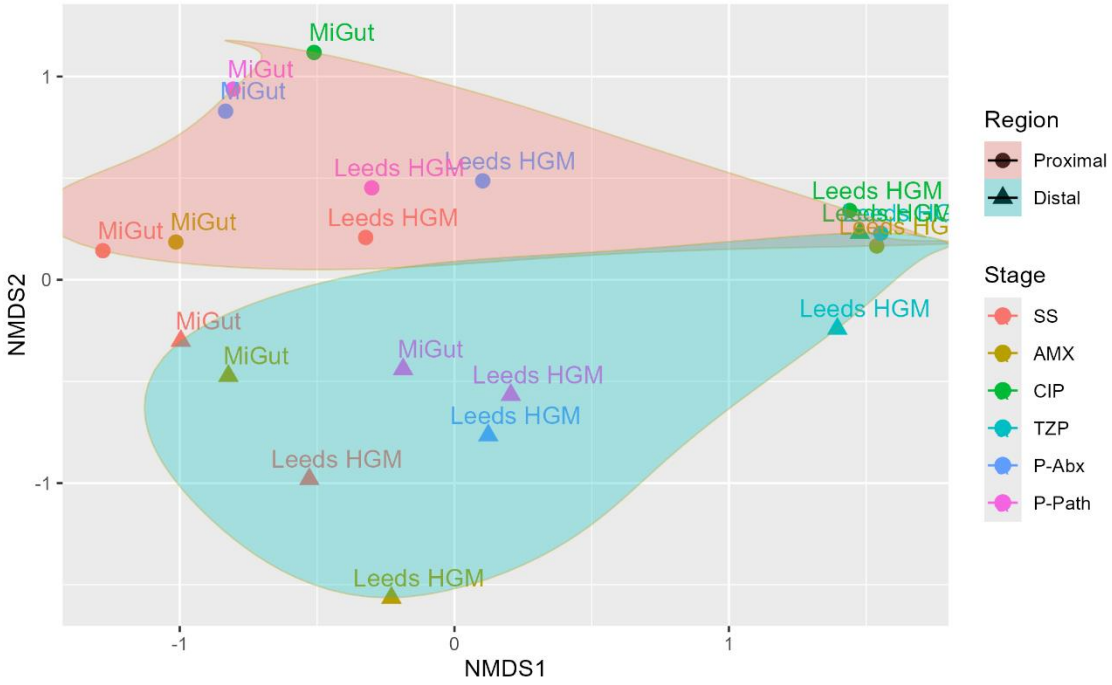

Figure 4-5 Genus Level Composition alterations, Vessel 1 vs Vessel 3. The Proximal(vessel 1)(Circle, Red boundary) and Distal (vessel 3)(Triangles, Blue Boundary) regions display distinct borders

Following the cessation of antibiotic treatment, the recovery trajectory trends towards a profile more similar to their region of origin. Although the disrupted microbial profiles are unique, exhibiting considerable deviation, there remains an underlying capacity for the microbiome to return to a composition more closely resembling its pre-antibiotic state.

4.5.1.2 Identifying Minority Genera

In the normalised dataset there were 76,344,629 reads assigned, resulting in 103 Taxa identified at Genus level. Within the 103 taxa, approximately 95 % of reads could be assigned to 25 genera presented in Figure 4-6 Distribution of assigned reads across all samples at the genus-level. Organisms which could not be identified to Genus Level were analysed together as “No Genus ID”. The genera identified as being responsible for the 95 % majority of assigned reads were: *Escherichia*, *Bifidobacterium*, *Lactobacillus*, *Hungatella*, *Bacteroides*, *Pediococcus*, *Schleiferilactobacillus, Lactiplantibacillus*, *Enterococcus*, No Genus ID, *Candida*, *Lacticaseibacillus*, *Enterocloster*, *Coprococcus*, *Shigella*, *Lachnoclostridium*, *Limosilactobacillus*, *Intestinimonas*, *Bariatricus*, *Faecalibacterium*, *Eisenbergiella*, *Parabacteroides*, *Anaerofilum, Clostridium*, *Sutterella*

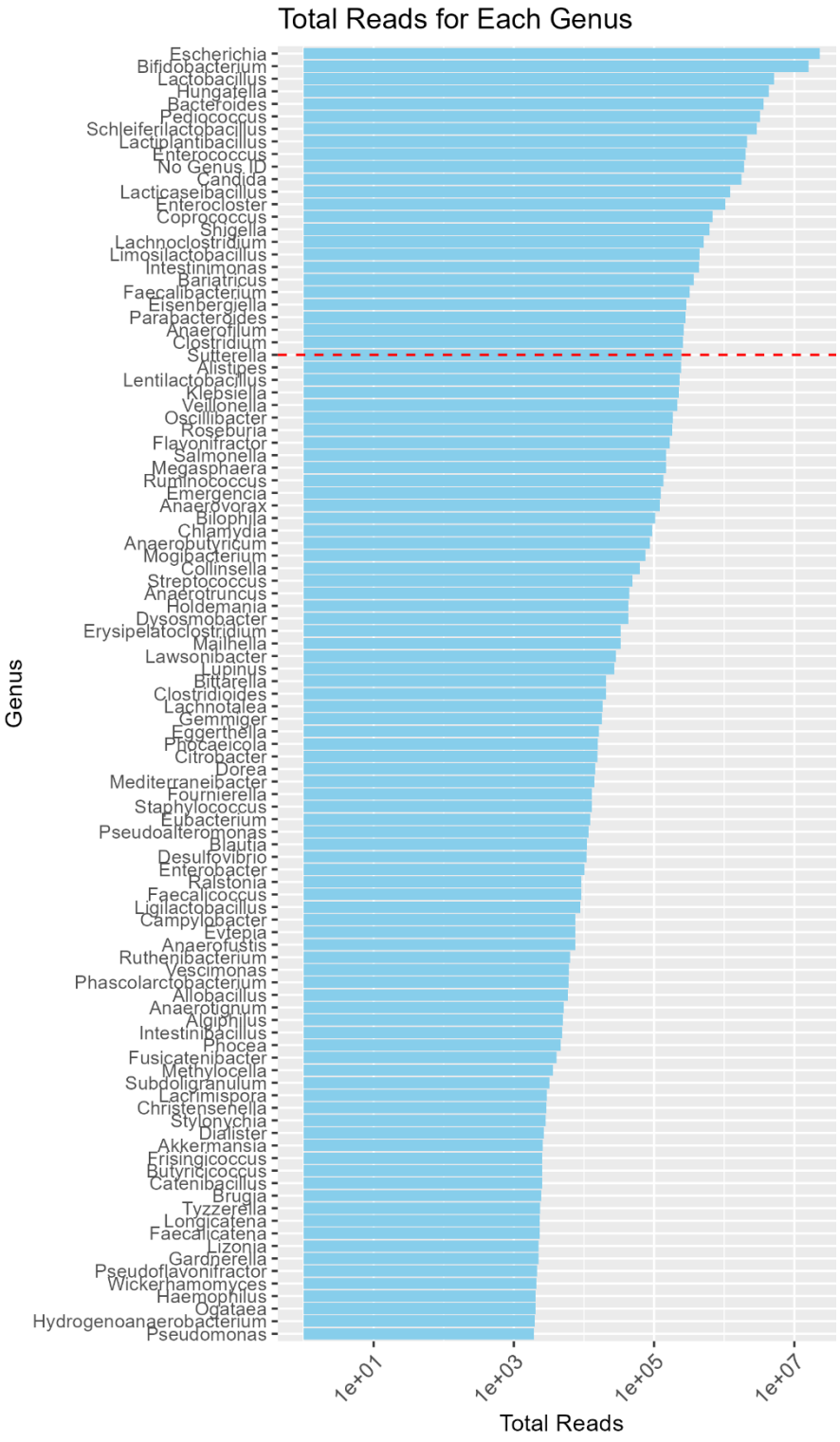

Figure 4-6 Distribution of assigned reads across all samples at the genus-level. The red dotted line represents a 95% read threshold, indicating the point at which 95% of the reads are allocated to a specific genus

4.5.1.3 Antibiotic Impact on Microbiome Diversity

The impacts of antimicrobial instillation of microbiome diversity (Shannon Index) were assessed for the 25 most abundant genera previously identified as being responsible for 95% of assigned reads, and minor genera. Analysis of microbiome composition was limited to the 9 most abundant genera, with others grouped as minor genera for clarity. Compositional changes are expressed as shifts in proportional abundances of genera by assigned reads at each experimental time point.

Vessel 1

Diversity and composition within the proximal (vessel 1) region of both models are presented in Figure 4-7 Vessel 1 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus. The MiGut sample corresponding to TZP is missing due to library failures discusses previously (4.4.3.2)

Following the equilibrium period the baseline microbiome displays notable variations between Leeds HGM and MiGUT replicates. Leeds HGM exhibited notably higher abundances of *Escherichia (*14.5%) and *Lactobacillus* (19.5%). In the MiGut platform at Steady State, the microbiome is dominated by Bifidobacterium, comprising 81.4%.

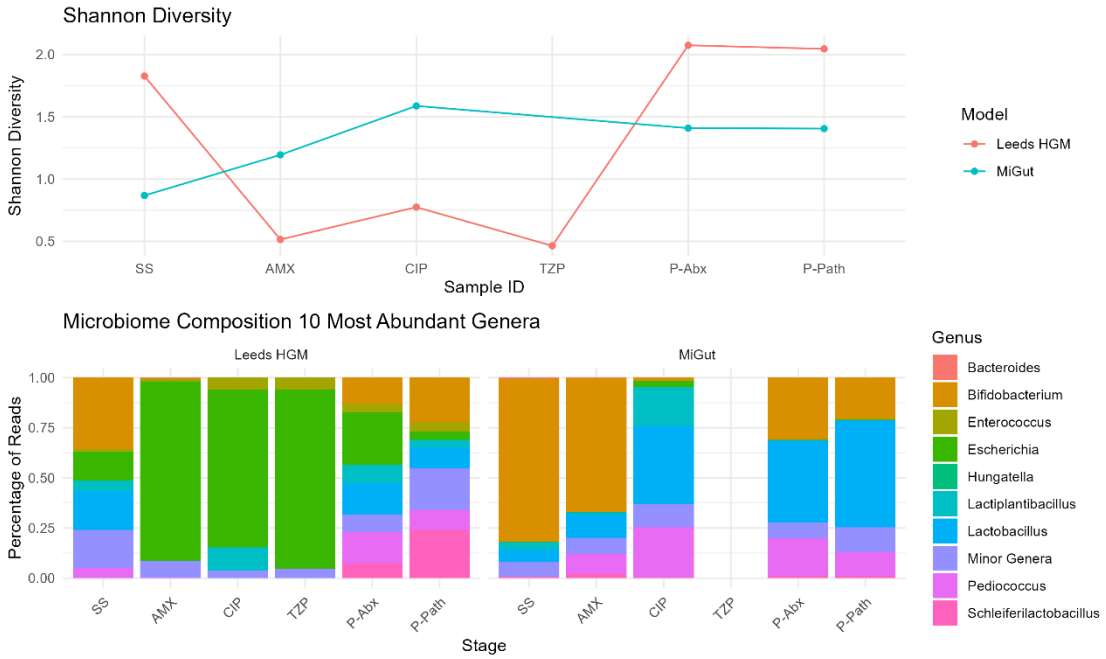

Figure 4-7 Vessel 1 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus

After AMX instillation, there was a drop in genus diversity in the Leeds HGM as the microbiome showed a proliferation of *Escherichia* (89.4%), with minor genera being the next most significant (8.3%). In the MiGut, there was an expansion of, *Pediococcus* (10.3%) and *Lactobacillus* (12.2%), coinciding with a reduction in *Bifidobacterium* (66.7%) and an increase in diversity as less prominent genera expanded.

During CIP instillation in the Leeds HGM, *Escherichia* remained the dominant genus (78.6%). There was also an expansion of *Lactiplantibacillus* (11.4%) and *Enterococcus* (6.1%). In the MiGut there was an increase in *Lactiplantibacillus* (19.3%), *Lactobacillus* (39.1%), and continuing expansion of *Pediococcus* (24.9%). There was an increase in diversity associated with the growth of less prominent genera.

TZP instillation caused significant disruption in the Leeds HGM microbiome, with many genera suppressed below detectable limits. *Escherichia* remained dominant (89.3%), followed by *Enterococcus* (6.1%) and minor genera (4.5%)

Following post-antibiotic recovery, diversity in the Leeds HGM was restored. All prominent genera seen prior to antibiotic instillation were present; *Bifidobacterium* (12.9%), *Escherichia* (26.3%), *Lactiplantibacillus* (8.8%), *Lactobacillus* (15.7%), minor genera (8.9%) and, *Pediococcus* (15.4%).

Enterococcus (4.3%), and Scheriferilactobacillus (7.5%) persisted following antimicrobial instillation. Post-recovery microbiota in MiGut also was reflective of pre-antimicrobial state, despite the loss of *Lactiplantibacillus*, *Bifidobacterium* (30.8%) remained present in addition to *Lactobacillus* (40.8%), Minor Genera (7.9%) and *Pediococcus* (19.0%).

Following pathogen exposure in the Leeds HGM microbiome composition included *Bifidobacterium* (22.2%), *Enterococcus* (4.5%), *Escherichia* (4.1%), *Lactiplantibacillus* (3.7%), *Lactobacillus* (10.4%), minor genera (20.7%), *Pediococcus* (10.1%), and *Schleiferilactobacillus* (24.0%). In the MiGut, the microbiome comprised *Bifidobacterium* (20.6%), *Lactobacillus* (53.3%), minor genera (12.2%), and *Pediococcus* (12.2%).

Vessel 3

Diversity and composition within the distal region of both models are presented in Figure 4-8 Vessel 3 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus. Sample points for MiGut corresponding to CIP,TZP and P-Abx are missing due to library failures discusses previously (4.4.3.2)

Following the equilibrium period, the distal vessel in both models exhibits greater baseline Shannon diversity than at steady state. The Leeds HGM displayed a high proportion of minor genera (68.6%) *Bacteroides* (14.7%), *Bifidobacterium* (2.1%), *Escherichia* (<2%), *Hungatella* (<2%), *Lactiplantibacillus* (2.6%), *Lactobacillus* (7.3%), and *Pediococcus* (<2%). MiGut exhibited a high proportion of Bifidobacterium and significant proportions of minor genera (23%) in addition to *Bacteroides* (5.4%), *Lactiplantibacillus* (2.5%), *Lactobacillus* (4.7%), and *Schleiferilactobacillus* (<1%)

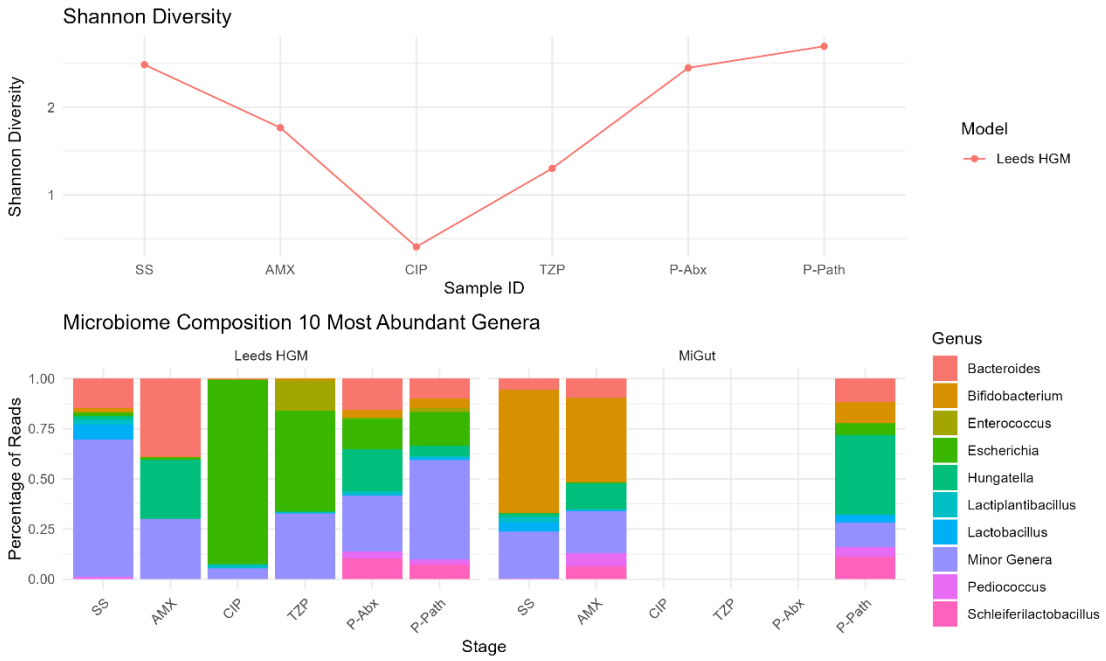

Figure 4-8 Vessel 3 Shannon Diversity and Microbiome Composition featuring 10 most abundant Genus. Due to the lack of sample points for MiGut, Shannon diversity is omitted

Following AMX instillation, Leeds HGM experienced a drop in Shannon diversity as the microbiome became dominated by Bacteroides (38.9%), *Hungatella* (29.4%), Minor Genera (30.2%), *Pediococcus, Lactiplantibacillus* and

Lactobacillus were suppressed below detectable limits. *Escherichia* remained low <2% and did not exhibit the same proliferation seen proximally in vessel 1.

Following CIP dosing, there was an emergence of *Escherichia* dominating the microbiome (91.1%). Similar to observations in vessel1 (proximal), although delayed, this dominance by *Escherichia* was associated with a notable drop in Shannon diversity. Minor genera were present in lower abundance (5.4%) and there was a re-emergence of *Lactiplantibacillus* (2.0%).

TZP saw a reduction in the abundance of *Escherichia (0.3%)* and an increase in diversity as *Enterococcus expanded (15.1%).* Minor Genera also expanded (32.7%)

Following the post-antibiotic recovery period, there was a significant increase in Shannon diversity within the Leeds HGM, returning to pre-antibiotic instillation levels. Microbiome composition seen post-antibiotic was similar to that seen prior to antibiotic treatment with some restructuring. *Bacteroides* (15.6%) and *Bifidobacterium* 3.5%) were established at levels similar to those seen prior to antibiotic intervention. However, *Escherichia* remained within the microbiome at an elevated level (15.3%)

Minor genera (28%) were present in lower abundance compared to the pre-antibiotic period. There was a restructuring of *Lactobacillaeceae* observer prior to antibiotic instillation, *Schleiferilactobacillus* (10.5%) replaced the previously dominant *Lactobacillus* species, which were present at <1% post-antibiotic. *Lactiplantibacillus* and *Pediococcus* were present at <2% and 3.3%, respectively, indicating lower abundance in these genera post-antibiotic treatment. *Hungatella* (21.2%) emerged as a more prominent strain post-recovery than prior to antibiotic treatment.

Following Pathogen exposure, the microbiome showed similar profiles to the post-antibiotic recovery structure. There was a similar composition between MiGut and Leeds HGM; the MiGut had a greater presence of *Hungatella* and a reduced proportion of Minor genera.

In the Leeds HGM, *Bacteroides* (9%) remained a significant portion of the microbiome. Escherichia (16.9%) continued to persist at similar levels, while *Hungatella* (5.3%) was present at a lower proportion. The composition of Lactobacillaceae remained stable, with *Lactiplantibacillus* (<1%), *Pediococcus*

(2.5%), *Lactobacillus* (<2%), and *Schleiferilactobacillus* (7.4%) being the most dominant genus. Minor genera accounted for 49.5% of the composition.

In the MiGut, *Hungatella* (39.8%) was the genus with the highest abundance. *Bacteroides* (11.7%), *Bifidobacterium* (10.4%), and *Escherichia* (6.0%) were also present. Minor genera accounted for 12.2% of the composition. The composition of Lactobacillaceae was comparable to that seen in the Leeds HGM. *Schleiferilactobacillus* (11.3%) was the dominant genus within the family, while *Lactobacillus* (3.8%) and *Pediococcus* (4.8%) were present but less abundant. There was a notable absence of *Lactiplantibacillus*.

4.5.1.4 Differential Abundance Analysis

Differential abundance analysis of the top 25 genera identified was performed with DESeq for proximal (vessel 1) and distal (vessel 3) regions. Genus falling below the 10th percentile base mean reads were filtered out, resulting in thresholds of 1112 for the proximal (vessel 1) region and 20806 for the distal (vessel 3) region. Differential abundances are reported as $\log_2$FoldChange for both Leeds HGM and MiGut platforms, comparing each experimental phase to the end of steady state.

Vessel 1

Vessel 1 differential abundance data for differentially abundant genera (Log2 fold change >1 and adjusted p-value <0.05) during the post-antimicrobial recovery period and post-pathogen inoculation are presented in Table 4-3 Vessel 1 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery and Table 4-4 Vessel 1 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation

Table 4-3 Vessel 1 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery

Genus	Log2FoldChange	Adjusted p-value
Coprococcus	-30.00	1.94E-09
Bacteroides	-29.33	4.38E-09
Faecalibacterium	-27.31	4.61E-08
Pediococcus	5	1.02E-02

Table 4-4 Vessel 1 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation

Genus	Log_2FoldChange	Adjusted p-value
Coprococcus	-30.62	3.40E-08
Bacteroides	-29.62	3.40E-08
Faecalibacterium	-27.84	1.85E-07

Vessel 3

Vessel 3 differential abundance data for differentially abundant genera (Log2 fold change >1 and adjusted p-value <0.05) during the post-antimicrobial recovery period and post-pathogen inoculation are presented in Table 4-5 Vessel 3 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery and Table 4-6 Vessel 3 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation.

Table 4-5 Vessel 3 (Leeds HGM + MiGut) Differential Abundance Post-Antibiotic Recovery

Genus	Log2FoldChange	Adjusted p-value
Sutterella	-30.00	$5.94x10^{-6}$
Coprococcus	-19.45	$6.39x10^{-4}$
Faecalibacterium	-18.91	$2.27x10^{-3}$
Shigella	16.37	$2.32x10^{-6}$
Escherichia	4.15	$4.22x10^{-3}$

Table 4-6 Vessel 3 (Leeds HGM + MiGut) Abundance Post-Pathogen Inoculation

Genus	Log_2FoldChange	Adjusted p-value
Sutterella	-30.00	$9.38x10^{-09}$
Coprococcus	-19.93	$5.20x10^{-06}$
Faecalibacterium	-19.39	$4.49x10^{-05}$
Shigella	15.44	$5.99x10^{-08}$
Escherichia	3.30	$6.94x10^{-03}$
Hungatella	4.32	$3.96x10^{-02}$

4.5.1.5 Species Diversity Across Experimental Phases.

Total Species Count

Quantifying the total number of identified species present at each experimental phase provides a detailed assessment of diversity, illustrating how the microbiome composition may alter in response to antibiotic instillation. The total number of unique species is presented in Table 4-7 Comparison of Species counts in Leeds HGM and MiGut platforms.

Table 4-7 Comparison of Species counts in Leeds HGM and MiGut platforms. Sample points without associated NGS data are described as not applicable (NA).

Stage	Leeds HGM		MiGut	
	Vessel 1	Vessel 3	Vessel 1	Vessel 3
SS	48	157	53	92
AMX	27	81	43	87
CIP	23	23	28	NA
TZP	21	45	NA	NA
P-Abx	40	75	28	NA
P-Path	37	90	28	49

As expected, vessel 3, representing conditions of the distal colon, harboured the highest number of unique species compared with vessel 1, reflecting the conditions of the proximal region.

Antibiotic instillation had a dramatic effect on the total number of unique species, which decreased in both the first and last vessels regardless of the model system used. In the Leeds HGM vessel 3 displayed the greatest reduction in species count following TZP dosing. In the Distal region, the greatest reduction coincided with CIP administration, with recovery observed during TZP dosing.

The steady-state phase had the highest abundance of unique species. Total Species number did recover following the cessation of antibiotics in both vessel and vessel 3; however, they did not return to their pre-antibiotic abundances.

Species Unique to Region

Vessel 1 and Vessel 3 were assessed for the proportions of the population unique to each region (Presented in Figure 4-9). At each time point, the distal region exhibited the highest species diversity, with 140 of the 182 species identified across both model platforms being exclusive to this region. In contrast, most species in the proximal (vessel 1) region were shared with the distal (vessel 3) region, with only 28 of the 70 species identified in the proximal region being unique to that location.

Figure 4-9 Comparison of Unique and Shared Species Counts Across Proximal (vessel 1) and Distal (vessel 3) Regions at Steady State, Post-Antibiotic Recovery, and Post-Pathogen Challenge

The reduction in species was not confined to any specific spatial region or within unique / shared populations. Species depletion was observed across all populations and regions following antibiotic instillation, which persisted through the recovery period and post-pathogen challenge.

4.5.1.6 Impact of Antibiotic Instillation on Species Diversity in Highly Diverse Genera

Antibiotic instillation is associated with changes in species diversity within highly diverse genera, presented inFigure 4-10 Heat map of species counts by genus across Leeds HGM and MiGut Platforms.

Proximal Region (vessel 1)

In the proximal region*, Lactobacillus species* were initially established as 11 species in Leeds HGM and 10 species in MiGut. During antibiotic instillation, populations of *Lactobacillus* were suppressed, falling below detectable levels. The impact of AMX and CIP was not as apparent in the MiGut platform, where species abundances remained consistent throughout the dosing period. The cessation of antimicrobials was associated with a recovery in detectable *Lactobacillus species* in the Leeds HGM.

Bifidobacterium established in both models to similar levels, 9 in Leeds HGM and 8 Species in MiGut. Antibiotic instillation led to a reduction in species diversity for both models. AMX had a greater impact on species diversity within the Leeds HGM. During CIP and TZP dosing, species numbers remained low across both platforms. Following the cessation of antibiotic instillation, there was a partial recovery in *Bifidobacterium species* in both models; however, neither reached the abundances seen during Steady State.

Veillonella was established in both models. The Leeds HGM initially had a higher level of species diversity. However, antibiotic exposure depleted *Veillonella* species below detectable levels in both models, remaining below detection for the continuation of the model

Distal Region

Bacteroides initially established in the Leeds HGM with 11 species; there was a lower initial level of diversity in *Bacteroides* species within the MiGut with 2 Species established. There was an initial increase in *Bacteroides* species during AMX instillation in the Leeds HGM. However, CIP and TZP saw severe depletion in *Bacteroides* species throughout dosing. *Bacteroides* were able to recover to pre-antibiotic levels of species abundance after the removal of antibiotics.

Enterocloster was established as 6 species in both Leeds HGM and MiGut models. AMX instillation had no impact on recoverable species of *Enterocloster*; however, CIP and TZP were reduced in detectable species. Following the cessation of antibiotic dosing, species diversity was able to recover to similar levels 5 Species in Leeds HGM and 4 in MiGut.

Bifidobacterium established in both models, 4 detectable species during Steady state in the Leeds HGM and 8 species in the MiGut. AMX and CIP saw depletion in the number of recoverable species of *Bifidobacterium.* There was some recovery during TZP. However, levels had reduced again following the post-antibiotic recovery period and remained below pre-antibiotic exposure levels for the continuation of the experiment.

Clostridium were established as 5 species in Leeds HGM and 4 in MiGut. There was a notable reduction in recoverable species during CIP and TZP instillation. AMX had less impact on *Clostridium species.* Populations in the Leeds HGM were able to recover to a greater extent to 4 species, MiGut populations remained low at 1 species.

Lactobacillus colonisation was similar across both models 6 species in Leeds HGM and 4 species in MiGut, AMX and CIP were associated with a depletion of species diversity; there was some recovery observed in Leeds HGM during TZP. However, there was no further recovery, and populations never returned to pre-antibiotic levels.

Oscillibacter colonised both models to a similar extent; 8 species in Leeds HGM and 7 species MiGut became severely depleted during antibiotic dosing in Both Leeds HGM and MiGut. AMX had a greater impact on diversity in Leeds HGM. *Oscillibacter* species remained below detectable levels for the remainder of the experiment following antibiotic dosing.

Faecalibacterium established as 4 species in both Leeds HGM and MiGut. Antibiotic dosing resulted in the depletion of all species below detectable levels. Neither model saw a resurgence of species following the cessation of antibiotics

Sutterella was only detectable in Leeds HGM at steady state. There was a notable presence in both models during AMX dosing. The reason for lack of detection in MiGut during Steady state not being evident. CIP and TZP saw the

depletion of *Sutterella* species below detectable levels where they remained for the continuation of the experiment

Alistipes and *Ruminococcus* only established detectable levels within the Leeds HGM, with 6 and 7 species being detected, respectively. Both genera were depleted below detectable levels following CIP administration and remained undetectable for the rest of the experiment.

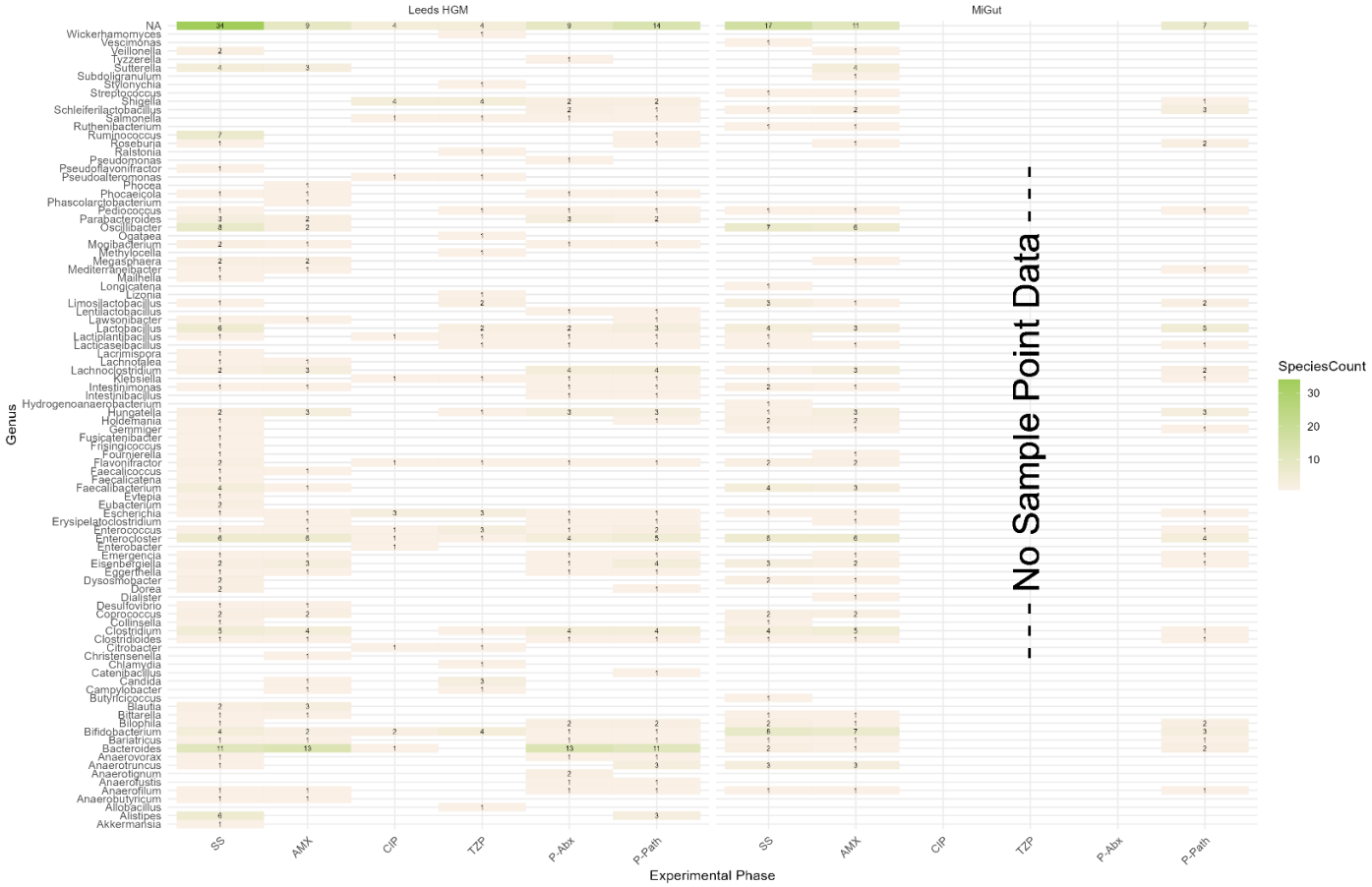

Figure 4-10 Heat map of species counts by genus across Leeds HGM and MiGut Platforms. Unique species frequency within the Distal portion of *in-vitro* gut models throughout experimental phase

4.5.2 Metagenomic Functional Profiling of the intestinal microbiome

MEGAN performs functional analysis on NGS data by binning sequence data into KEGG categories, utilising its functional hierarchies, KEGG Pathway and Brite Hierarchies. KEGG Pathway groups entities into pathways based on their specific biological functions, whereas Brite Hierarchies classifications are based upon broader biological functions. At steady-state there was a total of 41,371,104 pathway assignments. This includes 23,913,110 reads assigned to KEGG Pathway, 13,868,689 reads to Brite Hierarchies, and 3,589,307 reads categorized as Not Included in Pathway or Brite (Figure 4-17 Total Assigned Reads at Steady State for Leeds HGM and MiGut).

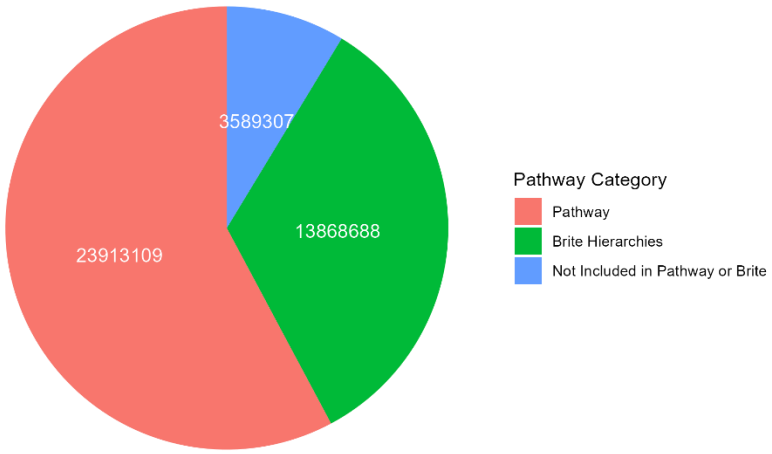

Figure 4-11 Distribution of Reads by Pathway Categories at Steady-state for both Leeds HGM and MiGut platforms. Breakdown of assigned reads between KEGG hierarchical classifications, KEGG Pathways (Red), Brite Hierarchies (Green), and Unassigned Categories (Blue)

Not Included Pathway or Brite assignments were discarded from further analysis. Brite hierarchies were excluded due to a lack of granularity in the structure, making it challenging to form a cohesive picture of functional changes within the microbiome.

4.5.2.1 KEGG Pathway

At Steady state, KEGG Pathway analysis identified 5560 enzymes across 314 sub-subpathways and 46 subpathways. Carbohydrate metabolism emerged as the most abundant subpathway with 5.59E+06 reads assigned (Figure 4-12 Total Reads Assigned to Subpathways within KEGG Pathways).

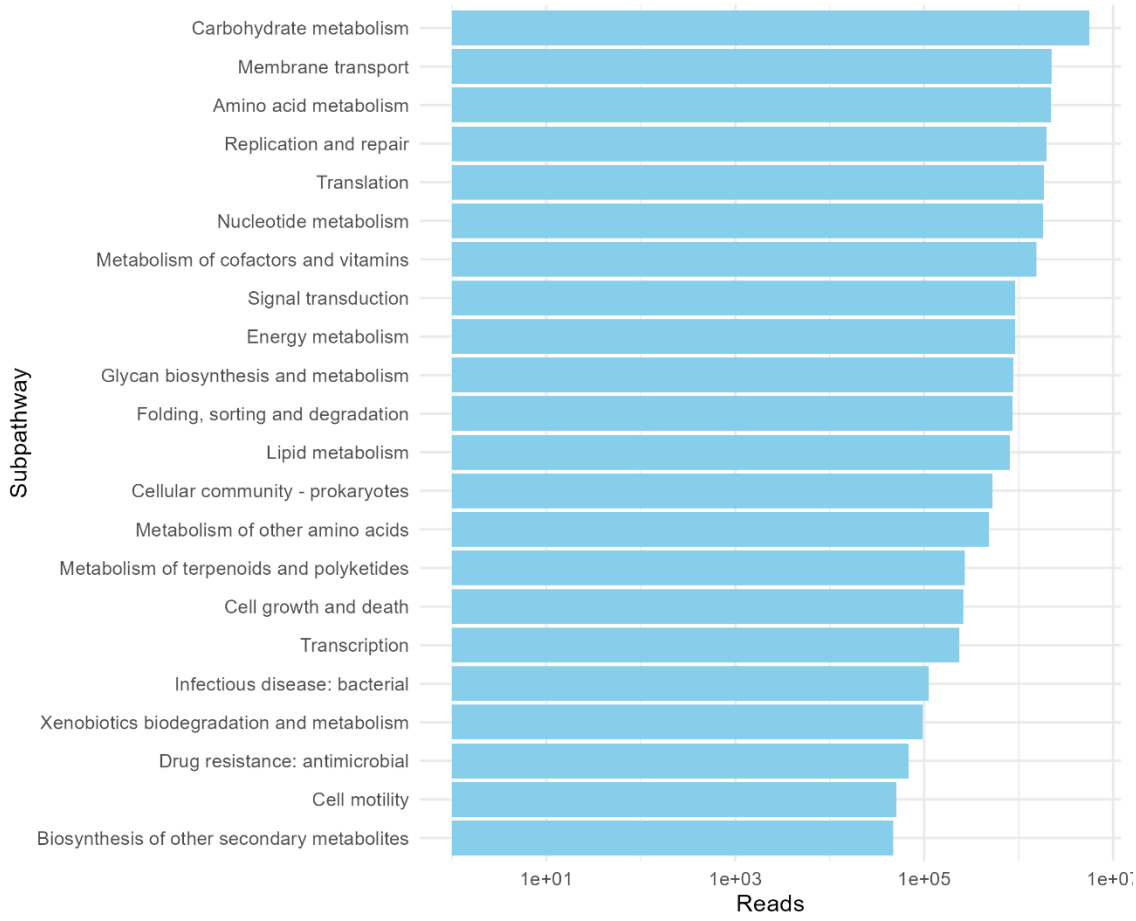

Figure 4-12 Total Reads Assigned to Subpathways within KEGG Pathways. Steady State Distribution of assigned reads within KEGG Pathways at Sub-pathway resolution.

Among the ten most abundant subpathways, six were associated with the Metabolism Pathway: Carbohydrate metabolism, Amino acid metabolism, Nucleotide metabolism, Metabolism of cofactors and vitamins, Energy metabolism, and Glycan biosynthesis and metabolism. 2 were associated with Environmental Information Processing pathway: Membrane transport and Signal transduction, and 2 were associated with Genetic Information Processing: Replication and repair, and Translation

Targeted Pathway Interrogation of the Metagenome: SCFAs

Several genera, including *Faecalibacterium*, *Intestinimonas*, *Coprococcus*, and *Lactobacillus*, which exhibited significant disturbances during antimicrobial instillation, are known to be associated with the production of short-chain fatty acids (SCFAs). To determine whether these genus-level alterations are reflected in changes to the functional capacity of the metagenome, a targeted approach was employed to evaluate specific metabolic pathways associated with SCFA production. SCFA are crucial intestinal metabolites produced by bacterial fermentation of non-absorbed carbohydrates such as resistant starches fructans, xylans, gums and pectins within the colon (see Figure 4-13 Outline of major Short-chain Fatty acid production.).

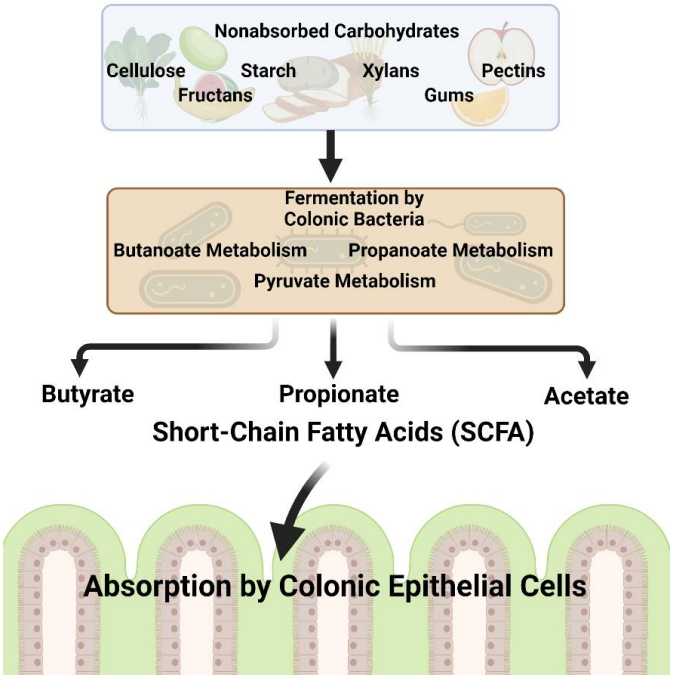

Figure 4-13 Outline of major Short-chain Fatty acid production. Image adapted from KEGG Carbohydrate digestion and absorption Reference pathway (ko04973) (KEGG, 2019).

The key functional pathways investigated include Butanoate Metabolism, Propanoate Metabolism, and Pyruvate Metabolism, with a specific focus on acetate production. These pathways are identified as prime candidates for more focused investigation, not only due to the functional implications of SCFAs as crucial metabolites for colonocytes and their role in gut homeostasis but also their potential to be altered based on taxonomic observations

Butanoate (Butyrate) Metabolism (ko00650)

The pathways for Butyrate synthesis are well-characterised and comprise four main routes: the acetyl-CoA, Glutarate, 4-aminobutyrate, and lysine pathways (Vital et al., 2014). Within the dataset, 54 unique enzyme assignments were identified within Butanoate metabolism. To assess potential changes to Butyrate production by the intestinal microbiome, differential abundance analysis was performed on all identified components. Assignments with low abundances, those falling before the 10th percentile of reads, were filtered out before analysis. Where applicable, assignments were categorised by their associated pathways, and log_2FoldChange plotted Table 4-8 Differentially Abundant Genes Within Butyrate Metabolism. Assignments within butanoate metabolism but outside the scope of this chapter's analysis were discarded from future consideration. Assignments were classed as differentially abundant if they had a log_2FoldChange greater than and an adjusted p-value less than 0.05. Those which met the criteria for being differentially abundant are annotated in Figure 4-14 Differentially Abundant Enzymes and Their Roles in Butyrate Metabolism

Table 4-8 Differentially Abundant Genes Within Butyrate Metabolism

Enzyme Name	Enzyme Commission number	KEGG Orthology Identifier	Log_2Fold Change	Adjusted p-value
Glutaconate CoA-transferase Subunit A[1] & B[2]	EC:2.8.3.12	K01039[1] K01040[2]	-5.00[1] -11.06[2]	$2.0x10^{-3}$ (1) $3.7x10^{-3}$ (2)
4-Hydroxybutyrate dehydrogenase / sulfolactaldehyde 3-reductase	EC:1.1.1.61, EC:1.1.1.373	K08318	4.06	$2.0x10^{-4}$
Acetolactate synthase II small subunit	EC:2.2.1.6	K11258	4.46	$3.0x10^{-4}$

[1]Values for Subunit A, [2] Values for subunit B

Notable changes were observed in genes associated with the acetyl Co-A pathway, with three out of four identified genes showing reductions and one gene exhibiting an increase in abundance. Despite, log_2 fold changes exceeding 1, none of these changes reached statical significance. Among the identified genes associated with acetyl-CoA metabolism, K00022 (3-hydroxyacyl-CoA dehydrogenase [EC:1.1.1.35]) exhibited an increase, while K00074 (3-

hydroxybutyryl-CoA dehydrogenase [EC:1.1.1.157]), K01715 (enoyl-CoA hydratase [EC:4.2.1.17]), and K07516 (3-hydroxyacyl-CoA dehydrogenase [EC:1.1.1.35]) showed reductions.

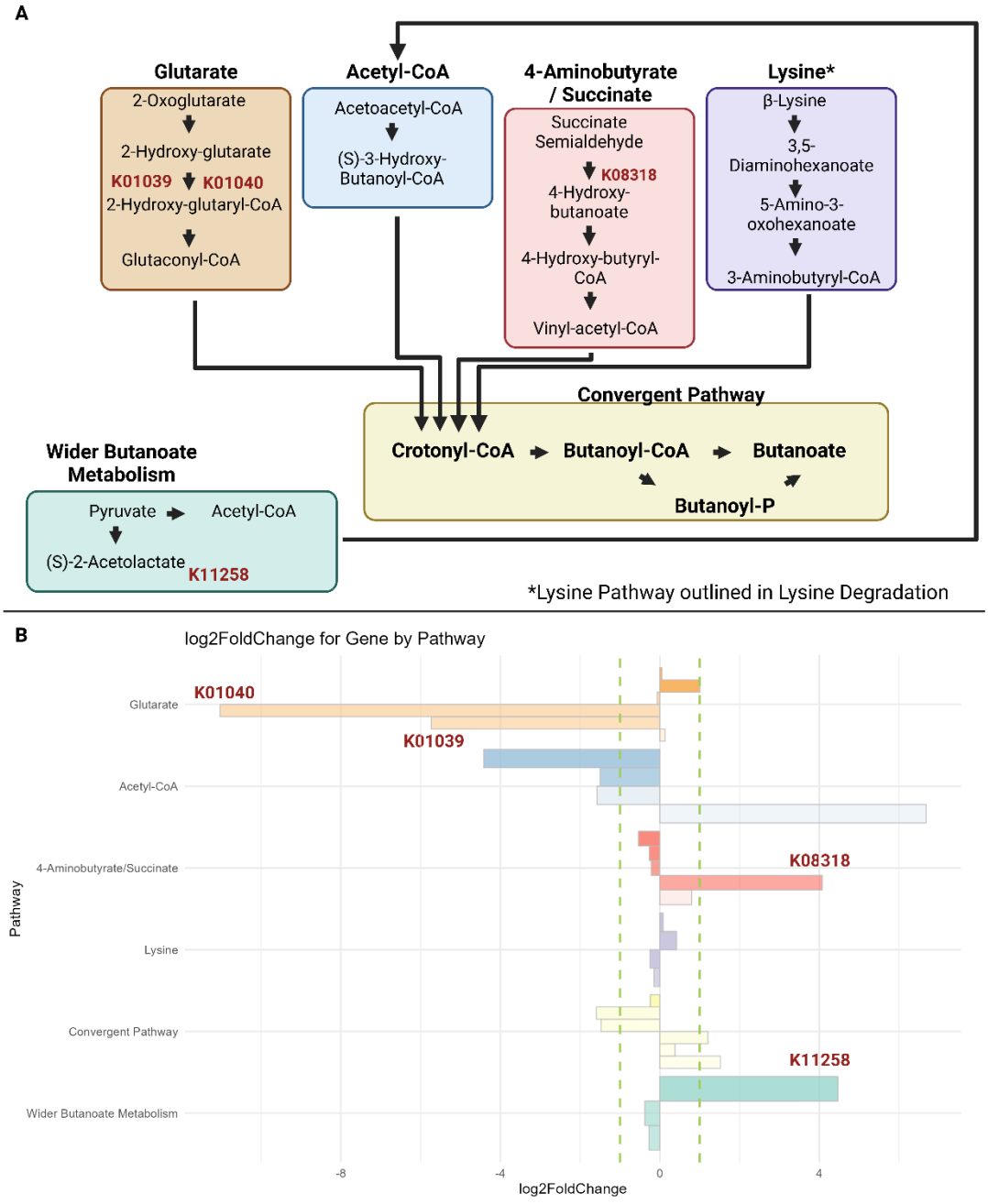

Figure 4-14 Differentially Abundant Enzymes and Their Roles in Butyrate Metabolism. (A) Outlines pathways discussed in this chapter highlighting key metabolic steps. (B) Depicts log_2FoldChange of enzymes within the identified pathways. Green lines indicate a log_2FoldChange greater than 1. Annotated enzymes; K01039 glutaconate CoA-transferase subunit A, K01040 glutaconate CoA-transferase, subunit B, K08318 4-hydroxybutyrate dehydrogenase / sulfolactaldehyde 3-reductase, K11258 acetolactate synthase II small subunit

Propanoate Metabolism (ko00640)

The KEGG Pathway ko00040 for propanoate metabolism highlights three primary routes for propionate production, which occur via lactate, succinate, and L-lactaldehyde. In the dataset, 46 distinct enzyme assignments were identified as being part of the propanoate metabolism pathway. To explore potential changes in propanoate metabolism, a differential abundance analysis was performed on all identified components. Enzyme assignments with abundances below the 10th percentile were filtered before analysis. Assignments were then processed in the same manner as those for butyrate metabolism. Several displayed $\log_2$-fold change greater than 1; however, they failed to an adjusted p-value less than 0.05. Those which met the criteria for being differentially abundant are annotated in Figure 4-15 Differentially Abundant Genes and Their Roles in Propanoate Metabolism. Following the antibiotic recovery period, there was a persistent decrease in enzymes associated with the synthesis of Propanoate from Lactate presented in Table 4-9 Differentially Abundant Genes Within Propanoate Metabolism.

Table 4-9 Differentially Abundant Genes Within Propanoate Metabolism

Enzyme Name	Enzyme Commission number	KEGG Orthology Identifier	Log_2Fold Change	Adjusted p-value
Butyryl-CoA dehydrogenase	EC:1.3.8.1	K00248	-2.42	$2.42x10^{-6}$
Multifunctional beta-oxidation protein	EC:4.2.1.-, EC:1.1.1.-	K14729	-5.27	$2.47x10^{-5}$
Propionate kinase	EC:2.7.2.15	K00932	3.14	4.57x10-^{4}

Methylglyoxal/glyoxal reductase (K23257 [EC:1.1.1.283, EC:1.1.1.-]) also showed an increase in abundance; however, this is discussed further within the context of acetate synthesis in Pyruvate Metabolism.

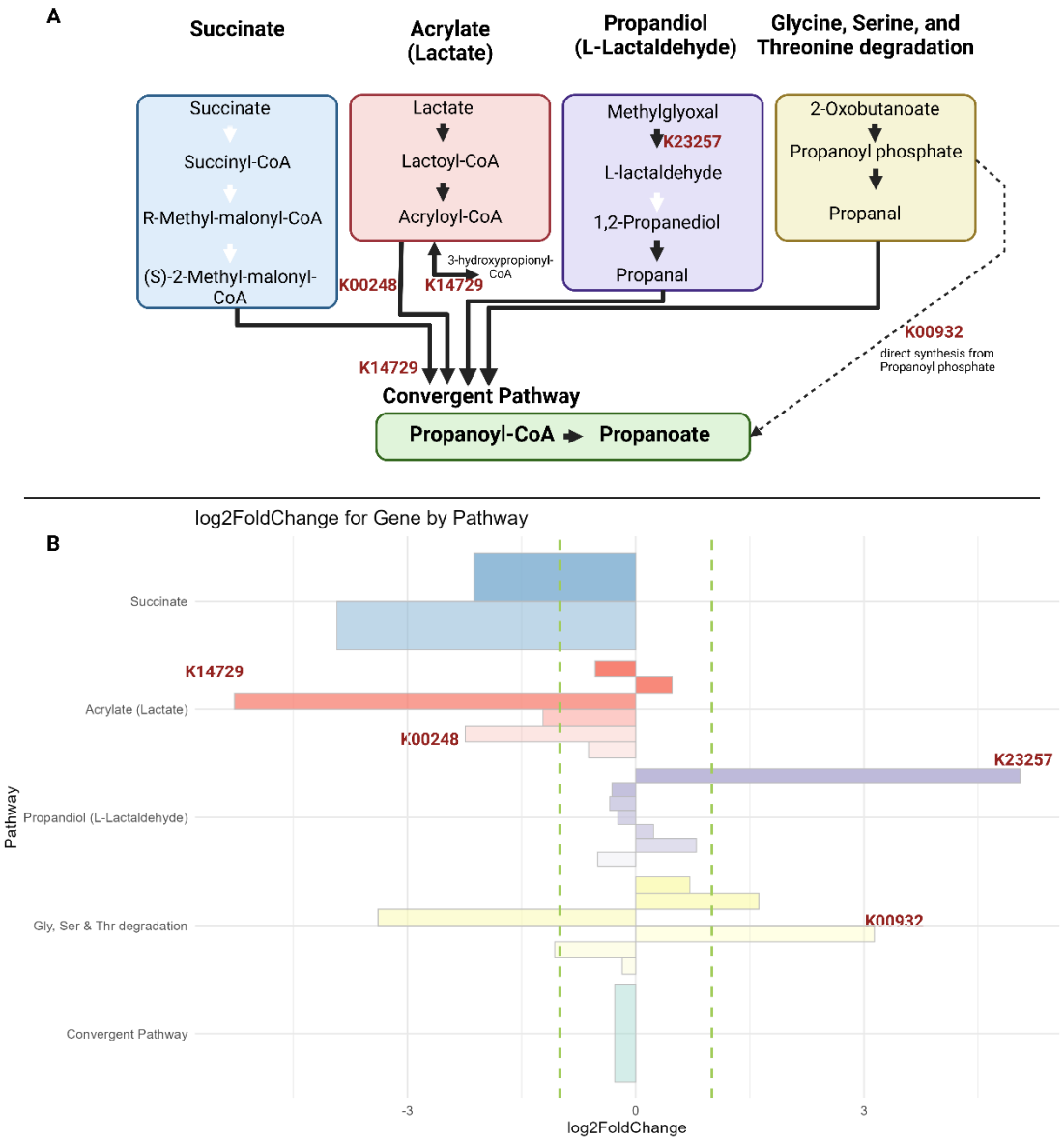

Figure 4-15 Differentially Abundant Genes and Their Roles in Propanoate Metabolism. Outlines metabolic pathways discussed in this chapter, highlighting key metabolic steps. White arrows depict, reaction steps for which no associated enzymes were identified within the dataset for any timepoint (B) Depicts log_2FoldChange of enzymes within the identified pathways. The dotted green lines indicate a log_2FoldChange greater than 1. Annotated enzymes; K00248 butyryl-CoA dehydrogenase, K14729 multifunctional beta-oxidation protein, K00932 propionate kinase, K23257 methylglyoxal/glyoxal reductase.

Pyruvate Metabolism (ko00620)

The pathways for acetate production are described within the KEGG map for pyruvate metabolism, with several pathways centred around acetyl-CoA contributing to acetate synthesis. This chapter focuses on lactate, which is closely tied to pyruvate and the central pathways for acetyl-CoA metabolism. This focus was determined post-analysis as the observations were limited primarily to these pathways and did not limit the scope of differential abundance analysis. Given the diversity of acetate production, it is more straightforward to discuss the findings within this context. To assess potential alterations in acetate metabolism, differential abundance analysis was conducted on all identified components. Data points with mean abundances below the 10th percentile threshold were filtered before analysis of results. Assignments were then processed in the same manner as those previously described. Those which met the criteria for being differentially abundant are annotated in Figure 4-16 Differentially Abundant Genes and Their Roles in Acetate/Pyruvate metabolism.

Table 4-10 Differentially Abundant Genes Within Acetate/Pyruvate Metabolism

Enzyme Name	Enzyme Commission number	KEGG Orthology Identifier	Log_2FoldChange	Adjusted p-value
methylglyoxal/glyoxal reductase	EC:1.1.1.283, EC:1.1.1.-	K23257	5.05	0.0403
L-lactate dehydrogenase (cytochrome)	EC:1.1.2.3	K00101	-9.30	0.0046
lactate 2-monooxygenase	EC:1.13.12.4	K00467	-3.45	0.0107

[1] Associated with the metabolism of L-lactate

K21618 (R)-2-hydroxyglutarate-pyruvate transhydrogenase [EC:1.1.99.40] exhibited a decrease in abundance similar to L-lactate dehydrogenase (cytochrome), with a log_2-fold change of -8.53. Although its P-Value initially failed to reach significance, this decrease persisted post-pathogen challenge, where its level of significance increased with additional replicates from the MiGut platform.

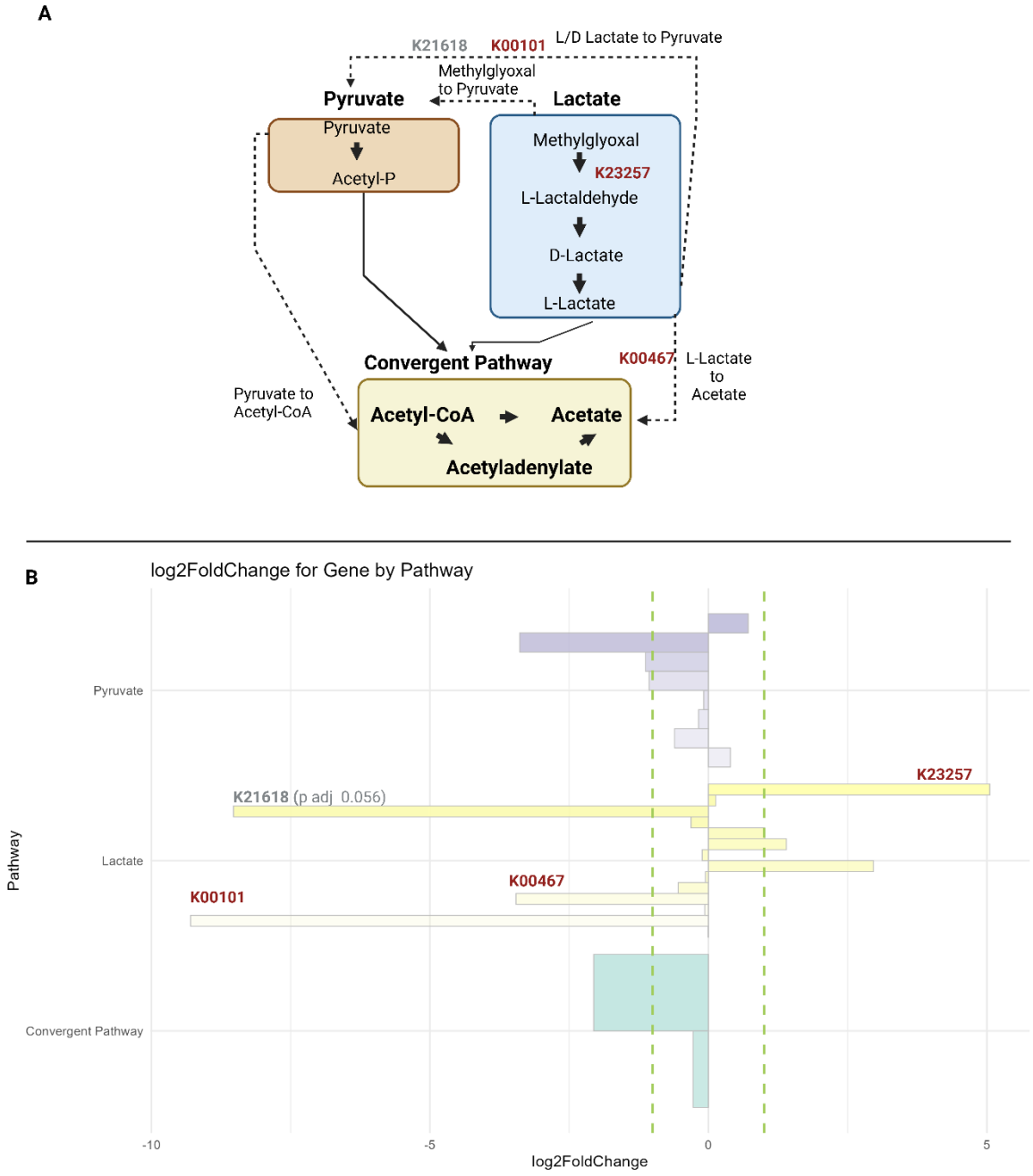

Figure 4-16 Differentially Abundant Genes and Their Roles in Acetate/Pyruvate metabolism. Outlines metabolic pathways discussed in this chapter highlighting key metabolic steps. (B) Depicts log_2FoldChange of enzymes within the identified pathways. The dotted green lines indicate a log_2FoldChange greater than 1. Annotated enzymes; K00101 L-lactate dehydrogenase, K00467 lactate 2-monooxygenase, K23257 methylglyoxal/glyoxal reductase, K21618 (R)-2-hydroxyglutarate-pyruvate transhydrogenase

4.6 Discussion

4.6.1 Steady state microbiome

The baseline microbiome was comparable between Leeds HGM and MiGut platforms. Key genera, particularly those monitored by culture, were established throughout both models by the end of the steady state phase, notably *Bacteroides*, *Bifidobacterium*, *Lactobacillus*, *Enterococcus*, and *Escherichia*.

There was an increased abundance of Bifidobacterium within the MiGut platform at a steady state. Increased *Bifidobacterium* colonisation was noted in 16s qPCR analysis of the experiment, albeit to a lesser extent (Davis Birch et al., 2023).

When viewing the baseline microbiome as total reads rather than proportionally (Figure 4-17 Total Assigned Reads at Steady State for Leeds HGM and MiGut), the comparability of the baseline composition before much more evident between the platforms.

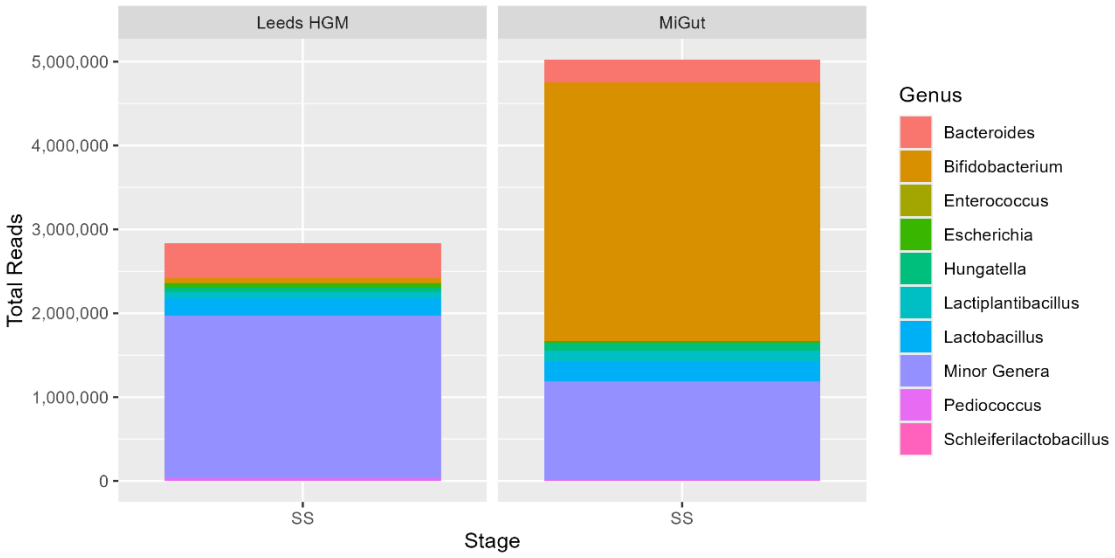

Figure 4-17 Total Assigned Reads at Steady State for Leeds HGM and MiGut. Total Genus level assignments at steady state shows a greater abundance of *Bifidobacterium* within the MiGut compared to Leeds HGM

It is unclear whether this increased abundance accurately represents the true compositional nature of the microbiome or if it is an artefact of DNA extraction and sequencing. This increase in the abundance of reads, particularly within Bifidobacteria, may influence diversity metrics. Typically, increased sampling

depth may have impacts on the ability to recover minority genera, increasing the diversity of the data. However, in this instance, it appears the only reads within *Bifidobacteria are inflated.* Increased abundances of *Bifidobacteria* within the model at steady state and AMX may particularly influence differential abundance data when compared to steady state. In this instance, while there were large changes within the abundance of Bifidobacteria, they did not reach statistical significance. This is likely due to the absence of this observation within the Leeds HGM.

Displaying microbiome data in an intuitive way may be challenging, and it is important to identify how choosing to represent compositional data may impact interpretation. Presenting the data as proportionate reads rather than total reads may be useful; however, in this instance, the increased abundance of Bifidobacterium compressed other genera in MiGut data.

In both models, proximal (vessel 1) and distal (vessel 3) regions show unique compositional profiles. Vessel 3 displays a greater genus-level diversity compared to Vessel 1. Reflecting culture base findings *Bacteroides* fail to become established in the vessel 1 proximal region. The failure of *Bacteroides* to effectively colonise vessel 1 contrasts with some findings in the literature (Ahmed et al., 2007) (Wang, X. et al., 2003). The lack of *Bacteroides* colonisation within vessel 1 may be attributable to a lack of seeding from a preceding vessel.

4.6.2 Impacts of Antimicrobials on the Intestinal Microbiome.

With the growing understanding of the potential contributions of the microbiome to human health, it is increasingly important to understand the potential unintended consequences associated with medical interventions.

The metagenomic data analysed in this chapter echoes many of the observations seen through culture-based analysis, expanding on them and offering additional levels of insight beyond key microbial species isolated for enumeration.

4.6.2.1 Compositional Shifts

Ordination of genus-level taxonomic data with NMDS explores the compositional trajectory throughout the simulated course of multiple rounds of

antimicrobial therapy, which might reasonably be experienced in chronically sick individuals, tracking the subsequence recovery phase.

There appear to be notable differences in the impact which different antibiotics introduced into the model have upon the microbiome, with amoxicillin in most distances displaying smaller compositional disruptions to the microbiome than ciprofloxacin or piperacillin-tazobactam. However, this may be partially attributed to the dosing regimen being designed not to allow complete recovery of the microbiome between therapies.

As previously mentioned, the proximal (vessel 1) and distal (vessel 3) appear to have compositionally distinct profiles. However, following antimicrobial instillation, the composition of the disrupted microbiomes becomes more similar to each other than their spatially related steady-state compositions. Despite the disrupted microbiome occupying a distinct cluster, they retain the underly diversity needed to return to compositions more similar to their distinct initial profiles, likely driven by the tight environmental controls within the system and differing nutrient availability.

4.6.2.2 Diversity

Diversity is a key metric within microbiome research, it offers the potential to describe the varying structure and composition of the microbial community. This chapter explored how antibiotic instillation can impact both the total abundance of species within individual genera and alterations to the genus-level composition of the microbiome. Total species diversity was important to the work within this thesis due to functional heterogeneity among species; the loss or gain of species within a genus may also indicate shifts in functional profiles. Genus diversity here was explored as a descriptive metric for understanding the restructuring of the microbiota coinciding with antibiotic instillation.

Species Diversity

In response to antimicrobial instillation, there are notable reductions in the diversity of species present for several genera, with others being lost entirely.

Lactobacillus within the Leeds HGM, there was a notable reduction in overall strain diversity. *Lactobacillus* species in the MiGut platform appeared to be more resilient to antibiotic instillation being able to recover with the emergence

of a previously low abundant strain. Lactobacilli are well documented as being potential probiotic strains responsible for the production of lactic acid and SCFAs. Their potential to produce SCFA has been documented with varying capacity even within strains of the same species, with co-cultures showing increased production over single strain cultures, strains associated with higher butyrate production also being shown to exhibit stronger antimicrobial activity against pathogens (Thananimit et al., 2022).

Bifidobacterium species diversity was severely depleted across both models investigated in this chapter, with only a single species remaining in each post-antibiotic instillation and recovery. *Bifidobacteria* are important fermentative constituents of the intestinal microbiome, responsible for the breakdown of non-digestible poly and oligosaccharides to fermentable monosaccharides, with varying metabolic profiles within the genus (Falony, G. et al., 2006; Scott et al., 2014; Dong et al., 2024) .Bifidobacteria typically lack the ability to produce butyrate directly, while Usta-Gorgun *et al*. (Usta-Gorgun and Yilmaz-Ersan, 2020) described their potential to produce butyrate in the presence of salep, they are primarily associated with the production of acetate and lactate (De Vuyst, L. et al., 2014). The production of acetate by bifidobacteria has been shown to cross-feed production by butyrate-producing species (Falony, G. et al., 2006; De Vuyst, Luc and Leroy, 2011). A reduction in *Bifidobacterium* species, may therefore result in a reduced metabolic profile for the utilisation of non-digestible carbohydrates within the microbiome, reducing metabolic cross-feeding and a reduction in SCFA production. The reduced *Bifidobacterium* diversity may have implications for the maintenance of the host immune system.

There was a significant reduction in species diversity within the genus *Ruminococcus*. The discussion around this genus is problematic within the literature, as it represents a polyphyletic classification of bacteria within the families *Oscillospiraceae* and *Lachnospiraceae*. A common topic of discourse in the literature focuses on enterotypes, which provide a simplified view of the gastrointestinal microbiome and heavily rely on genus-health correlations. The reductive nature of enterotypes raises concerns, particularly when genus-level classifications, upon which these correlations are based, may be inherently flawed.

Ruminococcus gnavus, a species heavily studied for its potential relationship to inflammatory bowel disease, is more appropriately classified within *Lachnospiraceae* as *Mediterraneibacter* in the taxonomic dataset. This species is transiently detected post-antibiotic treatment. *Ruminococcus gauvreauii*, another member of the family *Lachnospiraceae*, lags in reclassification within the taxonomic data and is the only surviving member of the genus post-antibiotic therapy.

Two members of the *Ruminococcus* genus lost during Antibiotic were *R.* bromii and *R. bovis*. *R. bromii* is a prominent degrader of resistant starches in the gastrointestinal tract, and whose activity is essential to the effective breakdown of resistant starches by other bacteria (Ze et al., 2012), *R. Bovis* is a closely related species to *R. bromii with* a similar fermentative capacity. The level of abundance of *Ruminococcus* within the data set had it fallen below the cut-off for prominent genera and differential abundance analysis. However, the loss of key species within the genus and the saccharolytic capacity could be an area for future investigations within the dataset, with genus-level observations guiding the exploration of related functional pathways.

Faecalibacterium is a genus known for its short-chain fatty acid producing capacity. Notable for the production of formate, butyrate and d-lactate. They are some of the most abundant producers of Butyrate within the gut(Mallott and Amato, 2022) Previously identified as *Fusobacterium*, the inaugural species *F. prausnitzii* was differentiated due to its *requirement* of acetate for growth (Duncan et al., 2002). Highly abundant within the healthy intestinal microbiome, the genus was severely depleted by the antimicrobial therapy used here.

Depletion of *Faecalibacterium* has been noted for its association with dysbiotic microbiomes in Crohn’s disease patients and is noted as possessing a potentially ameliorative capacity *in vitro* (Sokol et al., 2008). The reduction of *Faecalibacterium* strains noted here could lead to a net reduction of SCFA production capacity within the microbiome. It should be noted, however, that the genus was not identified within the MiGut Steady state sample selected for NGS analysis. Typically, present as an abundant genus within the microbiome and were identified within the Leeds HGM steady-state sample, this could signal a failure to sufficiently colonise one of the models in this instance, due to the lack

of data from additional models it's not immediately clear if this observation persists across all replicates or is an outlying observation.

Oscillibacter species were severely depleted through antibiotic instillation, closely related to the genus *Clostridium* and *Oscillospira*; there are functional associated with the production of the short-chain fatty acid Valeric acid from a variety of carbon sources (Iino et al., 2007) and Butyrate via the glutarate pathway (Mallott and Amato, 2022).

The genus *Coprococcus* saw a depletion of species in response to antibiotics. Two of the three identified species, *C. comes,* and *C. catus*, are present within the data set with both being lost following antibiotic instillation. Members of the genus are identified as being important producers of short-chain fatty acids within the microbiome (Mallott and Amato, 2022). *C. catus* produces butyrate and propionate while C. *come* produces butyrate with formate or lactate in a substrate dependent manner *(*Vital *et al., 2014).* The depletion of species within *Coprococcus* could signal the collapse of vital SCFA production pathways within the intestinal metabolome.

The reductions see in species-level diversity within the observed genus in many instances may not indicate a complete loss of capacity but rather a reduction in functional redundancy. They areas outlines above pay particular focus on saccharolytic and SCFA function within metabolic genus correlations. The reductions outlined here have the potential to limit pathways variations and metabolic cross-feeding, as a result of antibiotic administration, effect which may persist for multiple weeks post-therapy.

Genus Level Alpha Diversity

Due to issues during library preparation, there was a lack of data for MiGut at what would be presumed to be the peak disruption of the microbiome following TZP instillation. Consequently, it is not possible to determine whether this trend would continue or reverse resulting in a decrease in genus-level diversity following continued antibiotic instillation.

The Leeds HGM returned to Shannon diversity levels matching or exceeding steady state despite being severely depleted throughout antibiotic instillation. Despite this, the model lacked intact colonisation resistance to prevent the growth and proliferation of *C. difficile* or *K. pneumoniae* within the model. While

diversity is a useful and widely used metric for microbiome research, providing a quantifiable point of reference for the state of the microbiome, it is of limited use without context.

Differentially abundant Genus – Post recovery

Exploring differential abundance for the top 25 genus highlights changes in microbiome composition. As discussed previously both *Coprococcus* and *Faecalibacterium* were depleted below detectable levels, leading to them being identified as being differentially abundant post-recovery.

Bacteroides were identified as being differentially abundant with the proximal (vessel 1) region of the model. The lack of recovery in this vessel may be influenced by the fact that they were initially less abundant than in vessel 3 (distal). Further influenced by the fact there are fewer opportunities for reseeding. Populations within the distal region can be reseeded by residual populations in vessel 1, whereas there are no up-steam sources for vessel 1.

Sutterella levels were reduced in vessel 3. The genus and its contributions to the intestinal microbiome are less profiled than some of the genera within this chapter. There have been links between therapeutic failure in ulcerative colitis patients receiving faecal microbiota transplantation and its potential to degrade IgA (Kaakoush, 2020). The lack of information about *Sutterella* and its contributions to gut homeostasis, means that it is difficult to draw any meaningful functional insights which could be explored as a part of further metabolomics investigations or functional pathway analysis.

Escherichia and *Shigella* showed increased abundance in vessel 3 following antimicrobial instillation. *Escherichia,* in particular, remained one of the most prominent genera throughout antibiotic instillation in the Leeds HGM, in vessels 1 and 3. While levels decreased post-antibiotic instillation, they remained higher than pre-antimicrobial levels

Hungatella is associated with the production of SCFAs such as butyrate (Lewis, G. et al., 2019) and is among the two SCFA producers which increase in abundance post-antimicrobial instillation. Species diversity within the genus remained consistent throughout antibiotic instillation. *Pediococcus* is a genus linked to SCFA production, displaying an increase in abundance post-antimicrobial therapy. This increase was particularly notable in proximal regions

vessel 1) of both the MiGut and Leeds HGM platforms. Although expansion was observed in vessel 3, they failed to achieve statistically significant p-values despite having log2fold increase greater than one. The expansion of *Hungatella* and *Pediococcus,* along with their capacity to produce SCFAs, may provide some compensatory effects for the reduction in other genera.

Analysis of genera, though limited in its specificity, begins to outline a potential shift in the metabolic capacity of saccharolytic bacteria and those associated with the production of SCFAs. Initial exploration of the functional metagenome showed few statistically significant changes. However, the functional characteristics associated with disrupted genera outlined here allowed for a targeted bottom-up integration of functional pathways

4.6.2.3 Functional Analysis of the Metagenome

Metagenomics can offer valuable insights into the functional potential of microbial communities. It is useful for identifying potential functional pathologies in dysbiotic states though these impacts may be amplified or compensated down by downstream changes. The observations from Metagenomic studies can inform other mechanistic studies such as transcriptomics, proteomics, and metabolomics, which can reveal how metagenomic level changes manifest impacting host health. Observations from metagenomic analysis in this chapter will guide future metabolomic investigations.

Initial functional investigations at subpathway and sub-subpathway levels showed no sustained notable changes following antibiotic therapy and post antibiotic recovery period. This suggest that any lasting metabolic changes induced may be too subtle to detect at these levels and may have been masked by highly conserved enzymes within pathways remaining consistent.

As mentioned throughout taxonomic analysis, many of the key genera identified as being potentially affected by antibiotic instillation are associated with SCFA production. As it would not be feasible to explore all 314 sub-subpathways within KEGG Pathway identified in this chapter, those associated with the key metabolic pathways for SCFA.

Butyrate Synthesis and Metabolism

Within the context of butanoate metabolism outlined in KEGG pathway ko00650, there was a reduction in Glutaconate CoA-Transferase subunits A and B. This enzyme is a key step in the production of butyrate via the Glutarate pathway and is responsible for the conversion of 2-Hydroxyglutarate to 2-hydroxy-glutaryl-CoA producing Acetate as an additional product, outlined in Figure 4-18 Glutarate Pathway for Butyrate Synthesis. This reduction in glutaconate CoA-transferase suggest a potential bottleneck within this butyrate synthesis pathway which could lead to decreased Butyrate production.

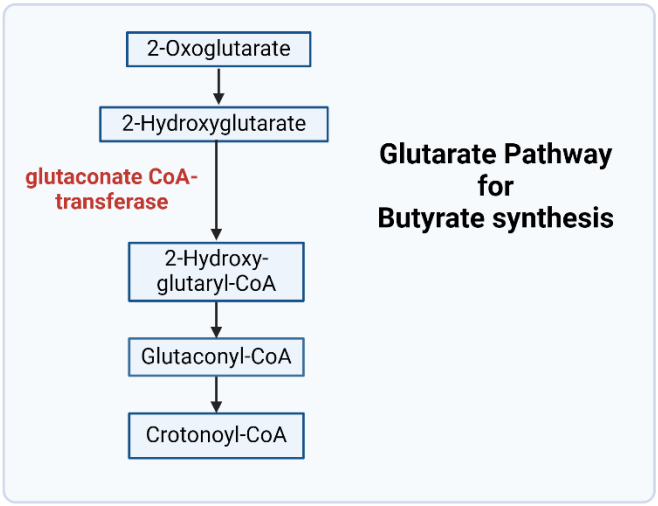

Figure 4-18 Glutarate Pathway for Butyrate Synthesis. Glutaconate CoA Transferase (EC:2.8.3.12), responsible for the conversion of 2-Hydroxyglutarate to 2-hydroxy-glutaryl-CoA identified in red. Outlining a potential bottleneck in the synthesis of Butyrate via this pathway. Figure outlines unique steps to Glutarate pathway to Crotonoyl-CoA where major synthesis pathways converge.

The glutarate pathway is not prominently associated with the bacterial genera identified as being reduced in abundance by DESeq analysis - *Coprococcus* and *Faecalibacterium* being more strongly associated with the acetyl-CoA pathway (Vital et al., 2014).Despite the depletion of genera associated with the production of butyrate via the acetyl-CoA pathway. The glutarate pathway appears to be conserved mainly within the *Fusobacteria and Oscillibacter (*Mallott and Amato*, 2022)*. The association with *Oscillibacter* being notable due to their depletion by antibiotic instillation. Depletion of *Oscillibacter* potentially driving the reduction in glutaconate CoA-transferase abundance. The pathway,

being restricted predominately these genera, may, by extension be vulnerable to disruption. Due to the low abundance of the pathway(Mallott and Amato, 2022), the impact that it may have on overall butyrate levels within the system remains unclear, it is hopeful that further metabolomic analysis could provide valuable insights, into how these changes in metabolic capacity may resolve.

Although not statistically significant, a trend towards the reduction of genes encoding enzymes involved in acetyl-CoA metabolism within butyrate metabolism was observed. This trend may be of interest, particularly in the context of the depletion *Fusobacteria* and *Oscillibacter,* both of which are associated with the acetyl-CoA pathway as described above. Further investigation is needed to determine whether the depletion of these genera may be associated with a significant loss of functional capacity within butyrate metabolism

Additionally, there was an increase in 4-hydroxybutyrate dehydrogenase / sulfolactaldehyde 3-reductase and acetolactate synthase II small subunit.

4-hydroxybutyrate dehydrogenase / sulfolactaldehyde 3-reductase is an enzyme with dual functionality. In the context of butanoate metabolism, it is responsible for converting succinate semialdehyde into 4-hydroxybutyrate (Nirenberg and Jakoby, 1960). The reaction is reversible and, depending on the presence of substrate, could lead to increased production of succinate semialdehyde or 4-hydroxybutyrate. Computational data from MetaCyc suggest that under normal conditions, the directionality of the reaction would favour the production of 4-hydroxybutyrate. Due the reversible nature of catalysis, additional study would be required to better characterise the functional implications of these findings under disrupted conditions. 4-hydroxybutyrate is significant as it is an intermediate compound in the synthesis of butyric acid and gamma-amino-butyric acid. A minor pathway for the synthesis of butyric acid, 4-hydroxybutyrate can either be converted directly through the terminal enzyme butyryl-CoA:4-hydroxybutyrate CoA transferase, or through the metabolism to crotonyl-CoA (Vital et al., 2014). This finding may not be highly impactful on the production of butyrate or other SCFAs within the but serves to highlight the complexity of metabolic interactions and the potential for systemic interactions affecting the body beyond the gut.

Acetolactate synthases are predominately associated with the formation of branched-chain amino acids, in addition to butyrate metabolism. Consisting of a large catalytic subunit and three smaller regulatory subunits which are involved in feedback inhibition binding to metabolic products such a valine. In addition to amino acid formation, acetolactate synthase has been identified as a possible method of pH modulation for the internal conditions of bacteria and cyanobacteria (Maestri and Joset, 2000), and its regulation may be increase in the presence of excess carbohydrates. An increase in the acetolactate synthase II small subunit could indicate the decreased availability of carbohydrates, possibly due to the reduction of saccharolytic bacteria within the microbiome. Therefore, there would be a lower requirement to reduce intracellular pyruvate. Observing a change in the regulatory subunit of acetolactate synthases, it is possible that this may be indicative of wider changes to nutrient availability, particularly related to the breakdown of nondigestible carbohydrates within the model.

Propanoate Synthesis and Metabolism

There were notable reductions in two metabolites associated with propanoate metabolism.

Butyryl-CoA dehydrogenase, an enzyme involved in the production of propionate from lactate, catalyses the conversion of acryloyl-CoA to propanoyl-CoA, a central metabolite across propanoate synthesis pathways.

Multifunctional beta-oxidation protein abundance was also reduced following the antibiotic recovery period; although not directly involved in propanoate synthesis, multifunctional beta-oxidation protein accepts hydroxy-propionyl-CoA as a substrate converting it to acryloyl-CoA. Notably, both of these enzymes are associated with the production of propionate from lactate via the acrylate pathway and are associated with the firmicutes, *Clostridiaceae* and *Clostridioides* (Reichardt et al., 2014). *Clostridioides*, species diversity and abundance appeared to be resilient to the courses of antimicrobial therapy employed throughout the dosing period of the model. *Coprococcus* species (*C. catus*) have been shown to potentially possess the capacity for propionate production from lactate, and their depletion in both overall abundance and species diversity may be attributable to this observation.

Propionate kinase was associated with an increase in abundance following the antibiotic recovery period. Propionate kinase is part of a metabolic pathway in *Escherichia coli*, synthesising propionate from L-threonine. It catalyses the terminal step of this pathway, converting propanoyl phosphate to propanoate. During antibiotic treatment persisting through the antibiotic recovery phase, there was a significant increase in the abundance of *Escherichia*, while *Coprococcus* decreased, indicating a potential shift in propionate metabolism.

This observation could be emblematic of functional redundancy within the microbiome, where different bacteria occupy similar metabolic niches. This shift may resolve by stabilising the microbiome to an extent, with *Escherichia* filling the metabolic void. There may, however, be additional ramifications if the microbiome transitions from utilising the lactate/acrylate pathway. The production of propanoate from propanoyl-phosphate via propionate kinase is unique in that it doesn't produce propanoyl CoA as an intermediary, a compound cross feeding other metabolic pathway.

This identifies as the abundance of propanoate within the model, a potential target for future metabolic investigations. There have been studies exploring the antimicrobial properties of propionate *in-vitro* and *in-vivo*, outlined effectively in a systematic literature review (Langfeld et al., 2021), with it displaying bacteriostatic and bactericidal properties to concentrations as low as 10mM.

Acetate Synthesis and Metabolism

Acetate production is detailed within the KEGG pathway outlining pyruvate metabolism. These processes are versatile, involving multiple pathways from a variety of starting substrates such as acetyl-CoA and ethanol. The range of substrate and enzymatic diversity provides an implicit degree of resilience to acetate synthesis, which may not be seen in other SCFA pathways. Acetate, however, is a crucial substrate in the pathways for the formation of other SCFAs, such as in the terminal step of butyrate production where butanoyl-CoA and acetate yield butyrate and acetyl-CoA, itself a central metabolite in acetate production. These inter-pathway interactions illustrate the complexity and cross-feeding of metabolic pathways, suggesting that fluctuations in acetate may be inherently linked to broader carboxylic acid synthesis within the microbiome.

Following the antibiotic recovery period, there was a significant increase in methylglyoxal/glyoxal reductase. This enzyme catalyses the reduction of methylglyoxal to (S)-lactaldehyde. Depending on substrate availability, this may result in an increased concentration in (S)-lactaldehyde, potentially increasing the availability of L-lactate, a precursor to acetate. The specific enzyme annotation K23257 identified in this instance has limited associated literature, described in *Bacillus* species (Sakai et al., 2001). In contrast, methylglyoxal/glyoxal reductase is broadly reported in *Saccharomyces* (Murata et al., 1985). Whether this increase has the capacity to induce changes in acetate production is unclear as there was no increase in the abundance of the (S)-lactaldehyde: NAD+ oxidoreductase, the enzyme responsible for the formation of (S)-lactate and antibiotic instillation was associated with reductions in the abundance of enzymes associated with the synthesis of acetate production from lactate.

A decrease in the abundance of L-lactate dehydrogenase (cytochrome) was observed. This enzyme plays a key role in the production of acetate from L-lactate via pyruvate and could represent a bottleneck in the acetate production pathway, despite potential increases in lactate production capacity by the microbiome. There was also a reduction in (R)-2-hydroxyglutarate: pyruvate transhydrogenase abundance, which catalyses the formation of pyruvate from R-lactate, though this did not reach levels of significance (>0.05). Additionally, lactate 2-monooxygenase was identified as being differentially abundant. However, this is discussed as being either a hypothetical protein or an uncharacterised protein, calling into question the validity of this assignment.

The changes induced by the potential for acetate production by the microbiome portray a complex web of metabolic interactions. Whilst there is a potential increase in the capacity for the microbiome to form key metabolites within the pathway, there are notable potential bottlenecks downstream.

4.6.3 Summary

While taxonomy is integral to the discussion in this chapter, it is insufficient to fully characterise dysbiosis. Significant alterations in enzyme abundance within pathways for butyrate propanoate and acetate metabolism highlight these compounds as candidates for metabolomic investigations.

Chapter 5
Metabolomics

5.1 Background

This chapter investigates changes to the *in-vitro* metabolome after antibiotic exposure, building on insights from previous chapters.

In Chapter 3, an *in-vitro* model experiment explored the compositional dynamics of the microbiome through the isolation and enumeration of key bacterial taxa during antibiotic instillation and recovery. Monitoring of key genera revealed notable shifts in the microbiome throughout antibiotic treatment. Following a 3-week recovery period, culture analysis showed that at a genus level, the microbiome returned to a composition similar to that observed prior to antibiotic instillation. However, challenging the microbiome with pathogens resulted in the colonisation and proliferation of introduced pathogens. This suggests that colonisation resistance was effectively reduced despite similar microbiota composition

In an effort to gain a greater understanding of the dynamics, a metagenomic analysis was performed, exploring both the taxonomic and functional composition of the microbiome Chapter 4. The increased sensitivity of the taxonomic analysis from next-generation sequencing identified a shift in the abundance of genera with important functional associations. Depletion of genera *Faecalibacterium and Coprococcus,* important SCFA producers, and reductions in species diversity within genera, such as *Oscillibacter* and *Ruminococcus*, suggest potential shifts in functional capacity or redundancy within the microbiota. This paints a more nuanced picture of the microbiome post-antibiotic therapy and recovery.

The metabolomic analysis in this chapter seeks to explore how these compositional and functional shifts resolve at the phenotypic level. Untargeted metabolomics was used to profile the gut metabolome, investigating gross changes across the metabolome and exploring metabolites within pathways identified as differentially abundant in the metagenomic analysis from the previous chapter. The aim is to identify putative functional biomarkers at the phenotypic level of antibiotic-induced gut dysbiosis.

5.2 Chapter Aims

- **Explore Phenotypic Resolution of microbial disruption** - Investigate how compositional and functional shifts identified in previous chapters resolve within the metabolome.
- **Targeted Interrogation of Functional Shifts**: Exploration of putative functional alterations observed through metagenomic analysis.
- **Identify Putative Biomarkers of Dysbiosis**:

5.3 Experimental Outline

Samples for LCMS analysis were collected from the *in-vitro* models described in Chapter 3. The analysis covers samples from the Leeds HGM model (outlined in Figure 3-2 The Leeds human gut model (HGM) and configuration overview) and three replicates of the MiGut platform (Figure 3-3 Details of the Migut platform and model.).

Although metagenomic analysis revealed some variation between Leeds HGM and MiGut, and the effect of antibiotics, both systems demonstrated a loss of colonisation resistance, demonstrated by the colonisation and proliferation of pathogens. Consequently, data from both systems were combined to identify a robust panel of putative metabolic markers of antibiotic-induced dysbiosis. Sample points were aligned with key experimental stages: following the end of the steady-state period, after each antibiotic instillation period, and immediately before the addition of the next antibiotic. Follow the three-week dysbiosis recovery period and nine days after the introduction of pathogens (*C. difficile* and *K. pneumoniae*).

Time points aligned with those used for metagenomic analysis with two exceptions. For the end of steady state, metagenomic analysis was conducted on samples taken on Day 10, while samples from Day 13 were used for metabolomic analysis for all models. A sufficient sample was not available for both metagenomic and metabolomic analysis, and therefore, the next closest date was chosen. This deviation was deemed unlikely to cause a significant impact on analysis; as indicated by culture analysis, the model was in a stable condition by this point with no intervention. For TZP, Day 34 was initially chosen for both metagenomic and metabolomic analysis. However, this date failed to

create libraries of sufficient quality for NGS, as described in Chapter 4 Metagenomics 4.3.3.1 Library Preparation.

For the analysis in this chapter, samples were taken from one Leeds HGM replicate and three MiGut replicates. Samples from vessel 3 were used for metabolomic analysis. Vessel 3 is representative of the distal colon and, therefore, the most appropriate for identifying faecal biomarkers. Samples from all key time points; steady state, antibiotic dosing (amoxicillin, ciprofloxacin, piperacillin-tazobactam), post-antibiotic recovery (P-Abx, and pathogen challenge(P-Path) (Figure 5-1 Experimental Outline: Metabolomic investigation into antibiotic-induced dysbiosis).

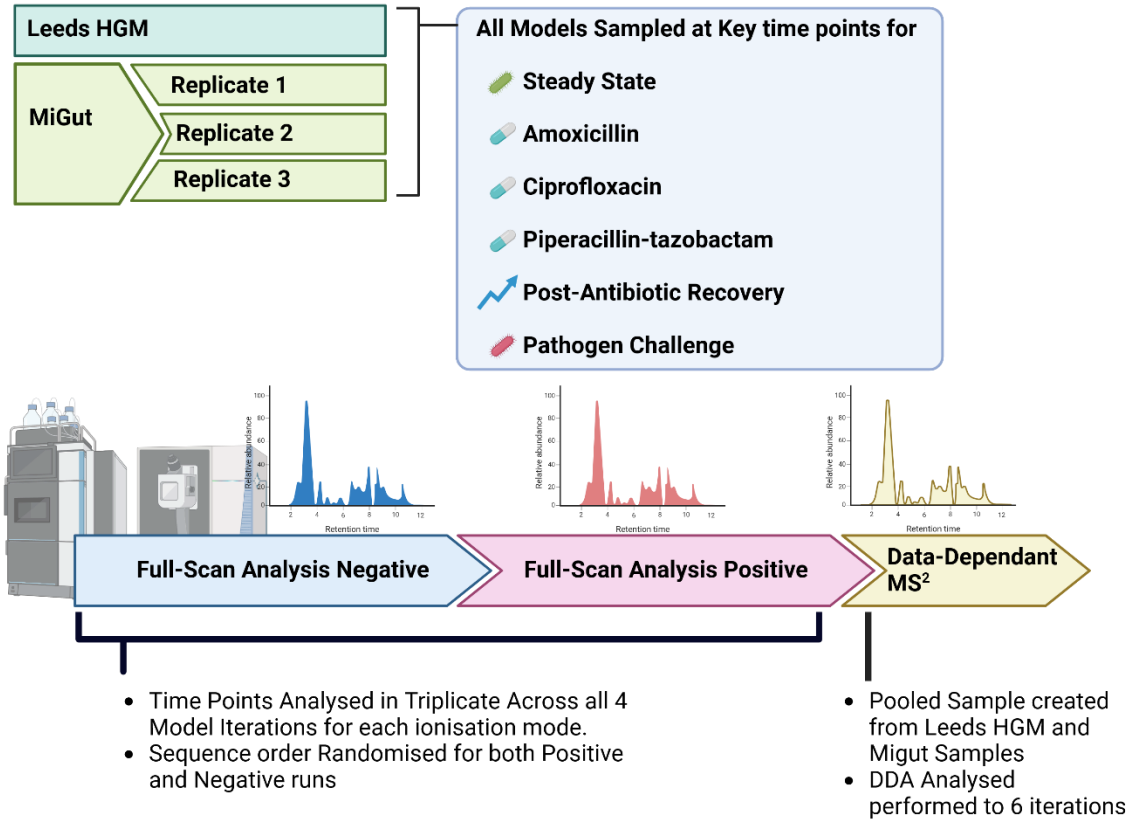

Figure 5-1 Experimental Outline: Metabolomic investigation into antibiotic-induced dysbiosis LCMS workflow for exploring metabolomic changes associated with antibiotic-induced dysbiosis. Samples from the Leeds HGM were taken from vessel 3 of a single model replicate. Samples were taken from vessel 3 of three MiGut replicates.

Sampling resulted in 24 samples in total, 6 from the Leeds HGM across all sample points and 18 across MiGut replicates. Samples were extracted and analysed in triplicate, with 3 x 5µL injections. Quality control and blanks were

injected every 5-6 samples. Positive and Negative analysis was performed in two consecutive experiments, followed by MS^2 data-dependent acquisition (DDA). All experiments consisting of approximately 190 injections were performed over a single MS^2 run spanning 3 days.

5.4 Methods

Methods used throughout this chapter were developed and validated in Chapter 2 Metabolomic Methods Development.

5.4.1 Sample Collection and Processing

5.4.1.1.1 Collection and Storage

Samples were collected throughout the *in-vitro* models at time points corresponding to major experimental phases. The were collected aseptically with sterile pipettes and stored in sterile 1.5mL microcentrifuge tubes. Samples were then centrifuged at 5000g for 20 minutes and filter sterilised through a 0.22µm polyether sulfone (PES) syringe filter into sterile 1.5mL microcentrifuge tubes. They were then stored at -80ºC until ready for extraction.

5.4.1.1.2 Sample Processing

Samples were batched until the end of the model run and processed using a solvent-based metabolite extraction, as outlined in detail in Chapter 2 Metabolomic Methods Development, 2.4.2 Solvent-Based Metabolite Extraction. Samples were prepared as described in a 1:3 sample-to-solvent ratio methanol extraction. After extraction, samples were resuspended in 5% methanol for LCMS analysis

5.4.1.1.3 Pooled Sample Creation

A 10µL aliquot of each of the 24 experimental samples was used to create a pooled sample. This pooled for QC correction across the LCMS run and served as a sample for compound identification within the DDA experiment.

5.4.2 LCMS Analysis

LCMS analysis was performed, with untargeted analysis conducted on all experimental time points, as outlined in Chapter 2 Metabolomic Methods Development, 2.5 LCMS Parameters A DDA Deep scan experiment was performed on the pooled sample to generate MS^2 data for compound identification and annotation.

5.4.2.1 High-Performance Liquid Chromatography

HPLC was performed using the parameters outlined in Chapter 2 Metabolomic Methods Development, 2.5.1 High-Performance Liquid Chromatography , using C18 columns (Hypersil GOLD C18 Selectivity Reverse Phase column, 2.1 mm x 150 mm) and a guard column (2.1 mm x 10 mm).

5.4.2.2 Mass Spectrometry

Mass Spectrometry analysis was conducted on the Exploris 240 Orbitrap system (Thermo Fisher Scientific), using heated electrospray ionisation (ESI) as the ionisation source. The parameters used are the same as those outlined in Chapter 2 Metabolomic Methods Development 2.5.2Mass Spectrometry

5.4.2.2.1 Full Scan

Full scan Experiments in positive and negative ionisation modes were performed on all experimental samples across all models.

5.4.2.2.2 Deep Scan MS^2

A DDA MS^2 experiment was performed to provide MS^2 data for compound identification and annotation. Three deep scan iterations were used for compound coverage.

5.4.2.3 Data Processing

Data Processing and analysis were performed as outlined in Chapter 2, 2.6 Data Processing and Analysis. Metabolomic data was processed using compound Discoverer v3.3. In-house metabolite libraries generated from IROA GUTMLS and BACSMLS libraries, described in Chapter 2, 2.8 1.8 Local Library Development – Processing and Annotation, were used for compound annotation.

- **HMDB** – Version 5.0 was used as a basis for compound classification
- **MzCloud** – Final access date 01/11//2024 – was used for compound annotation

5.4.2.4 Statistics and Analysis

5.4.2.4.1 Distance Matrices

A Bray-Curtis dissimilarity matrix was calculated separately for metabolite data from positive and negative ionisation modes. No data transformation or normalisation was applied. Distance matrices were calculated in R with the `vegan` package function `vegdist` with `method="bray"`.

5.4.2.4.2 Non-metric Multidimensional Scaling (NMDS)

NMDS was performed to visualise the relationship between sample points based on the generated Bray-Curtis dissimilarity matrices. NMDS was chosen due to its effectiveness in resolving non-linear, complex, and high-dimensional relationships. NMDS analysis was performed in R with the vegan pack function `metaMDS.` Analysis was performed with default parameters (20 Runs for each data set. Data points were extracted and plotted using `ggplot2`.

5.4.2.4.3 Permutational Multivariate Analysis of Variance (PERMANOVA)

PERMANOVA was Implemented to test the significance of differences within the metabolome based on the experimental Stage (Steady-State, Amoxicillin, Ciprofloxacin, Piperacillin-tazobactam, Post-Antibiotic Recovery, Post Pathogen) and Model type (Leeds HGM and MiGut). PERMANOVA analysis was performed in R using the `adonis` function within the `vegan` package. Default parameters were used performing 999 permutations, and a significance threshold of <0.05 was used.

5.4.2.4.4 Differential Abundance Analysis

Differential abundance analysis was performed using Compound Discoverer 3.3 (Thermo Fisher) within the post-processing descriptive statistics workflow, employing default parameters. Compound discoverer was chosen over DESeq (used in Chapter 4 Metagenomic investigation of the *In-vitro* Intestinal microbiome. and typically employed for count-based data), as it is specifically designed for LCMS data analysis, making it more suitable for this task.

5.4.2.4.5 Box and Whisker plots

Box and whisker plots for compound abundances were generated using geom_`boxplot` from the R package `ggplot.` Lower and upper hinges

represent the 25th and 75th percentiles. The upper whisker extends to the largest value within 1.5 of the interquartile range. The lower whisker extends to the smallest values within 1.5 of the interquartile range. Data beyond 1.5 of either inter-quartile range as designated as outliers and plotted individually

5.5 Results

5.5.1 Overview of Compound Features Detected by LCMS

The metabolomic data obtained from the LCMS was analysed based on the mode of ionisation. A total of 31,860 features were detected across all models and sample points over both ionisation modes, with 15,342 features in negative ionisation mode and 16,518 in positive ionisation mode. Of the compounds detected, 2207 (14.39%) in negative Mode and 2103 (12.73%) in positive mode were assigned named annotations (Confidence level 3 as discussed in 2.9.3 Compound Annotation Confidence).

Data Dependent MS^2 acquisition resulted in spectral data for the majority of features within the data set; Negative ~60% (preferred ion 7539, other ion 1913, no MS^2 5890) and Positive ~71% (preferred ion 6514, other ion 5240, no MS^2 4764). The MS^2 data resulted in 225 spectra-based annotations in negative mode and 218 spectra-based annotations in positive mode (Figure 5-2 Experimental compound annotations by highest-level identification.

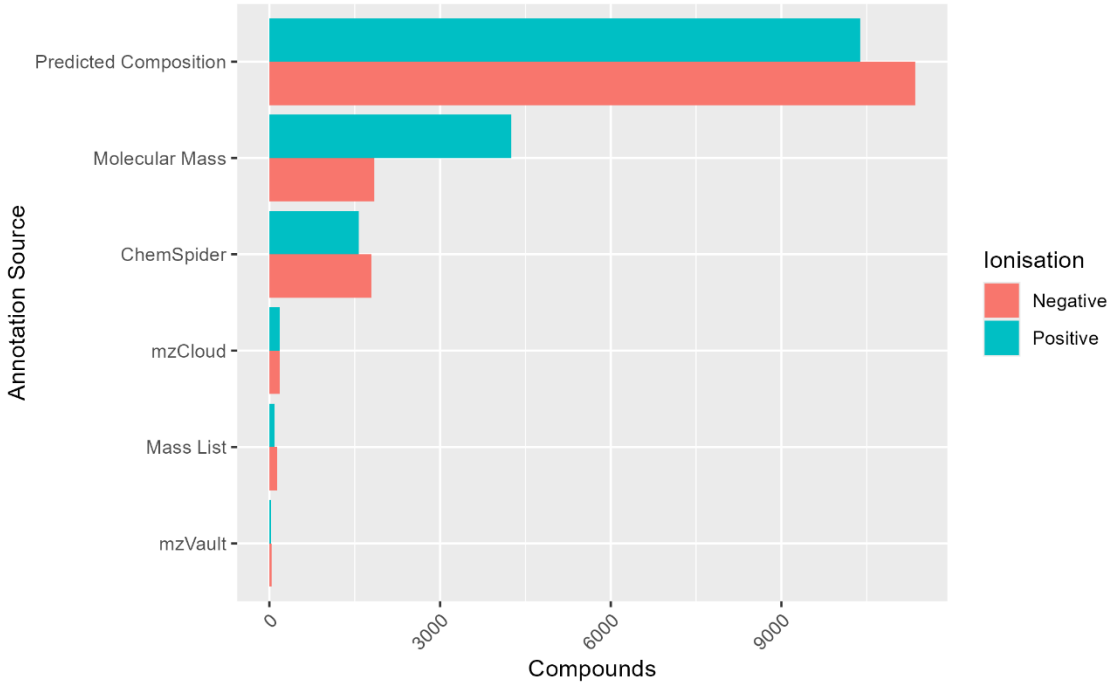

Figure 5-2 Experimental compound annotations by highest-level identification. Distribution of compound annotations across data sources

5.5.2 Compositional Alterations to the *In-Vitro* Metabolome

Pairwise distance matrices were computed between all Leeds HGM and MiGut samples using the Bray-Curtis dissimilarity index for data obtained from positive and negative LCMS analysis. Non-metric multi-dimensional scaling (NMDS) was used to represent data. Stress values were obtained for positive and negative data sets, which indicated reliable representations of the data in reduced dimensions (negative 0.1019533, positive 0.0954939). As with metagenomic data reducing the dimensionality provides an intuitive assessment of beta diversity, alluding to compositional alterations throughout antimicrobial instillation.

5.5.2.1 Compositional Alterations Detected by Negative LCMS

Negative ionisation mode effectively captures the shifts occurring in the in-vitro metabolome. The NMDS plot reveals a clear delineation between two groups of sample points. (Figure 5-3 NMDS Analysis of Metabolome Changes Detected in Negative Ionisation mode). Group 1 comprises steady-state and AMX sample points, whereas Group 2 includes points associated with the antibiotic dosing periods for CIP, TZP and the P-Path phases following the 3-week antibiotic recovery period and pathogen challenge.

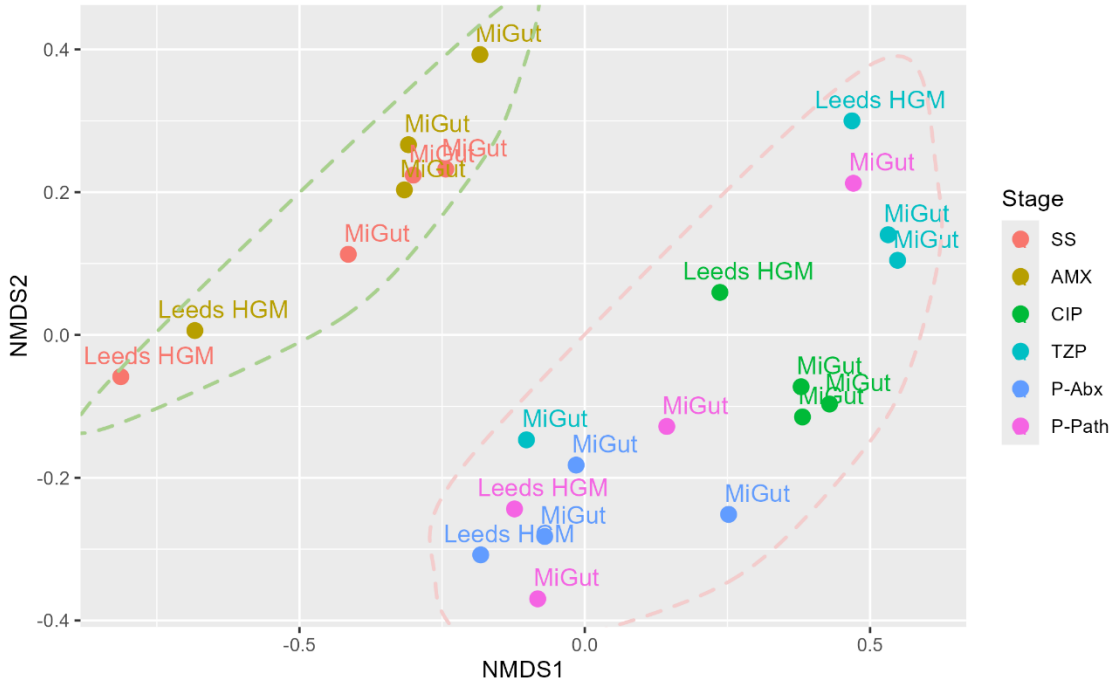

Figure 5-3 NMDS Analysis of Metabolome Changes Detected in Negative Ionisation mode . Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green) , Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink). Sample Point Grouping: Group1: SS and AMX (Green), Group2: CIP,TZP,P-Abx and P-Path (Red)

Among the antibiotics tested, samples following amoxicillin instillation show the least deviation from the pre-antibiotic steady state, aligning with observations in the changes to taxonomic composition detected through metagenomic analysis.

Within Group Two, three distinct clusters appear to emerge: one predominantly comprising TZP samples, another of CIP samples and a final cluster comprising P-Abx and P-Path samples. Notable, there are two samples from TZP and P-Path time points from a single MiGut model that deviates from the clustering of

other data points for these periods. These samples display a reverse pattern relative to the clustering observed with the rest of TZP and P-Path samples.

5.5.2.2 Compositional Alterations Detected by Positive LCMS

The separation between the two groups identified in negative ionisation mode is less distinct in the analysis conducted in positive ionisation mode. The composition of the clusters also differs from the compositions identified in negative ionisation mode (Figure 5-4 NMDS Analysis of Metabolome Changes Detected in Positive Ionisation mode).

The first cluster includes samples from the steady-state, AMX, P-Abx and P-Path, and the second cluster comprises samples from CIP and TZP. The MiGut samples identified in negative ionisation as deviating from the clustering pattern exhibit a similar deviation in positive ionisation mode.

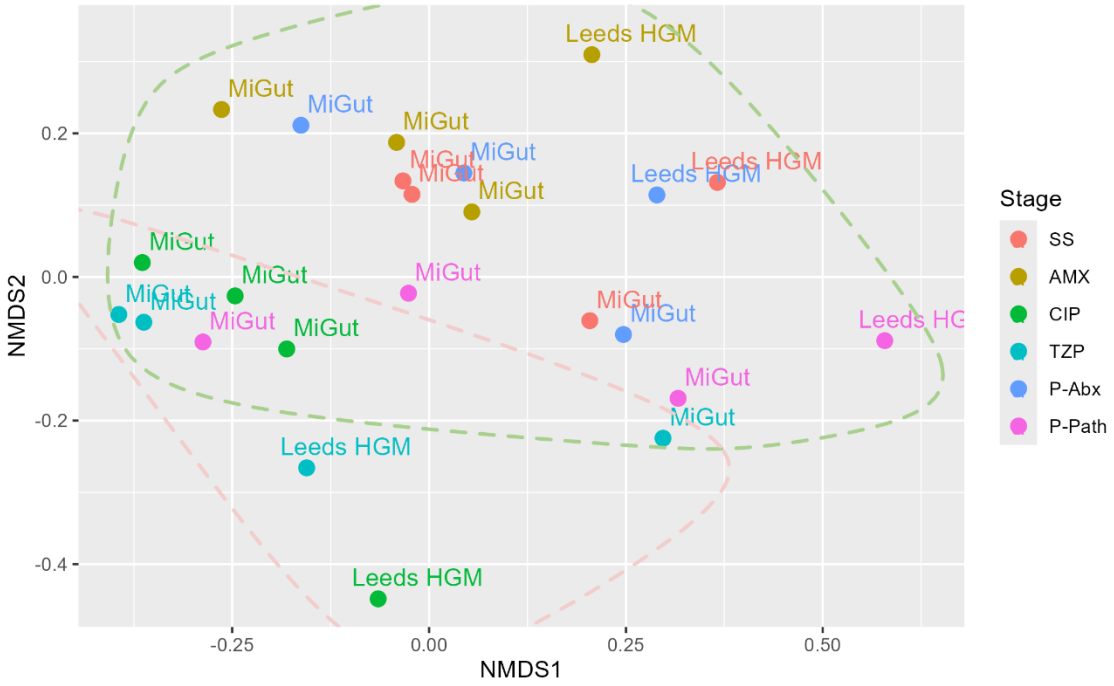

Figure 5-4 NMDS Analysis of Metabolome Changes Detected in Positive Ionisation mode. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink). Sample Point Grouping: Group1: SS and AMX (Green), Group2: CIP, TZP, P-Abx and P-Path (Red)

In positive ionisation mode, the P-Abx time points are positioned closer to the steady-state samples compared to the other potentially disrupted time periods, reflecting a trend more consistent with metagenomic taxonomic analysis.

5.5.2.3 Comparing the Resolution of Metabolomic Differences: PERMANOVA Analysis of Positive and Negative Ionisation Modes

PERMANOVA analysis was used to evaluate the effectiveness of positive and negative ionisation modes in resolving metabolic shifts resulting from antibiotic therapy. In the negative ionisation mode, experimental stage (SS, AMX, CIP, TZP, P-Path, P-Abx) explained 54.48% (F=0.001) of the variation, while in the positive ionisation mode, PERMANOVA analysis explained 35.91% (F=0.007) of the variation, suggesting that the negative ionisation mode may be more effective at describing the metabolomic changes observed under these conditions.

5.5.3 Antibiotic-Induced Alterations in Short-Chain Fatty Acid Pathways

This section explores the phenotypic resolution of the metabolomic changes observed in previous chapters, focusing on the changes in key metabolites in SCFA synthesis pathways.

Given that the compounds discussed within this section are carboxylic acids, they exist in protonated and deprotonated forms depending on pH. For the sake of clarity and consistency within the context of this research, these compounds will be referred to by their biologically relevant names, corresponding to their deprotonated forms under physiological conditions presented in Table 5-1 Key Metabolites Presented and Discussed in This Chapter, Along with Their Protonated and Deprotonated Form.

Table 5-1 Key Metabolites Presented and Discussed in This Chapter, Along with Their Protonated and Deprotonated Form

Metabolite	Protonated form	Formula (Deprotonated)	Biological Context
Acetate	Acetic Acid	CH_3COO^-	SCFA
Propanoate	Propanoic acid	$C_2H_5COO^-$	SCFA
Lactate	Lactic acid	$C_3H_5O_3^-$	Precursor Metabolite
Butyrate	Butyric acid	$C_4H_7O_2^-$	SCFA
2-Hydroxyglutarate	2-Hydroxyglutaric acid	$C_5H_7O_6^-$	Precursor Metabolite
2-Oxoglutarate	2-Oxoglutaric acid	$C_5H_5O_5^-$	Precursor Metabolite
Valerate	Valeric Acid	$C_5H_{10}O_2$	SCFA

5.5.3.1 Acetate & Propanoate

Acetate and propanoate metabolism proved challenging to investigate due to the diverse molecular weights of compounds within these pathways. Acetate (MW: 59.04) and propanoate (MW: 74.079) fall below the analytical window for the LCMS experiment performed (80-800 Da) was not detected. Compounds conjugated to coenzyme A (CoA, MW: 767.5) fall above the detectable range.

5.5.3.2 Lactate

A recurring theme in Chapter 4 was the potential changes in the abundance of compounds associated with lactate pathways within both acetate and propanoate metabolism. Lactate was therefore examined in detail, as it fell within the analytical window and could potentially be influenced by multiple pathways.

Lactate was detected across all models and present across all stages in negative ionisation mode. The compound identified as lactate had a m/z 89.02436 ([M-H]-1 ion), and a calculated molecular weight of 90.03164, an annotated mass difference of -0.64ppm and a retention time of 1.949min (Figure 5-5 Extracted Ion Chromatogram (XIC) for Compound Identified as Lactate (m/z 88.05246) In Negative Ionisation)

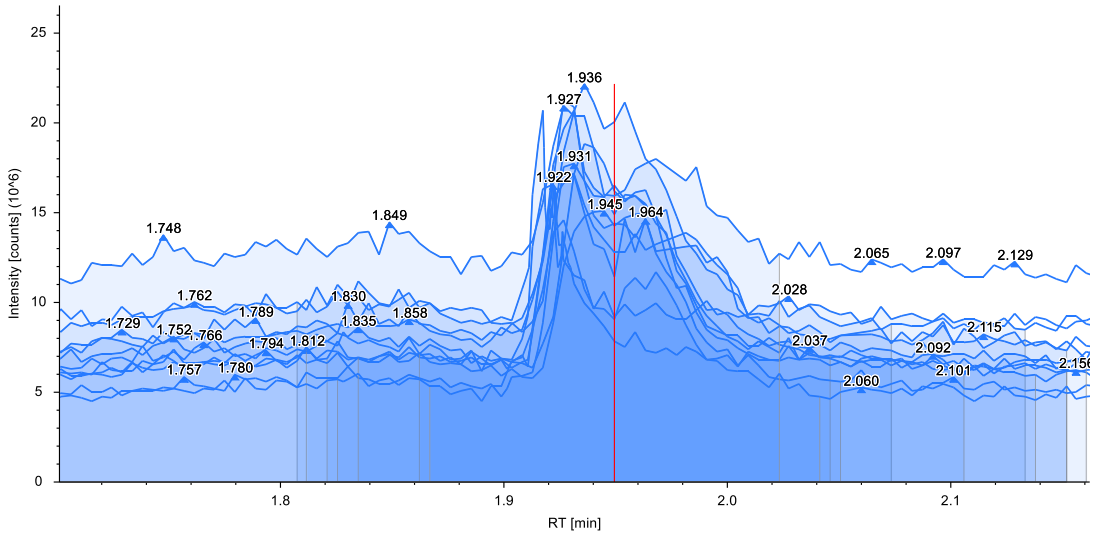

Figure 5-5 Extracted Ion Chromatogram (XIC) for Compound Identified as Lactate (m/z 88.05246) In Negative Ionisation. The region highlighted (1.7-~2.1min) corresponds to the retention time identified for Lactate. The red line at 1.949min indicates the apex of the integrated chromatographic peak.

Lactate was selected in the DDA and Fragment Ion Search (FISh) analysis identified complete coverage (FISh Coverage: 100.00) with 5 Matched fragments, MS^2 spectra presented in Figure 5-6 MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Lactate.

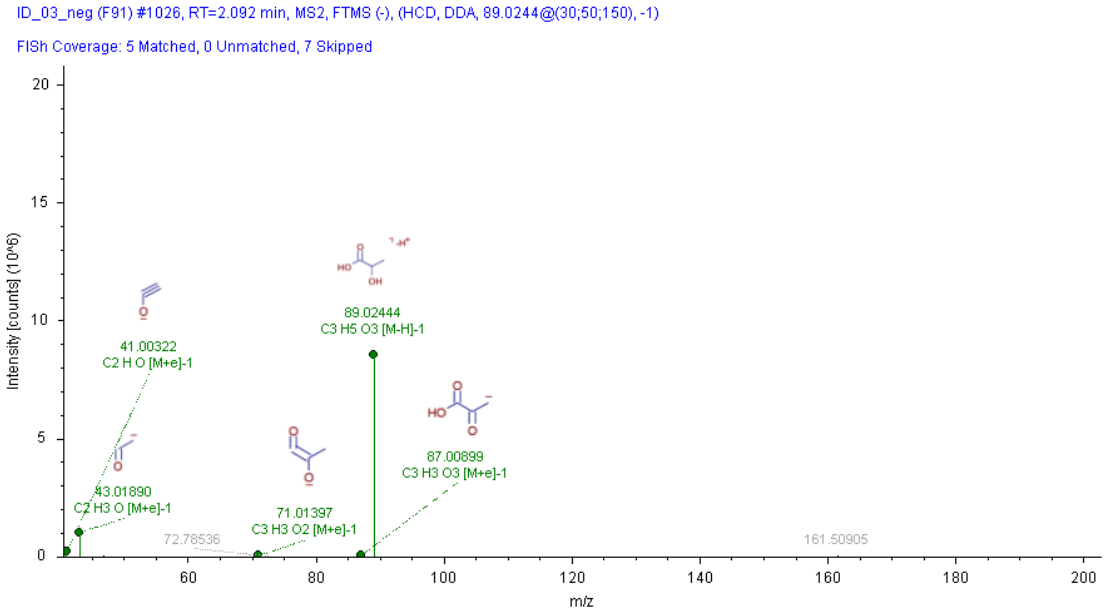

Figure 5-6 MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Lactate. Showing matched peaks and their associated fragments

Lactate: Changes in Compound Abundance Coinciding with Antibiotic Treatment and Recovery

Lactate abundance was compared across experimental phases for all models (Figure 5-7 Lactate Abundance Across Experimental Stage). Levels were lowest during steady state and remained low during AMX dosing. During CIP instillation there was an increase in lactate abundance, and levels remained at an elevated state for the remainder of the experiment.

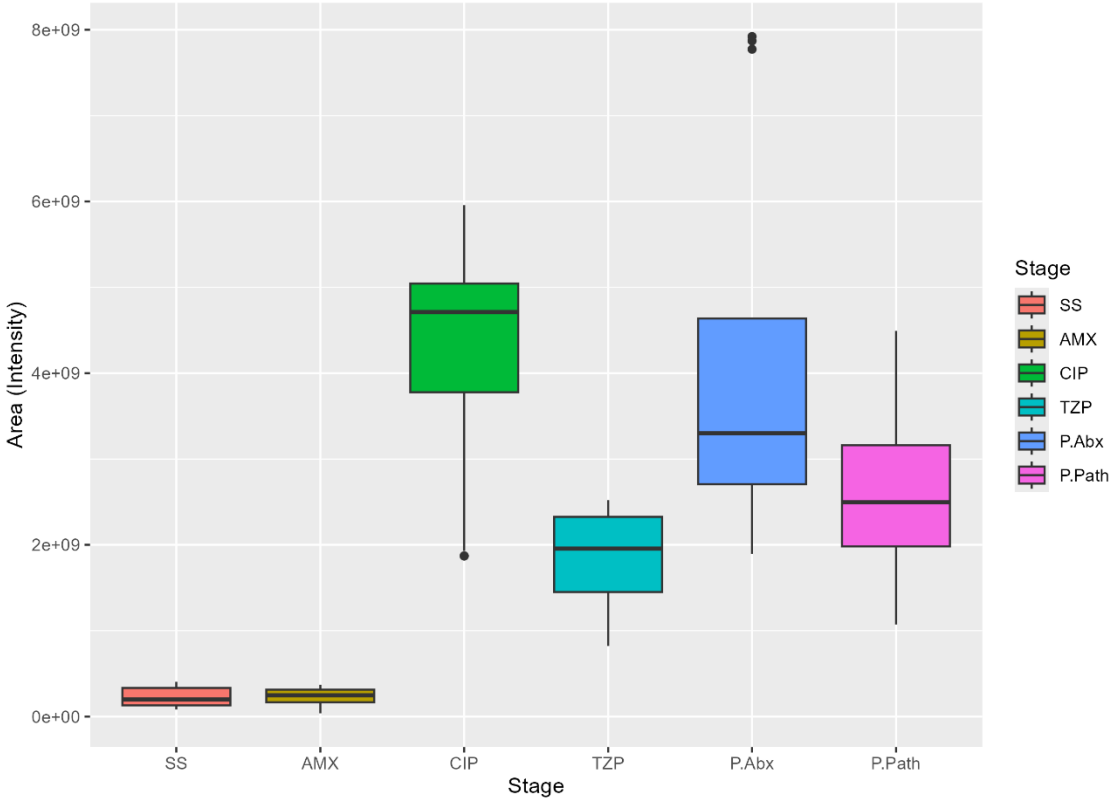

Figure 5-7 Lactate Abundance Across Experimental Stage. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green) , Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Lactate abundance showed little variation at steady state, and during AMX instillation; however, lactate abundance increased in all model systems after exposure to ciprofloxacin and remained elevated in subsequent stages.

The trend of elevated lactate levels following repeated exposure, which failed to return to baseline, was consistent across all model replicates. The MiGut systems shower greater fluctuations in lactate abundance during the disrupted

phase spanning TZP – P.Path (Figure 5-8 Lactate Abundance Across Experimental Stage and Model Replicates)

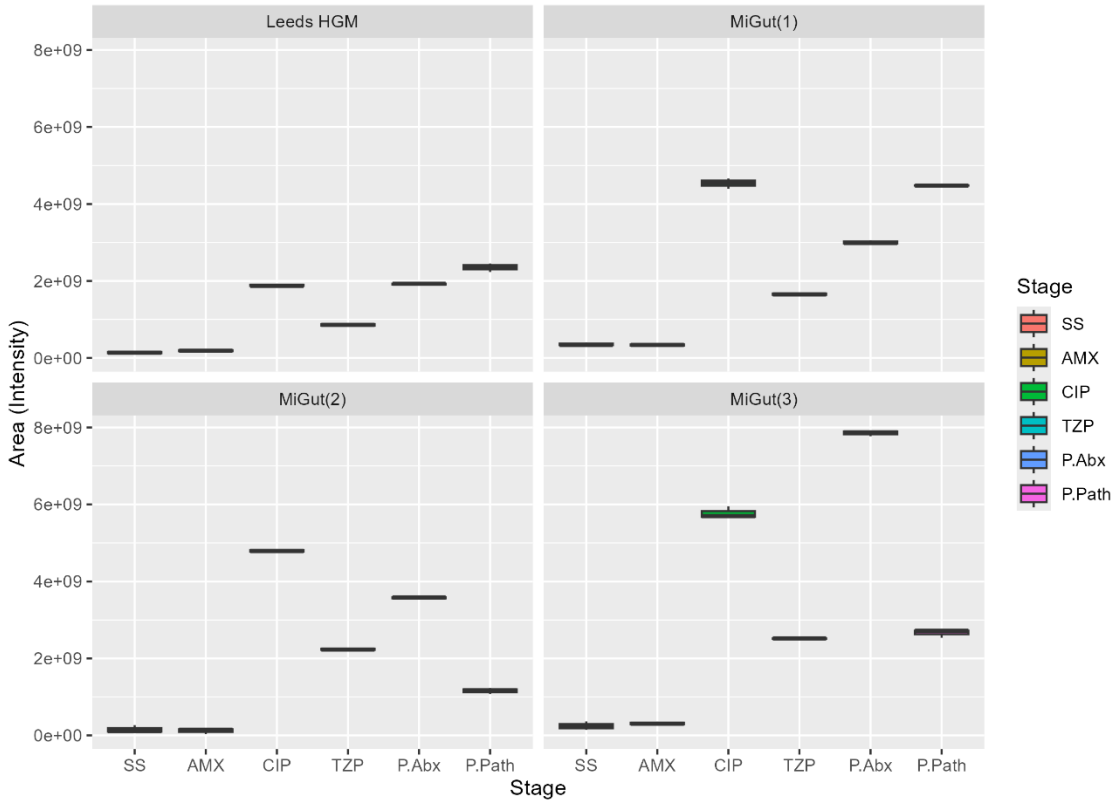

Figure 5-8 Lactate Abundance Across Experimental Stage and Model Replicates. Steady State (SS, Red), Amoxicillin (AMX, Olive), Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink) Differential abundance analysis of lactate revealed a significant increase in abundance (adjusted P-value <0.05) with a >3 Log2Fold increase observed across all model replicates at P-Abx and P-Path challenge time points.

Table 5-2 Differential Abundance of Lactate, P. Abx and P. Path vs SS

	P. Abx		P. Path	
	Log2 Fold Change	Adj. P-Value	Log_2 Fold Change	Adj. P-Value
Leeds HGM	3.79	0.007	4.12	0.008
MiGut (1)	3.14	0.043	3.73	0.045
MiGut (2)	4.98	0.016	3.24	0.041
MiGut (3)	5.15	0.0132	3.51	0.153

Blue: Increase in abundance

5.5.3.3 Butyrate

Butyrate was detected across all models and at all stages in negative ionisation mode. The compound identified as butyrate acid was characterised by an m/z of 87.04515 ([M-H]-1 ion), a calculated molecular weight of 88.05246, an annotated mass difference of 0.34 ppm and a retention time of 4.504 minutes (Figure 5-9 Extracted Ion Chromatogram (XIC) for Compound Identified as Butyrate (m/z 88.05246) In Negative Ionisation).

. The region highlighted (4.4-4.9min) corresponds to the retention time identified for butyrate acid. The red line at 4.504 indicates the apex of the integrated chromatographic peak.

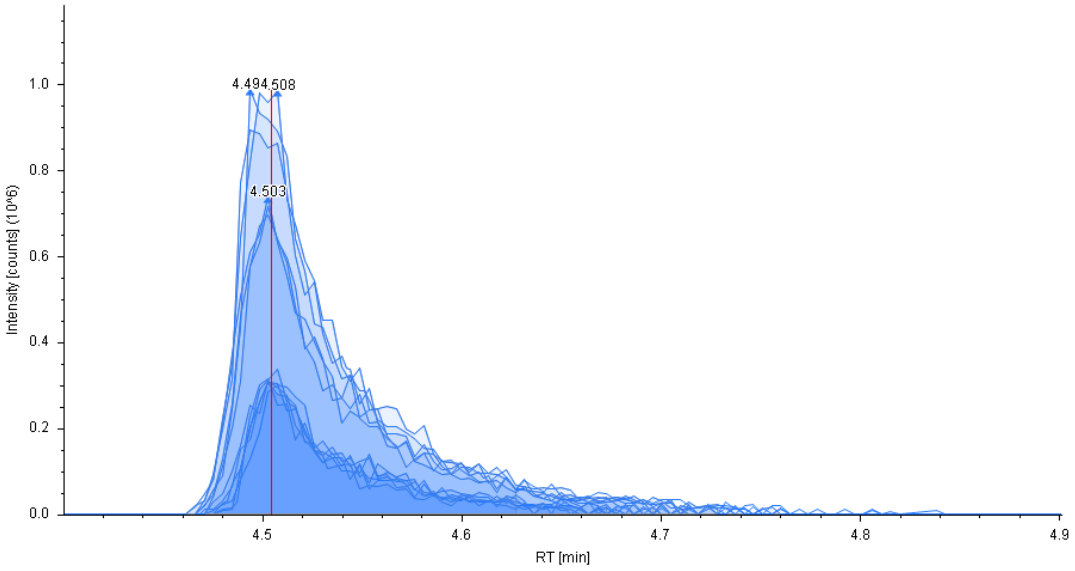

Figure 5-9 Extracted Ion Chromatogram (XIC) for Compound Identified as Butyrate (m/z 88.05246) In Negative Ionisation. The region highlighted (4.4-4.9min) corresponds to the retention time identified for butyrate acid. The red line at 4.504 indicates the apex of the integrated chromatographic peak.

In negative mode, this compound was not selected for data-dependent acquisition and consequently lacked MS^2 data. Identification was, therefore, based on calculated molecular weight and predicted molecular formula.

Additional hits for this compound were observed in negative mode, but these displayed lower signal areas. One of these lower hits was selected for DDA. However, Fragment Ion Search (FISh) analysis provided no matches against butyrate. These compound features were excluded from further analysis on the basis of lower signal strength and poor overlap with predicted fragments.

Four putative identifications for butyrate were made in positive LCMS analysis. Of the four identifications, two were selected for MS^2 DDA, corresponding to the ions $[M+DMSO+H]^+$ and $[M+ACN+N]^+$. Despite the presence of MS^2 data, there was no spectral match made for any of the compounds and annotations were made based on calculated MW and predicted formula.

Butyrate: Changes in Compound Abundance During Antibiotic Treatment and Recovery

Butyrate levels were highest during steady state and AMX dosing phases. A reduction in detected levels was observed during CIP instillation and continues through to the end of TZP dosing. Following the cessation of antibiotics, butyrate levels increase during the P-Abx period and persisted throughout the pathogen challenge phase. However, butyrate levels failed to reach those observed prior to antibiotics Figure 5-10 Butyrate Abundance Across Experimental Stage.

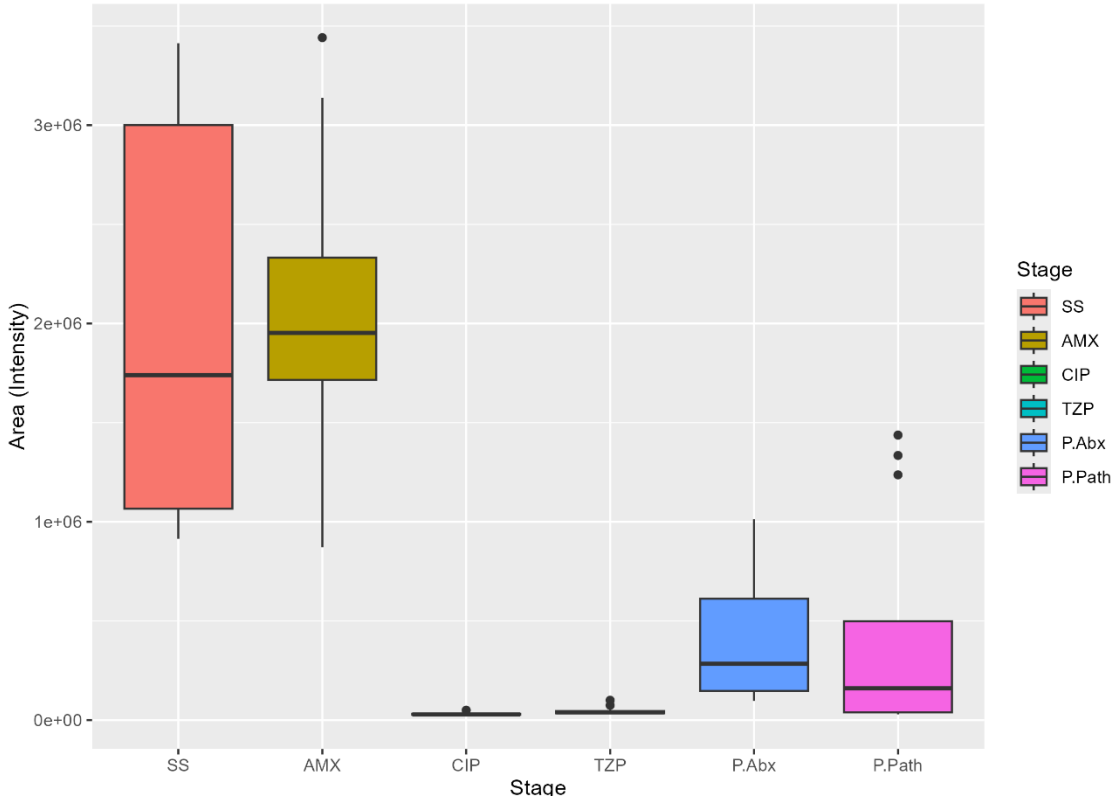

Figure 5-10 Butyrate Abundance Across Experimental Stage. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green) , Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

A reduction in levels of butyrate was observed during CIP and TZP across all models, with some recovery following the cessation of antibiotic instillation (Figure 5-11 Butyrate Abundance Across Experimental Stage and Model Replicates) This recovery was most prominent in the Leeds HGM

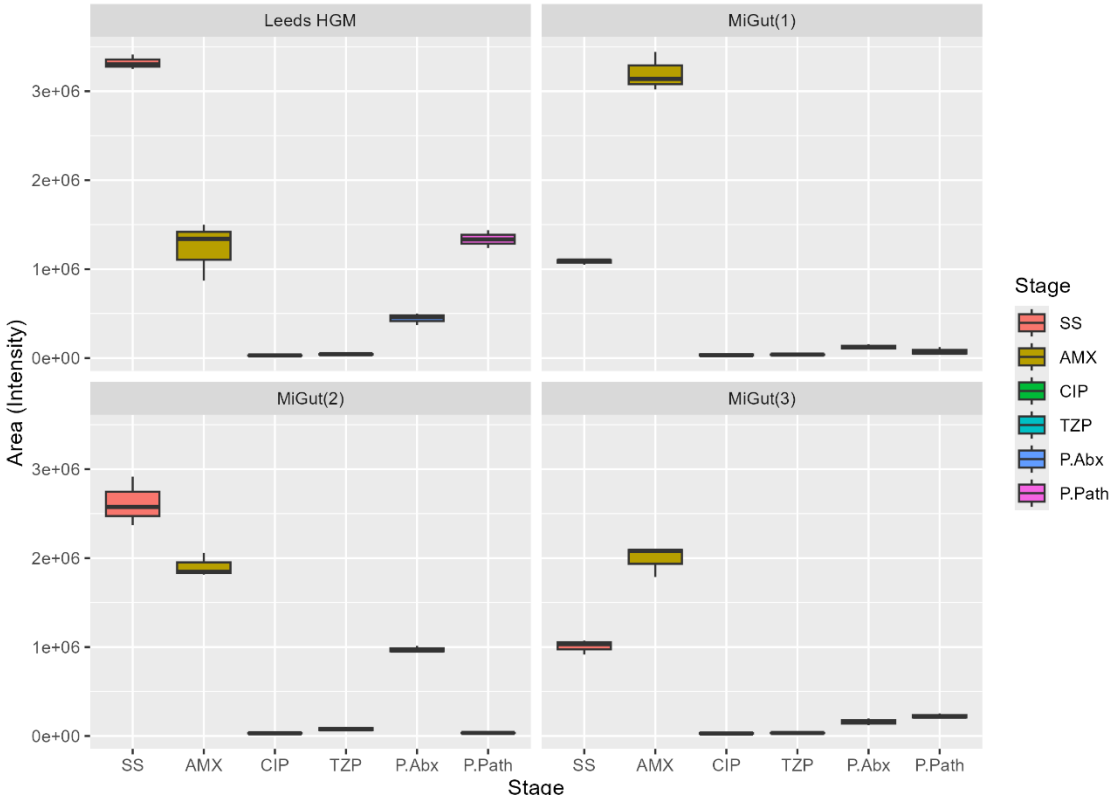

Figure 5-11 Butyrate Abundance Across Experimental Stage and Model Replicates. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Differential abundance analysis revealed statistically significant reductions in butyrate acid (adjusted P-value <0.05) across all models at the end P-Abx and following P-Path.

Table 5-3 Differential Abundance of Butyrate, P. Abx and P. Path vs SS

	P-Abx		P-Path	
	Log2 Fold Change	Adj. P-Value	Log_2 Fold Change	Adj. P-Value
Leeds HGM	-2.85	0.014	-1.31	0.018
MiGut (1)	-3.13	0.017	-4.18	0.036
MiGut (2)	-1.43	0.015	-6.36	0.003
MiGut (3)	-2.76	0.013	-2.08	0.016

Orange: Decrease in abundance

5.5.3.4 Associated Changes in Butyrate Metabolism

Due to the observed reductions in butyrate levels, compounds with the pathways for butyrate metabolism were explored to determine the presence of a metabolic bottleneck in the pathway. These compounds are part of the glutarate pathways for butyrate production. This pathway was chosen based on the association with glutaconate-CoA transferase (EC:2.8.3.12), which was identified as decreasing in abundance within the metagenome following antibiotic instillation.

5.5.3.4.1 2-Hydroxyglutarate

2-hydroxyglutarate was identified across all models and experimental phases. The compound feature was characterised by an m/z of 147.03714 ([M-H]-1 Ion), a calculated molecular weight of 148.03714, and an annotated mass difference of -0.02ppm. 2-Hydroxyglutarate eluted with a retention time of 1.532 minutes

Figure 5-12 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Hydroxyglutarate (m/z 148.03714) In Negative Ionisation

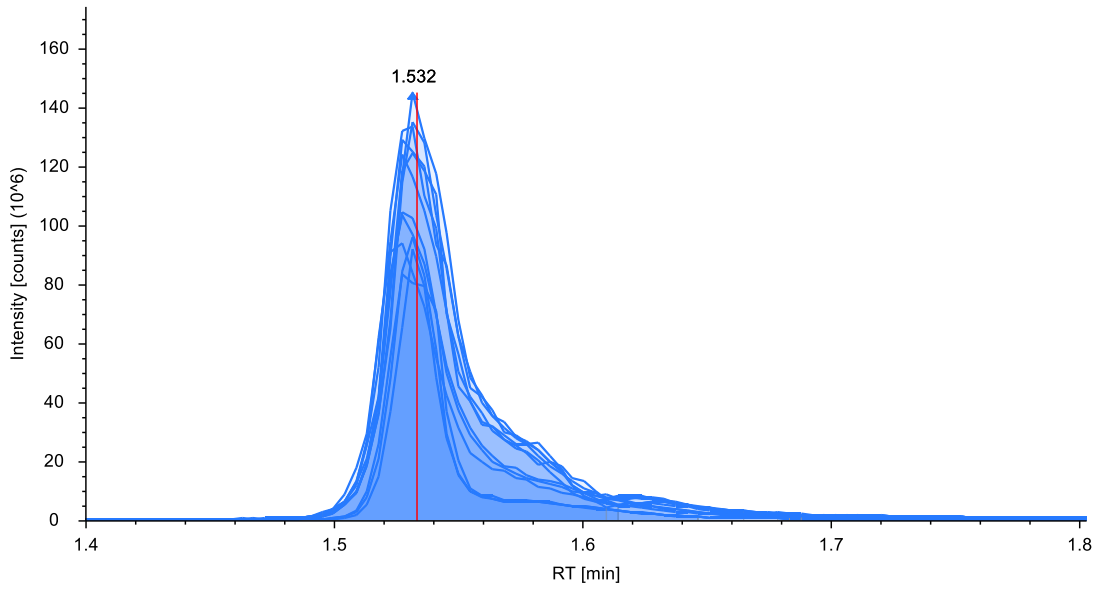

Figure 5-12 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Hydroxyglutarate (m/z 148.03714) In Negative Ionisation. The region highlighted (1.4-1.8min) corresponds to the retention time identified for 2-Hydroxyglutarate. The red line at 1.532 indicates the apex of the integrated chromatographic peak.

2-Hydroxyglutarate was selected for MS^2 analysis in the data-dependent experiment and FISh analysis reaffirmed the compound annotated with complete coverage (FISh coverage 100%) with 5 matched compound fragments (Figure 5-13 MS² Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Hydroxyglutarate

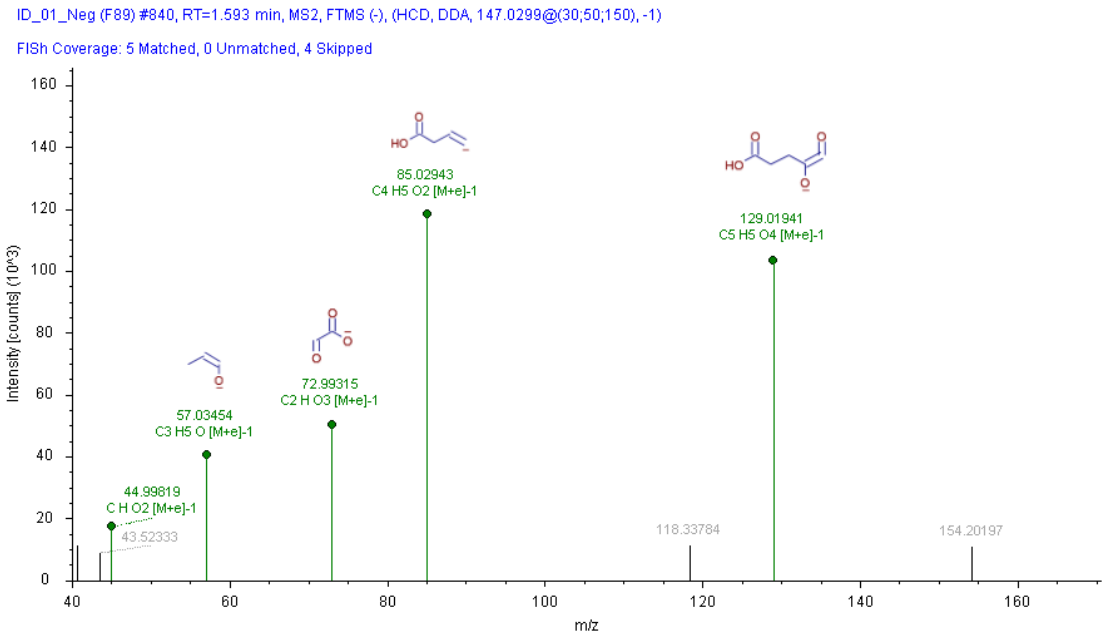

Figure 5-13 MS² Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Hydroxyglutarate. Showing matched peaks and their associated fragments

The abundance of 2-hydroxyglutarate remained stable during the steady-state baseline and AMX instillation. From CIP instillation onwards, the observed shift in abundance was inverse to that of butyrate. Elevated levels persisted through TZP dosing, P-Abx, and P-Path Figure 5-14 2-Hydroxyglutarate Abundance Across Experimental Stage.

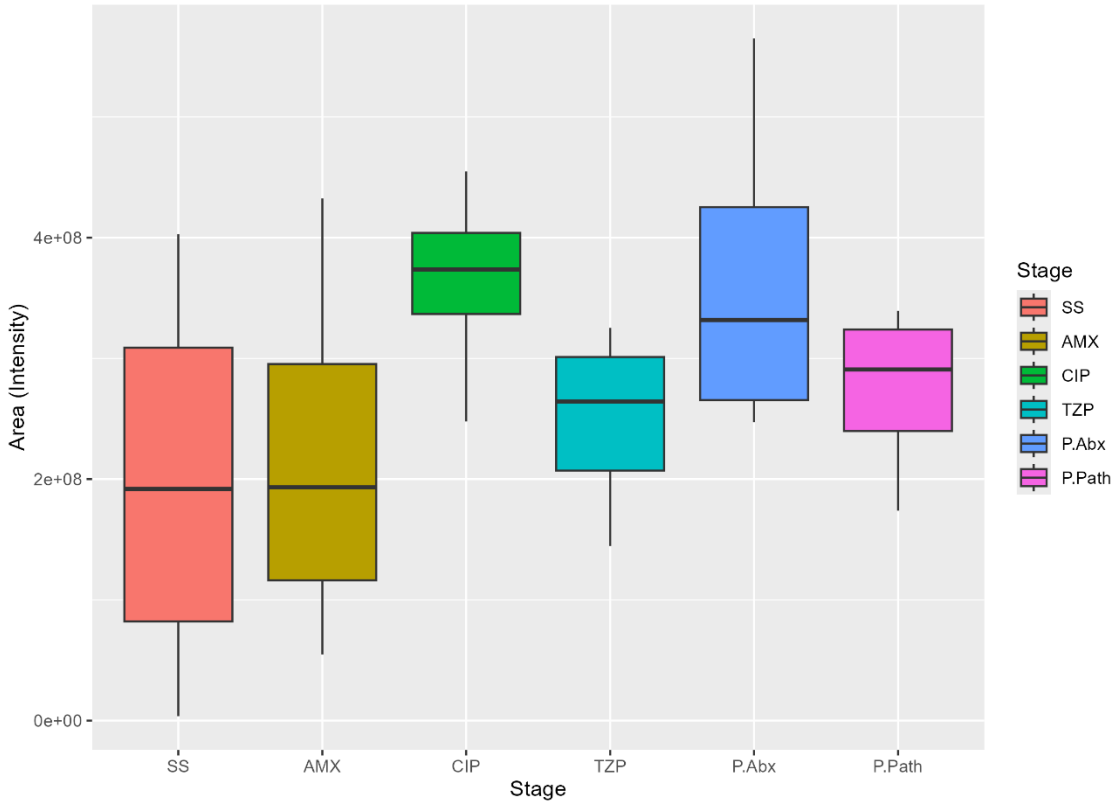

Figure 5-14 2-Hydroxyglutarate Abundance Across Experimental Stage . Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

The variance in 2-hydroxyglutarate abundance appeared to be less conserved between models compared to butyrate. Across all MiGut replicates, levels of 2-hydroxyglutarate were higher compared to the Leeds HGM. In the Leeds HGM from CIP instillation onward, there was a notable increase in the relative abundance of 2-Hydroxyglutarate. This trend was also noted in MiGut (3). However, P-Path levels decreased to those more reflective of steady state.

In MiGut replicates 1 and 2, the observations noted in Leeds HGM and MiGut (3) are not conserved. These replicates both displayed the greatest initial abundance of 2-hydroxyglutarate.

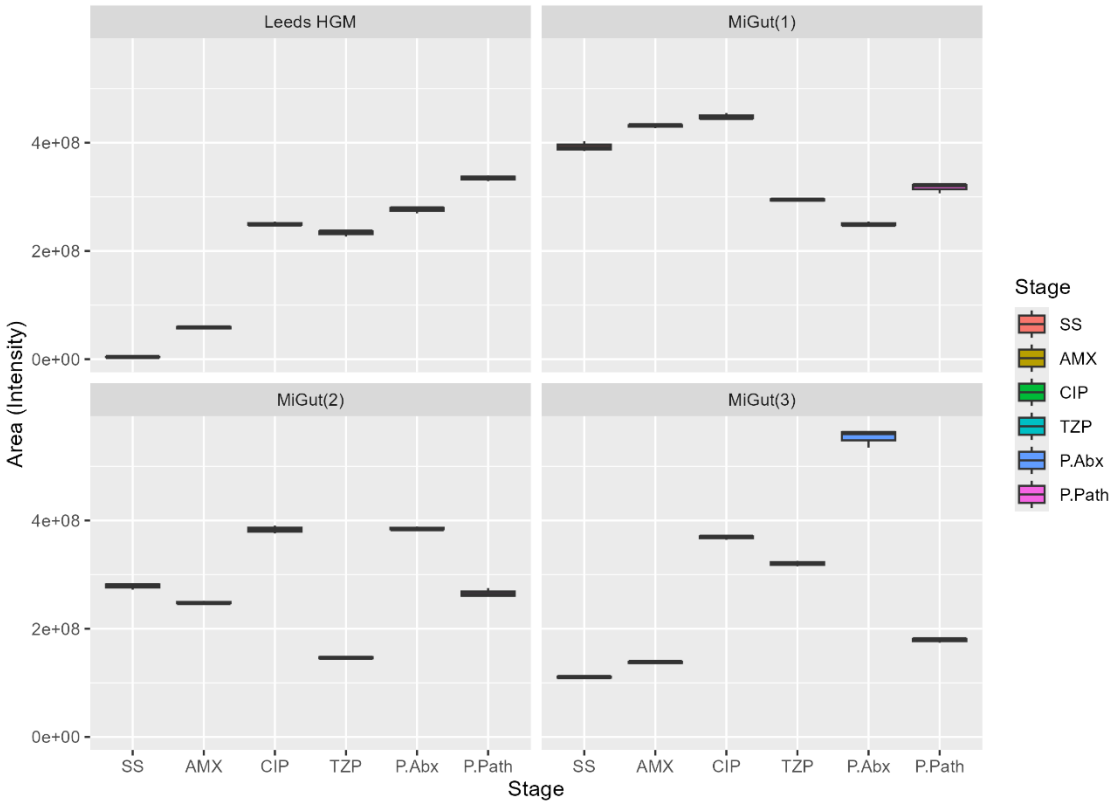

Figure 5-15 2-Hydroxyglutarate Abundance Across Experimental Stage and Model Replicates. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Differential abundance revealed significant increases in abundance (adjusted P-value <0.05) of 2-hydroxyglutarate across 3 of the 4 models following the P-Abx (Table 5-4 Differential Abundance of 2-Hydroxyglutarate, P. Abx and P. Path vs SS). The greatest increase was observed in the Leeds HGM with a Log2Fold

increase of 6.13, MiGut Replicates (2) and (3) both showed smaller increases in abundance.

P-Path, 2-hydroxyglutarate levels remained elevated relative to steady state in two of the four replicates. In the Leeds HGM, levels remained high, while in MiGut (3) levels decreased from those seen in the P-Abx sample. MiGut (2) levels trended towards those seen at steady state. MiGut (1) levels remained similar to those seen following antibiotic recovery.

Table 5-4 Differential Abundance of 2-Hydroxyglutarate, P. Abx and P. Path vs SS

	P. Abx		P. Path	
	Log2 Fold Change	Adj. P-Value	Log_2 Fold Change	Adj. P-Value
Leeds HGM	6.13	0.002	6.37	0.002
MiGut (1)	-0.64	0.009	-0.32	0.026
MiGut (2)	0.47	0.014	-0.07	0.150
MiGut (3)	2.34	0.002	0.72	0.015

Blue: Increase in abundance, Orange: Decrease in abundance, Grey: Adjusted P-Value > 0.05

5.5.3.4.2 2-Oxoglutarate (alpha-Ketoglutaric acid)

2-Oxoglutarate was identified across all model replicates and all experimental phases. The feature identified as 2-oxoglutarate was characterised by a m/z of 145.0143 ([M-H]-1 Ion), a calculated molecular weight of 146.02151. and an annotated mass delta of -0.12ppm. 2-Oxyglutarate eluted at 1.542 minutes

Figure 5-16 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Oxoglutarate (m/z 145.0143) In Negative Ionisation

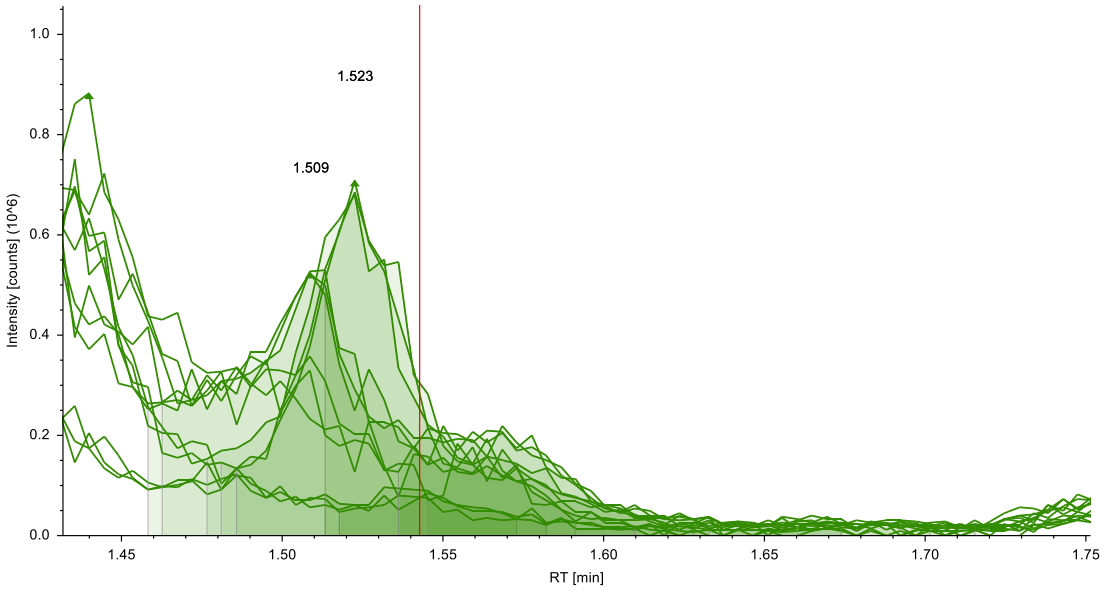

Figure 5-16 Extracted Ion Chromatogram (XIC) for Compound Identified as 2-Oxoglutarate (m/z 145.0143) In Negative Ionisation. The region highlighted (1.446- 1.75min) corresponds to the retention time identified for Lactate. The red line at 1.543min indicates the apex of the integrated chromatographic peak

2-Oxoglutarate was selected for MS^2 analysis within the data-dependent experiment scoring high in FISh analysis (FISh coverage 75%) matching 3 compound fragments (Figure 5-17 2-Oxoglutarate MS² Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Oxoglutarate.) increasing confidence in the assigned annotation.

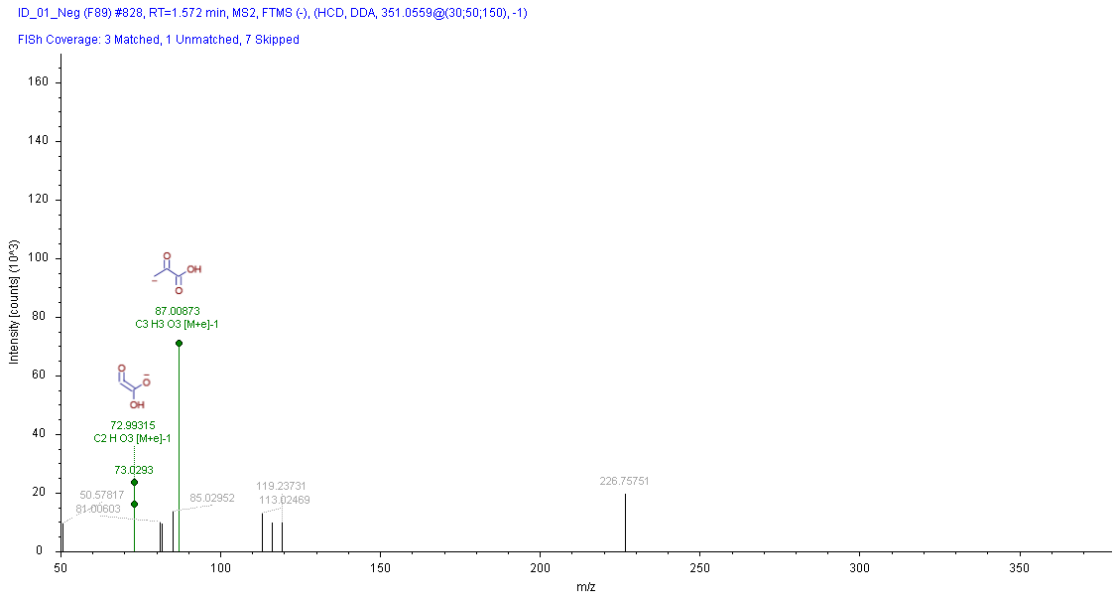

Figure 5-17 2-Oxoglutarate MS² Coverage from DDA Experiment with Matched Fragments via FISh Scoring for 2-Oxoglutarate. Showing matched peaks and their associated fragments.

The abundance of 2-oxoglutarate increased during CIP instillation and remained elevated for the remainder of the experiment. Significant variance was observed at the TZP and P-Path time points, while levels appeared more stable at the SS, AMX, and P-Abx (Figure 5-18 2-Oxoglutarate Abundance Across Experimental Stage).

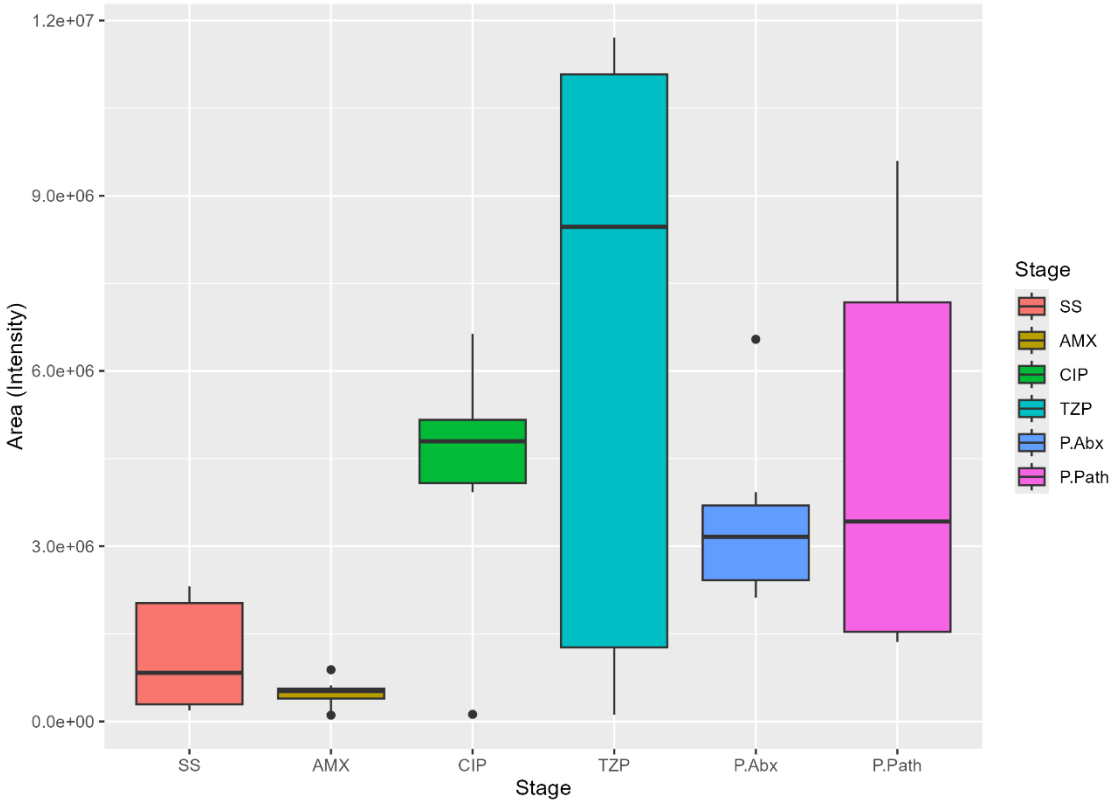

Figure 5-18 2-Oxoglutarate Abundance Across Experimental Stage. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Some inter-model variation in 2-oxoglutarate abundance was observed across gut model replicates (Figure 5-19 2-Oxoglutarate Abundance Across Experimental Stage and Model Replicates). In the Leeds HGM and MiGut (1), 2-oxoglutarate abundance increased following CIP instillation and remained at an elevated level for the duration of the experiment. MiGut (2) followed a similar trend with a temporary drop in 2-oxoglutarate abundance during TZP. MiGut (3) showed an initial increase in abundance during CIP, further increasing during TZP, but returned to levels comparable to steady state following the cessation of antibiotic instillation and remained stable for the remainder of the experiment.

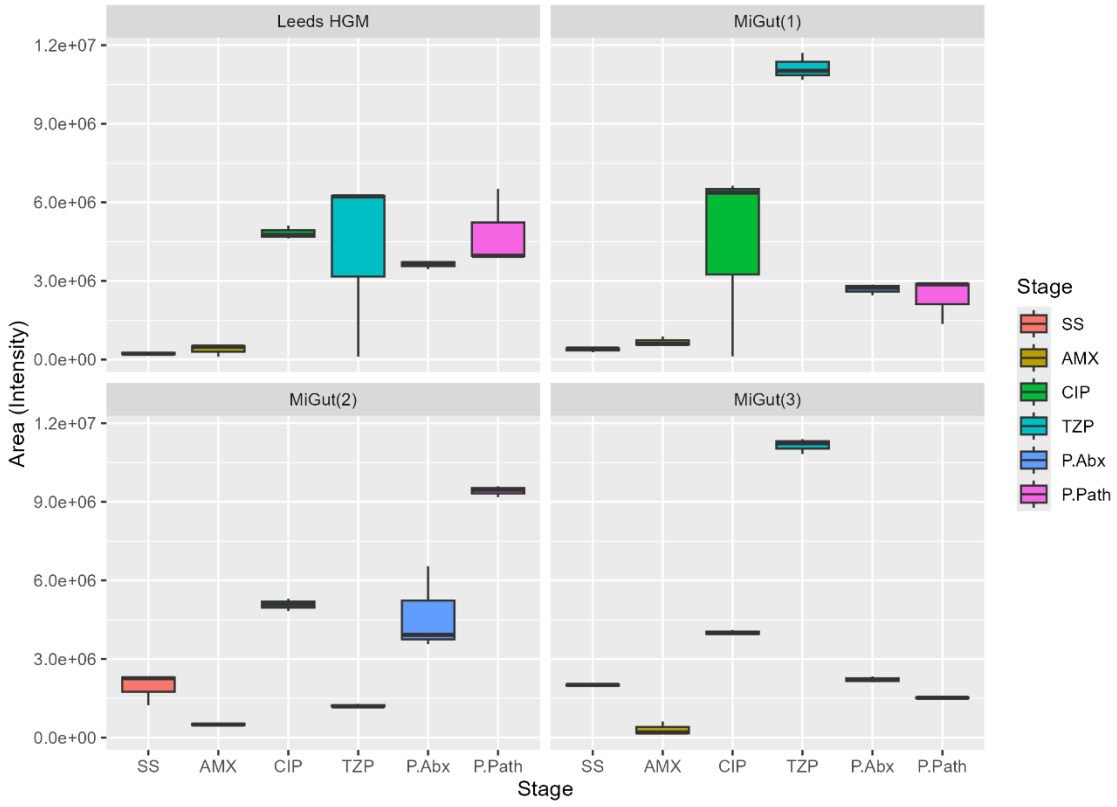

Figure 5-19 2-Oxoglutarate Abundance Across Experimental Stage and Model Replicates. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green) , Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

Differential abundance analysis showed that P-Abx 2-oxoglutarate levels were elevate across all models. Three models exhibited Log_2 fold changes greater than 1, with statistical significance (adjusted P-value <0.05) achieved in two models: Leeds HGM and MiGut (1). Following P-Path, 2-oxoglutarate levels were elevated across three models: Leeds HGM, MiGut (1) and MiGut (2), with Log2 fold changes exceeding 2 in all instances, each achieving statistical significance (adjusted P-value <0.05).

Table 5-5 Differential Abundance of 2-Oxoglutarate, P-Abx and P-Path vs SS

	P. Abx		P. Path	
	Log2 Fold Change	Adj. P-Value	Log_2 Fold Change	Adj. P-Value
Leeds HGM	3.95	0.014	4.41	0.006
MiGut (1)	2.72	0.023	2.74	0.016
MiGut (2)	1.5	0.089	2.03	0.037
MiGut (3)	0.17	0.159	-0.42	0.018

Blue: Increase in abundance, Orange: Decrease in abundance, Grey: Adjusted P-Value > 0.05

5.5.3.5 Valerate

Although not highlighted as a compound of potential interest through the differential abundance analysis of KEGG pathways in Chapter 4, valerate was investigated due to an interest in the implication of antibiotic instillation on SCFAs abundance.

A compound feature identified as valerate was detected across all sample points and model replicates. It was characterised by a m/z of 249.13569 ([2M+FA-H]-1 Ion), a calculated molecular weight of 102.06839, an annotated mass delta of 3.05ppm and eluted at 3.617minutes (Figure 5-20 Extracted Ion Chromatogram (XIC) for Compound Identified as valerate (m/z 249.13569) In Negative Ionisation)

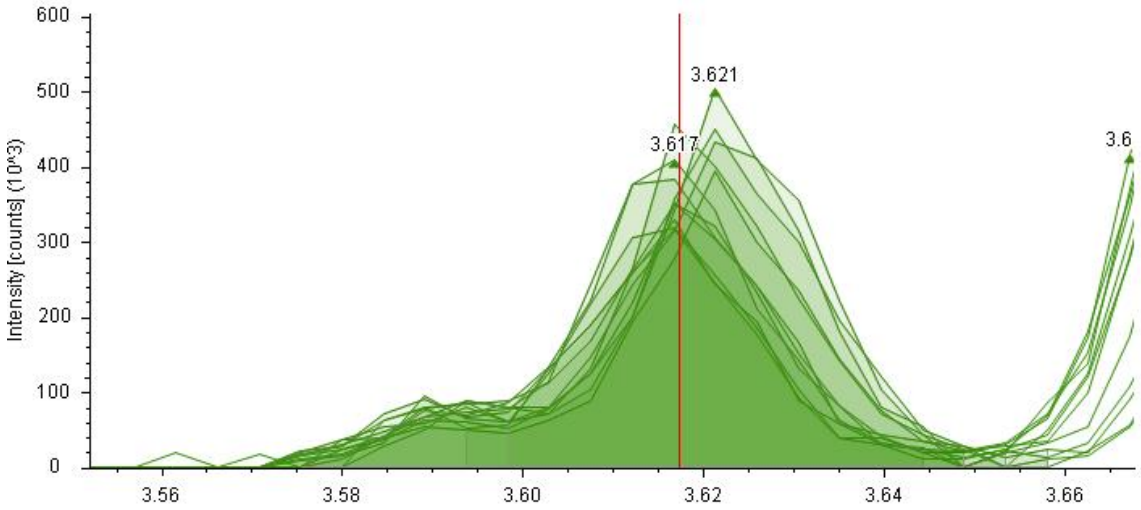

Figure 5-20 Extracted Ion Chromatogram (XIC) for Compound Identified as valerate (m/z 249.13569) In Negative Ionisation. The region highlighted (3.56 - 3.66min) corresponds to the retention time identified for valerate. The red line at 1.543min indicates the apex of the integrated chromatographic peak.

Valerate was selected for fragmentation in the DDA experiment. FISh analysis yielded a coverage score of 28.75, successfully matching 5 compound fragments (Figure 5-21 Valerate MS^2 Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Valerate. Mass accuracy and FISh coverage for valerate were lower compared to other compounds identified in

this chapter, as a result annotation confidence for valerate is lower relative to the compounds discussed previously.

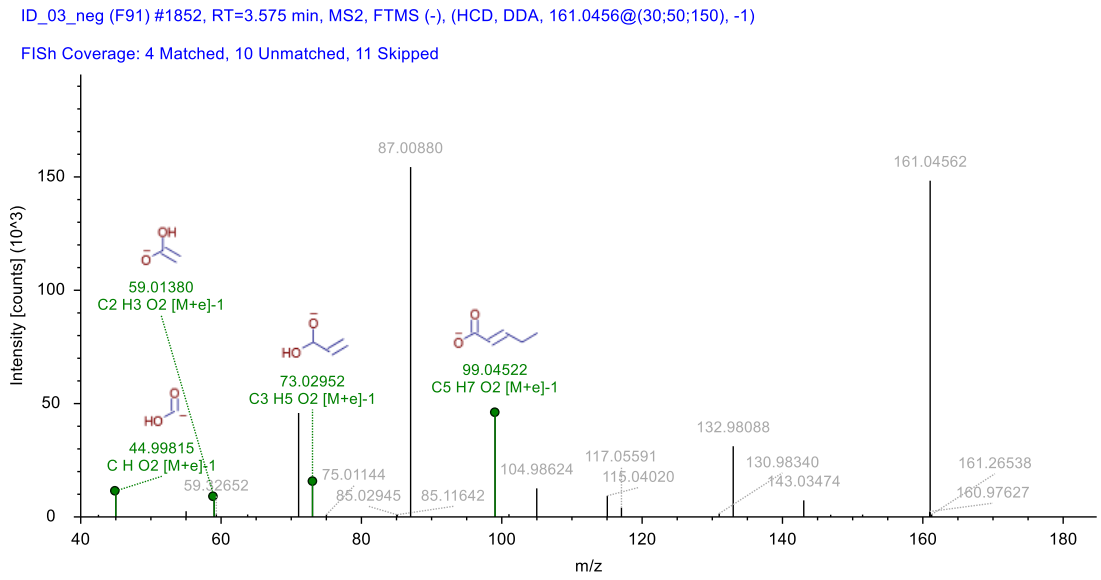

Figure 5-21 Valerate MS² Coverage from DDA Experiment with Matched Fragments via FISh Scoring for Valerate. Showing matched peaks and their associated fragments.

Valerate levels were highest during steady state but following AMX dosing levels, they declined and remained low for the remainder of the experiment Figure 5-22 Valerate Abundance Across Experimental Stage.

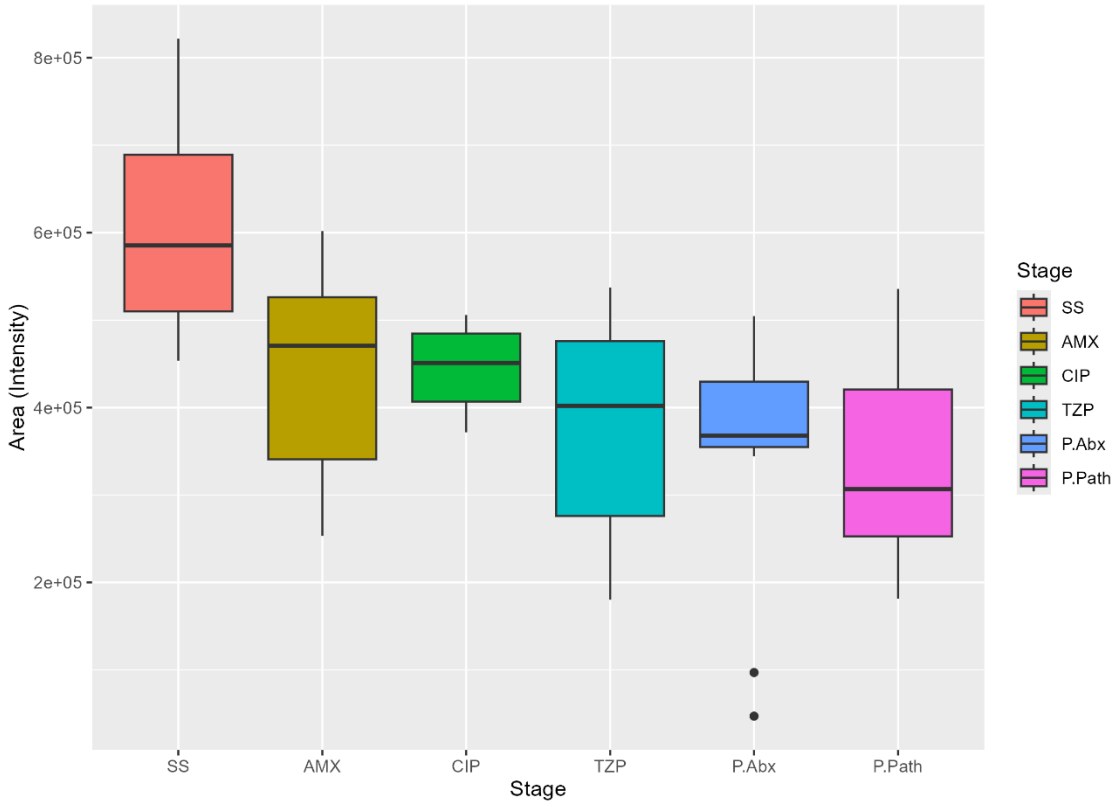

Figure 5-22 Valerate Abundance Across Experimental Stage. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green), Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink)

The trends in valerate levels were observed across the majority of models (presented in Figure 5-23 Valerate Abundance Across Experimental Stage and Model Replicates). Leeds HGM MiGut (1) and MiGut (2) exhibited consistent reductions in valerate abundance following antibiotic instillation, with levels remaining suppressed for the remainder of the experiment. The disruption of valerate was less pronounced in MiGut (3).

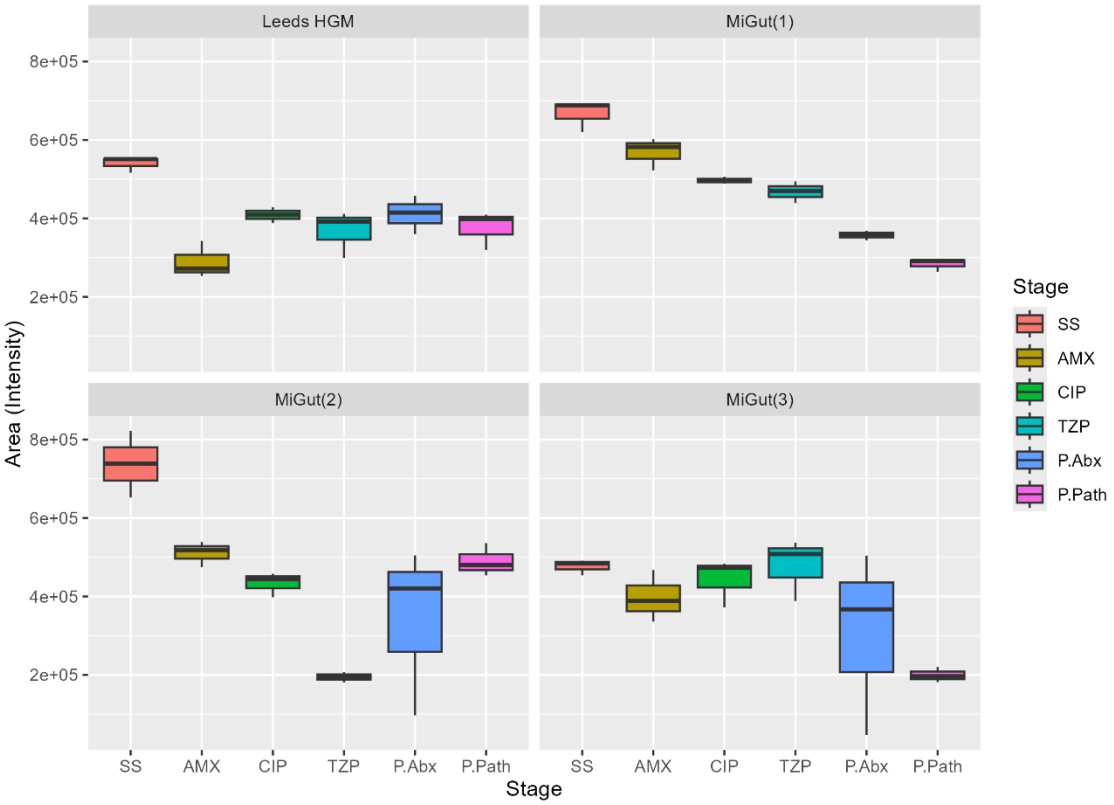

Figure 5-23 Valerate Abundance Across Experimental Stage and Model Replicates. Steady State (SS, Red) Amoxicillin (AMX, Olive) Ciprofloxacin (CIP, Green) , Piperacillin/tazobactam (TZP, Teal), Post Antibiotic (P-Abx, Blue) Post Pathogen (P-Path, Pink). Differential abundance analysis for valerate revealed a universal decrease in abundance across all models at both P-Abx and P-Path time points. While decreases in abundance were observed in all samples, Log2fold changes were modest compared to the other compounds discussed in this chapter and remained below 1.5 (

Table 5-6 Differential Abundance of Valerate, P. Abx and P. Path vs **SS**). Among the replicates, MiGut (1) experiences the greatest reductions in valerate abundance, with Log_2 fold Changes of -0.91 and -1.23 at P-Abx and P-Path time points, respectively.

Table 5-6 Differential Abundance of Valerate, P. Abx and P. Path vs SS

	P. Abx		P. Path	
	Log2 Fold Change	Adj. P-Value	Log2 Fold Change	Adj. P-Value
Leeds HGM	-0.41	0.162	-0.47	0.058
MiGut (1)	-0.91	0.018	-1.23	0.003
MiGut (2)	-0.81	0.277	-0.52	0.049
MiGut (3)	-0.4	0.447	-1.32	0.017

Orange: Decrease in abundance, Grey: Adjusted P-Value > 0.05

5.6 Discussion

This chapter set out to explore the phenotypic resolution of antibiotic induced disturbance to the *in-vitro* gut microbiome. Whilst culture-based analysis was effective at tracking the dynamics of key bacterial taxa in response to antimicrobial instillation, it failed to fully capture the loss of colonisation resistance evidence by colonisation and proliferation of *C. difficile* and *K. pneumoniae.*

Metagenomic analysis provided a more comprehensive analysis of the *in-vitro* microbiome, identifying perturbations to key bacterial genera, such as *Faecalibacterium* and *Coprococcus*. These observations, combined with functional analysis of KEGG pathways, revealed a narrative of potential long-term disruptions to the metabolic capacity of the microbiome in the production of SCFAs; identifying acetate, propionate, and butyrate pathways as being significantly affected. Based on these findings, a metabolomic analysis was conducted to observe how shifts in the metabolic capacity were resolved.

LCMS analysis successfully recovered a substantial number of compounds across both LCMS ionisation modes. However, with datasets of this size, challenges may arise in uncovering a focused biological narrative.

In the dataset, only a relatively small percentage of compounds received named annotations, 14.39% in negative mode and 12.73% in positive mode. This leaves approximately 85% of the data without named annotations. During data analysis, compounds were initially filtered and identified by predicted formula, which covers the majority of the data set, and of the compounds of interest recovered all of them received named annotations to provide an additional level of confidence. Even without names, putative identifications can be constructed for compounds based on their molecular formula. However, this requires significant manual data curation and is not feasible across the entire data set. Consequently, top-level analysis, looking for biological trends related to compound classes may be missing significant observations. Within this chapter, being directed by metagenomic analysis, it was possible to perform this manual interrogation on compounds within the identified SCFA pathways. This was also a significant factor in deciding to primarily focus on a subset of the data

To streamline the analysis in this chapter, the outcomes of the previous metagenomic investigation were used to direct the focus and discussion throughout.

In some instances, one of the injection replicates does not align with the other two. This discrepancy is particularly noticeable in box and whisker plots, which display changes in compound abundance by model. Upon review of the data, these inconsistencies appear to be attributable to single outlying injections. Given that the remaining two injections typically align closely, in both the median and quartile range, it is likely that this variation is an error introduced by a single injection, rather than a wider issue with sample preparation or machine accuracy. This could be attributable to insufficient sample volume within the HPLC vial, leading to an incomplete draw of intended 5µL sample volume. This is reinforced as extended whiskers typically extend below the expected range. Although the variation is significant, non-parametric analysis utilising median values has been applied throughout this chapter. The use of median values helps to mitigate the impact of outlying injections on overall data analysis (Reinhold et al., 2019).

5.6.1 Variations to the Microbiome

Antibiotics induced detectable changes in the metabolome in both positive and negative ionisation modes. However, the negative mode appeared more effective at resolving the persistent functional alterations that remained after the P-Abx period.

5.6.1.1 Negative NMDS analysis

In NMDS analysis, metabolomic data from the negative mode LCMS experiment revealed a potential dichotomy between two groups of sample points.

Group 1 comprised samples potentially representative of intact homeostatic balance: steady state samples collected prior to any antibiotic instillation, and AMX time points taken following a single course of antibiotics.

Associating AMX with a homeostatic microbiome is complex; as previous analysis demonstrates, some disruption to the microbiome did occur. Culture analysis revealed that AMX caused disruption to the genera *Lactobacillus*, *Bifidobacteria* and *Enterococcus* within the distal region (vessel 3) of the Leeds HGM, although total anaerobes and total facultative anaerobe populations remained high. Metagenomic analysis showed that AMX was associated with a decrease in Shannon diversity, an increase in the proportion of Bacteroides and Hungatella, and a decrease in the proportion of minor genera. There were also reductions in species diversity within *Bifidobacterium*, *Lactobacillus*, *Oscillibacter* and *Faecalibacterium,* identified as a key producer of SCFAs (Duncan et al., 2002; Vital et al., 2014)

However, it remains unclear whether the disruption caused by AMX, or a single antibiotic is sufficient to classify a microbiome as dysbiotic or whether the disruptions described above simply leave the microbiome increasingly vulnerable to future perturbations

Previous studies utilising the Leeds HGM demonstrated that a 7-day, 3-time co-amoxiclav (AMC) regime was sufficient to induce CDI alone (Chilton, Caroline H. et al., 2012). However, these findings were under different experimental conditions. The study by Chilton *et al.* explored the potential for AMC to induce CDI in an instance of pre-existing *C. difficile* colonisation, whereas in this study, *C. difficile* was not inoculated into the model until day 55. Lekang et al. (Lekang

et al., 2022) describe the impacts of AMX on the murine microbiome following a 7 or 14-day course of antibiotic treatment. They found that mice in both groups experienced losses in diversity following 5 days of treatment (56% in 7 days, 40% in 14 days). The loss of diversity did not appear to be correlated with treatment length; identifying no significant diversity between days 5,9 and 13 for 14 day treated mice suggested that the 5 day dosing period in the present study would likely be sufficient to observe peak AMX-induced disruption. The length of recovery was 6 days within 7 day treatment group and 7 days in the 14 day treatment group, indicating that CIP instillation in the present study will compound the impacts of AMX. Taxa which decreased in abundance reflect changes in the abundance of genera identified in the present study, with lekang *et al.* noting changes in *Lachnospiraceae* (Present study *Hungatella*) , *Oscillospiraceae* (present study *Oscillibacter*) and *Ruminococcaceae* (Present study Faecalibacterium) Elvers et al (2020) systematic review, describes in consistent observations between studies for AMX regards observations of *Lactobacillus, Bifidobacterium* and *Bacteroides.* Typically, they were described as decreasing in abundance as observed within the present study, however there were notable instance where they increased was also noted (Christensson et al., 1991). Elvers *et al.*, noted a lack of research and understanding into commonly prescribed antibiotic on the microbiome, highlighting the value of in-vitro studies exploring antibiotic-induced dysbiosis.

Slimings and Riley (2014) in a systematic review, describe penicillin combinations, such as AMC as increasing the risks of CDI (Odds ratio 1.4). Previous exposure to penicillin/clavulanate combinations was identified as those being the most likely to induce CDI within this group (Odds ratio 4.6). However, the window defined for previous antibiotic exposure within the analysis varies significantly (28 days – 3 months), which may impact findings. Miller *et al*, (2023) also highlighted the difference between AMC and AMX in their potential to induce CDI, with AMC being significantly higher (Adjusted OR 8.53 95% CI 8.23 – 8.85) than AMX (Adjusted OR 1.96, 1.88 – 2.04). With these observations in mind, the disruption and potential for AMC to illicit CDI, as demonstrated by Chilton *et al.*, may not be representative of the extent to which AMX disrupted the microbiome in the present study, where AMX may truly align more closely with steady state sample as described in the NMDS analysis.

Group 2 comprises samples taken following the instillation of 2 or more antibiotics (CIP, TZP, P-Abx, and P-Path). Due to the compounding effects of antibiotic exposure, it is possible that these samples represent a spectrum of dysbiotic microbiomes. Cumulative dose, number of antibiotics and days of exposure have all been demonstrated to be major risk factors for CDI (Stevens, Vanessa et al., 2011) (Dysangco et al., 2022).

Group 2 is notable for three potential sub-clusters, primarily consisting of samples from CIP, TZP and post-antibiotic phases (P-Abx and P-Path). Within group 2, there are two time points of particular note, both relating to the same MiGut (2) replicate for the P-Path and TZP time points. As mentioned previously, these time points deviate from the pattern exhibited by other TZP and P-Path sample points. Whilst it is not possible to prove, there is a possibility that the data for these samples are reversed. This may have occurred at multiple stages due to the complexity of the analysis pipeline, including sample collection extraction, LCMS analysis or data processing. To confirm this, additional experiments would be necessary to verify within sample point similarity.

Within negative analysis, the shift of the microbiome from homeostatic to potentially dysbiotic occurs during CIP instillation. Previously CIP has been shown to cause significant disruption within the *in-vitro* microbiome to induce CDI in the same model system (Saxton et al., 2009). Metagenomic analysis of vessel 3, revealed that CIP was associated with the loss of *Coprococcus* which has been associated with SCFA production (Singhal et al., 2021; Frolova et al., 2022).

TZP samples analysis also cluster within Group 2 (with the exception of the possible MiGut(2) sample discussed above), indicating further distinct shifts associated with cumulative antibiotics. Previous Leeds HGM studies have explored the potential for TZP to induce CDI (Baines, S. D. et al., 2005). Baines *et al.*, documented depletion of all enumerated populations below limits of detection for the entirety of the instillation period. Although TZP caused profound disruption to the microbiome in this instance, it was not associated with the toxin titres indicative of CDI within the two-week monitoring period. It is possible that a two-week monitoring period is not sufficient to entirely rule out CDI, with current model protocols requiring a 4-week monitoring period.

Venturini et al. (2021) document the potential for TZP to disrupt colonisation resistance in mice following a 5-day course of antibiotics. TZP-induced disruption, associated with reductions in *Bacteroidetes* and an increase in *Lactobacillaes*, *Proteobacteria* and notably *Enterobacteriaceae,* was shown to facilitate the temporary colonisation by multidrug resistant *E. coli* and *K. pneumonia* for up to 4 weeks.

While these previous studies have documented the potential for TZP to impact colonisation resistance both *in vitro* and *in vivo*, the present study builds on these observations identifying the potential for TZP as a part of an antibiotic dosing regimen involving the sequential administration of 3 distinct antibiotic to contribute to persistent functional dysbiosis, rendering the microbiome more permissive to colonisation by exogenous pathogens.

Interestingly, negative NMDS associated P-Abx and P-Path samples within potentially dysbiotic group more closely than metagenomic genus level NMDS analysis, suggesting that metabolomics may provide more specific markers for dysbiosis, further highlighting the limitations of defining and identifying dysbiosis taxonomic markers only.

5.6.1.2 Positive NMDS analysis

While showing sharing some similarities with negative mode NMDS, clustering steady state samples with AMX counterparts, positive mode NMDS analysis aligned the time points for P-Abx and P-Path more closely with the previously identified homeostatic samples. PERMANOVA analysis identified negative mode as being better able to account for time-point associated metabolome changes identified by LCMS in this study, with experimental stage accounting for 54.48% variance in negative LCMS and 35.91% in positive LCMS.

As discussed within the method development chapter, the analytical workflow developed by the C18 LCMS pipelined may be limited in its scope to recover certain compounds such as Sterols and fatty acyls. Compound classes such as carboxylic acids and derivatives are particularly well resolved using this column chemistry. This increased capacity to detect specific chemical properties, such as their propensity to deprotonate under certain pH conditions, may lead to some features of dysbiosis being more prominent in analysis. Acidic compounds tend to deprotonate, forming negative ions, and it is this

characteristic which may render them more prominent in negative ionisation mode (Cech and Enke, 2001; Malm et al., 2021).

5.6.2 Putative biomarkers of Antibiotic Induce dysbiosis

5.6.2.1 Butyrate

Butyrate abundance was investigated due to the identification of differentially abundant genes associated with butyrate metabolism in Chapter 4. These shifts in abundance reached statistical significance for key steps within the glutarate and 4-aminobutyrate/succinate pathways and within wider butyrate metabolism. There were also notable trends towards reductions in the abundance of three out of four of the genes identified associated with acetyl-CoA metabolism. These reductions, along with decrease in key butyrate-producing genera, *Fusobacteria* and *Oscillibacter,* suggest a potential reduction in the microbiomes capacity for butyrate production.

Within this study, a reduction in butyrate appeared to be a putative marker of antibiotic-induced dysbiosis using different model systems and across both post-antibiotic instillation time points. With reductions being conserved across all model replicates.

The reductions in butyrate abundance here are supported by published literature. Reductions in SCFA associated with high and low CIP dosing regimens in mice were described by Zhu et al. (Zhu et al., 2020). Both high and low CIP dosed mice initially showed decreases in the abundance of butyrate. Low dose CIP was associated with faster and more complete recoveries in butyrate following antibiotic cessation. Liu, X. et al. (2023) documented reductions in butyrate associated with depletion of microbiota species in mice following a 10 day course of antibiotics. However, the antibiotic treatment documented (ampicillin, vancomycin, neomycin sulphate and metronidazole all at 10mg) is extreme and unlikely to be representative of clinical dosing regimens.

Within the present study and wider literature, butyrate has been discussed as a potential marker for a range of conditions associated with the intestinal microbiome.

Butyrate has been previously discussed as a potential biomarker for CDI in a study by Bosnjak et al. (2023) who recorded increases in butyrate in patients colonised with *C. difficile*. The study compared 6 patient cohorts, with and without antibiotic-associated diarrhoea and *C. difficile* colonisation. Bonsiak *et al.* attributes the increase in butyrate potentially to *C. difficile* ‘s contribution to SCFA via fermentative pathways as an alternative to Stickland reactions (Hofmann et al., 2018). While butyrate abundance was increased in the presence of *C. difficile* colonisation, the practical application of butyrate as a marker of CDI is problematic when assays for specific markers, such as *C. difficile* toxin exist. Interestingly, within the study, patients with antibiotic-associated diarrhoea exhibited lower butyrate levels than the healthy control group, which may indicate a butyrate reduction as being associated with antibiotic-associated diarrhoea rather than an increase being an effective marker for CDI.

Interestingly butyrate has also been shown to have a protective effect against CDI *in vitro.* Pensinger et al. (2023) demonstrated that increased butyrate levels lead to increased lag phase and decreased maximum growth rates observed across 10 PCR ribotypes.

As described by Hayashi, A. et al. (2021), *Clostridium butyricum* MIYAIRI588 was shown to be preventative against CDI in mice, through the recruitment of neutrophil and t cell immunity activation. Neutrophil recruitment was suggested to be likely attributable to butyrate produced by *C. butyricum.* This protective mechanism while of note, is note at play within the current study due to the lack of host cells within the model. Instead, the findings highlight the potential for microbiota-specific mechanisms observed within the *in vitro* models presented here to be amplified *in vivo.*

A reduction in butyrate has also been discussed as a potential biomarker of colorectal cancer, along with acetate and propanoate (Adewiah et al., 2018).

McOrist et al. (2011) described high variability in butyrate concentration in participants free from recent antibiotic usage and bowel disorders, ranging from 3.5-32.6 mmol/kg in faecal samples.

The overall specificity of butyrate as a specific marker for antibiotic-associated dysbiosis may be limited, rather serving as an overall biomarker of gut microbiota health.

The lack of MS^2 information from the DDA experiment may be accountable to the selection of an earlier potentially isometric compound. FISh analysis on this earlier compound was performed for all proposed structures and failed to identify any matching fragments. Due to supporting observations from the literature, it was decided that the annotation of butyrate is highly liked for the identified compound feature. Rather than relying on data-dependent acquisition in the future for all compound IDs, it may be more useful to carry out semi-targeted analysis with a defined list of compounds of interest to ensure full MS^2 coverage.

5.6.2.1.1 Butyrate Associated markers: 2-Hydroxyglutarate and 2-Oxoglutarate

2-hydroxyglutarate, is commonly discussed and reviewed for its association with cancers (Dang et al., 2009; Foskolou et al., 2023) due to its accumulation within cancer cells due to mutations in isocitrate dehydrogenase. Here, it was investigated as a potential marker of antibiotic-induced dysbiosis as a specific marker of the glutarate pathway for butyrate production, due to the potential for disruption associated with the observed reduction in glutaconate CoA-transferase [EC:2.8.3.12] through metagenomic analysis. Within the Leeds HGM, 2-hydroxy glutarate appeared to be an effective marker for antibiotic-induced dysbiosis following both the P-Abx phase and P-Path. These trends were less conserved between the MiGut replicates than those observed with butyrate. Metagenomic analysis was not performed across all MiGut replicates due to financial limitations, so it is not possible to be sure that alterations observed via metagenomic analysis on Leeds HGM and MiGut (3) were the same across all replicates. Interestingly, 2-hydroxyglutarate saw similar increases across both replicates used for metagenomic analysis.

2-Oxoglutarate was also evaluated for its potential association with the glutarate pathway. 2-Oxoglutrate appeared to be elevated across most model replicates, following antibiotic instillation recovery, although it failed to reach significance in two replicates MiGut (2) and MiGut (3). Following P-Path addition 2-oxoglutarate was detected at elevated levels across 3 of the 4 replicates

Leeds HGM, MiGut (1), MiGut (2) (with a minor decrease in abundance in one, MiGut (3)) achieving statistical significance across all. While it demonstrated potential within the present study, its prevalence across a number of metabolic pathways as outlined in KEGG (Kanehisa and Goto, 2000) suggest it may be more suitable as a marker for pathway disruption within a comprehensive marker panel, rather than as a sole diagnostic marker for dysbiosis.

5.6.2.1.2 Lactate

Lactate abundance in the model was explored due to the distribution of lactate-associated pathways within acetate and propanoate metabolites. As a potential proxy marker for acetate and propanoate metabolism, lactate saw statistically significant increases in abundance across all model replicates following both the P-Abx and P-Path. Lactate is a common human metabolite in its L-lactate form; however, its D-lactate form is commonly attributed to the microbiome (Talasniemi et al., 2008). If the rise in lactate observed following antibiotic instillation is attributable to its D-lactate form, then its accumulation may be a useful microbiota-specific marker of dysbiosis. With the analytical pipeline utilised in this chapter L- and D- chirality are not resolvable by MS^2 analysis, so a more specific experimental analysis may be needed in the future.

Accumulation of lactate may be exploited by pathogens, including *C.* difficile, as discussed in a review of Microbial lactate utilisation by Louis et al. (2022), this may potentially provide a competitive advantage for *C. difficile* proliferation within the gut following antimicrobial instillation.

In addition to the potential alterations of lactate pathways of acetate and propanoate metabolism, depletion in commensal Clostridia in murine models has been associated with increased host lactate, supporting *Salmonella* (Gillis et al., 2018)

If the increase in lactate is attributable to a reduction in downstream processes, such as cross-feeding leading to acetate or propionate production, perturbations to SCFA pathways may have additional impacts, such as those described above.

5.6.2.1.3 Valerate

Although alterations to valerate pathways were not identified in the metagenomic analysis Chapter 4, valerate was investigated within the metabolic data set due to observed changes in other compounds related to SCFA production in this chapter, alongside the variations associated with other SCFA pathways identified in the metagenomic data. The identification of valerate was less definitive relative to other compounds discussed in this chapter, with higher mass deltas and reduced FISh coverage from MS^2 analysis.

Valerate was reduced across all antibiotic instillation periods and remained so for the continuation of the experiment. This is in line with observations in mice by Zhu *et al.* describing CIP-induced disruptions to SCFA. (Zhu et al., 2020). The authors detected reductions in valerate similar to those in butyrate. Despite the limitations in compound annotation, supporting evidence from the literature (Zhu et al., 2020) alongside the experimental data from this current study present compelling evidence of valerate's potential as a bacterial marker of antibiotic-associated dysbiosis.

5.6.3 Summary

Identifying core taxonomic markers of gut dysbiosis is challenging due to interindividual variability. However, the functional redundancy among microbial constituents suggests that metabolomics may provide a more specific avenue for biomarker identification. This rationale was the basis for the analysis conducted in this chapter.

Although pathway characteristics may not be universal, there is a general trend towards the persistent reduction in the microbiota's capacity to produce SCFA as detected by LCMS metabolomic analysis. This is evidenced by reductions in butyrate and valerate abundance and an increase in pathway precursors 2-hydroxy butyrate and 2-oxoglutarate. Metabolomics offers the potential to reveal insights into dysbiosis, which are not effectively captured by metagenomic or culture-based approaches. The analysis in this chapter presents several putative metabolic markers for future studies, which could provide an opportunity for a panel-based biomarker approach to identifying gut dysbiosis. This may offer greater power and specificity compared to single-marker or taxonomic methods.

It is important to consider that the biomarkers discussed in this chapter have been identified on the basis of relative abundance rather than absolute quantification. It would be necessary to conduct further experiments to determine absolute quantification. This quantification would help establish biomarker ranger, and potentially provide insight and direct future mechanistic studies on the potential in-vivo impacts of metabolite changes, such as how changes in SCFA production may impact colonocytes.

Chapter 6
Conclusions

6.1 Concluding remarks

The purpose of this thesis was the use of different *in-vitro* models to explore bacterial markers of antibiotic-induced dysbiosis, with the goal of identifying putative biomarkers and establishing a framework for future multi-omics analysis of the microbiome using the Leeds HGM and MiGut platforms.

The stated aims of this thesis were:

1. To develop an effective and scalable pipeline for metabolomic analysis of *in-vitro* gut models and human faecal samples.
2. To determine whether *in-vitro* gut models can be used for the investigation of antibiotic-induced dysbiosis
3. To identify putative microbial markers of antibiotic-induced gut dysbiosis, using multi-omics approaches (metagenomics and metabolomics) in *in-vitro* gut models

The aims of this thesis were addressed through the development and validation of an untargeted LCMS-based metabolomics pipeline, exploring scalability and effective recovery and identification of metabolites for *in-vitro* model samples and from human faecal samples. *In-vitro* models of antibiotic-induced dysbiosis were employed to analyse the persistent impacts of clinically reflective antibiotic dosing regimens upon the microbiome. Loss of colonisation resistance and functional dysbiosis was confirmed through the colonisation and proliferation of *Clostridioides difficile* and a carbapenemase-producing *Klebsiella pneumoniae*. Metagenomic analysis identified differentially abundant genera through taxonomic analysis and metabolic pathways of interest via KEGG pathway analysis. Metabolomic analysis built upon the metagenomic findings by exploring metabolites of interest based on their associations with differentially abundant taxa and KEGG pathways.

6.1.1 Pipeline Development & Validation of LCMS

The development of a versatile pipeline for the untargeted analysis of the intestinal metabolome in Chapter 2 aimed to balance the breadth of data generation with cost and time efficiency. A C18 column was selected for its versatile chemistry and effective resolution of amphiphilic compounds in aqueous samples (Jandera, 2013).

The analytical limitations of C18 chromatograph were explored using a commercially available metabolite library. This exploration aimed to identify compound classes which may be underrepresented in data generated with this pipeline. The metabolite library also served as the basis for an in-house spectral library aimed at enhancing compound identification. Within this analysis, some compound classes were identified as being poorly resolved by the C18 method (fatty acyls and steroids, steroid derivatives and organooxygen compounds). This underrepresentation was attributed to their highly polar or non-polar nature, making them less suitable for C18 chromatography. While the underrepresentation of these compounds was deemed acceptable in this instance, future studies may explore additional column chemistries if these compound classes become of particular interest.

Bead beating was identified as a scalable alternative to manual disruption of faecal samples, making it more suitable for large-scale extractions as it reduces variation in sample processing potentially introduced via human errors in manual preparation.

The complexity of chemical structures and the vast pool of potential compounds limited the annotation of compounds in his study. Compound annotation remains a significant challenge within the field of metabolomic research (de Jonge et al., 2022), as evidenced by the post-processing steps of pipeline validation, where only 20% of compounds received named annotations. The majority of compounds were identified to predicted compositions, meaning that compounds of interest not matched to databases could be filtered and identified manually. Top-level data analysis exploring trends related to compound classes may miss significant amounts of information, as it is not feasible to employ this level of data curation for all compounds with the dataset.

6.1.2 Utility of Leeds HGM and MiGut Platform for In Vitro Biomarker Screening

The Leeds HGM, developed from models described by Macfarlane et al. (1998), has been utilised extensively in gut microbiome research, numerous publications supporting its application (Baines, S. D. et al., 2005; Freeman et al., 2005; Crowther et al., 2014). Th limitations of HGM system has been overcome through the development of the MiGut platform, which has also been validated for microbiome research (Davis Birch et al., 2023).

In-vitro studies offer advantages over human studies in a number of areas which make them an effective complementary approach to human studies. They address ethical concerns of studies requiring human participants. They offer improved and simplified longitudinal studies, as repeated spatial-temporal measurements of a single system can be taken within a controlled laboratory environment, also avoiding challenges with participant adherence and dropout. Additionally, *in-vitro* models also allow for the control of environmental variables such a nutrients and medication as well as excluding potential health conditions which may introduce additional variables. This allows *in-vitro* models to study isolated effects on the microbiome, making them powerful complementary tools in understanding the complex biological interactions within the microbiome.

There was some variation noted between MiGut model replicates, a challenge associated with novel systems and applications. However, metagenomic and metabolomic analyses were able to broadly identify biomarkers with consistent trends between models. As experience builds with the MiGut platform, it is likely that these variations may become reduced; the additional certainty these replicates give to the Leeds HGM providing additional samples and the capacity to run multiple models at once makes it an exciting candidate for Biomarker screening in future studies.

6.1.3 Proposed Bacterial Markers.

The work in this thesis identified a panel of **9 putative biomarkers** of antibiotic-induced dysbiosis, by applying a multi-omics workflow to these in vitro model systems. Table 6-1 Putative Biomarkers Identified in Antibiotic-Induced Gut Dysbiosis.

The key narrative that emerged when exploring the functional aspects of antibiotic-induced dysbiosis was a reduction in key taxa and functional pathways associated with SCFA production by the microbiome. These were further explored through metabolomics, which sought to identify key metabolites within the pathways which may be indicative of altered metabolic capacity.

Two taxonomic markers were identified with differential abundance analysis and highlighted for their associations with SCFA metabolism.

Faecalibacterium* and *Coprococcus were both identified by differential abundance analysis and selected as potential taxonomic markers due to their association with SCFA production (Sokol et al., 2008; Miquel et al., 2013; Vital et al., 2014).

KEGG pathway analysis identified differentially abundant identifications within the pathways for acetate/pyruvate, propanoate and butyrate metabolism and compounds within these pathways were selected as potential metabolic markers.

Table 6-1 Putative Biomarkers Identified in Antibiotic-Induced Gut Dysbiosis

Biomarker	Observed Trend	Biological Context	Discussed
2-Hydroxyglutarate	Increase	Metabolic intermediate	Chapter 5
2-Oxoglutarate	Increase	Metabolic intermediate	Chapter 5
Acetate[1]	Decrease[1]	SCFA	Chapter 5
Butyrate	Decrease	SCFA	Chapter 5
Coprococcus	Reduction	Microbiota	Chapter 4
Faecalibacterium	Reduction	Microbiota	Chapter 4
Lactate	Increase	Metabolic intermediate	Chapter 5
Propanoate[1]	Decrease[1]	SCFA	Chapter 5
Valerate	Decrease	SCFA	Chapter 5

[1]Biomarker and trend not observed due to experimental limitations, proposed trend based on wider SCFA observation. Candidate for targeted follow-up study

Butyrate was effectively captured by the LCMS method and was reduced across all models replicated following multiple antibiotics. Butyrate identification

was particularly notable due to its importance as an energy source for host colonocytes (Roediger, 1980) and antimicrobial properties associated with the inhibition of opportunistic pathogens (Chang Kai et al., 2024) (Sorbara et al., 2019).

Metabolic intermediates **2-hydroxyglutarate** and **2-oxoglutarate** were selected due to their association with the glutarate pathway for butyrate synthesis. The trend towards increases in these compounds was less conserved between model replicates. However, that may be attributable to varying mechanisms resulting in decreased butyrate productions between model replicates. The prevalence of CoA-conjugated metabolites across the SCFA pathways meant that many of the compounds fell outside the analytical window of LCMS analysis.

Lactate was identified as a putative marker of dysbiosis based on KEGG pathway analysis which revealed changes within the acrylate/lactate pathways for acetate and propanoate metabolism, these findings suggest lactate may indicate broader SCFA-associated disruptions, as well as potentially highlighting a niche for exploitation intestinal pathogens in dysbiotic microbiomes (Gillis et al., 2018). The increases in lactate abundance were conserved across all model replicates, strengthening its potential as a marker of antibiotic-induced dysbiosis.

Valerate showed consistent reductions across the majority of model replicates following antibiotic instillation. Although it was not identified in metagenomic analysis, broader trends in SCFA metabolism suggest that valerate is a metabolite worthy of further investigation.

While **acetate** and **propanoate** were not identified in metabolomic analysis, they were highlighted in metagenomics as pathways of interest. The cross-feeding of SCFA pathways and the observed reductions in butyrate and valerate abundance indicate that both acetate and propanoate merit further investigation in future models, with experimental parameters to ensure their recovery.

6.1.4 Limitations

This work focused on identifying bacterial markers of antibiotic induced dysbiosis through the application of multi-omics methodologies.

A primary limitation of this work is its focus on bacterial markers of dysbiosis. While the microbiome comprises bacteria, archaea, viruses and eukaryotes, the work primarily focuses on bacterial aspects. Metagenomic analysis in this thesis addresses prokaryotic taxonomy and KEGG metabolism within its bacterial associations. Metabolomics was used to explore the phenotypic resolution of antibiotic-induced dysbiosis, which does not actively exclude the impacts of another organism; the discussion within the thesis is primarily limited to bacterial functionality. Additionally, the *in vitro* models utilised in this thesis lack human cell lines. While this may be seen as a benefit in some instances, allowing the exploration of isolated bacterial mechanisms, further experiments are needed to confirm how these changes would impact host cells.

Given the size of the datasets generated by metagenomic and metabolomic analysis, it was not feasible to discuss all aspects and potential markers of dysbiosis within this thesis. Therefore, a choice was made to focus on alterations to SCFA metabolism and explore this in detail.

Due to the cumulative effects of successive antimicrobials, the emergence of identified biomarkers coinciding with a particular dosing period cannot be definitively attributed to the impact of that antibiotic alone.

While this thesis aimed to develop an untargeted LCMS pipeline for metabolomic analysis, practical considerations meant this was not entirely achievable. LCMS was selected over Gas Chromatography due to the relative ease of sample preparation (fewer requirements for sample derivatisation), suitability for less volatile compounds, and shorter chromatographic run times. As discussed in the metabolomic development chapter, there are limitations to what any single column chemistry can effectively resolve. This thesis identified compound classes which are not effectively captured by the C18 LCMS method utilised in this thesis, and therefore their contributions and potential as biomarkers may be missed. Additionally, when designing experimental parameters a scan range must be defined. In this instance, it was 80-800 Daltons, which led to some compounds such as acetate, propanoate and CoA-conjugated metabolites, falling outside the analytical window. Future studies should, employ additional scan ranges or column chemistries to effectively capture these metabolites. These concessions were made due to the considerable increases in machine time and financial investments required to

accommodate multiple column chemistries and scan ranges, which were outside the capacity of this thesis.

Finally, financial constraints associated with next-generation sequencing meant that metagenomic analysis was not feasible for all models. As a result, taxonomic and functional analysis is limited to the Leeds HGM and MiGut (3) replicate. As a result, it remains unclear whether the changes in taxonomic composition and functional capacity within the microbiome observed were conserved across model replicates.

6.1.5 Future Work

This thesis identified a panel of nine putative taxonomic and metabolic bacterial markers of antibiotic induced dysbiosis. Future work would aim to:

1. **explore the clinical potential of these biomarkers,** and the validation of these markers was planned in a study of patient samples, using excess diagnostic material. Ethical approval was submitted and granted via the Integrated Research Application System (IRAS) ID 307667. The study aimed to determine if and how these markers presented in clinical samples. The study aimed to collect samples submitted to Leeds Teaching Hospital Trust for *Clostridioides difficile* diagnostic testing, over a 9-month period. While the study received ethical approval, the initiation of sample collection was delayed due to COVID-19 restrictions and a lack of quoracy at data access committees. These delays ultimately resulted in insufficient time to collect the required number of participants for the study.
2. **apply these biomarkers to future *in-vitro* studies** This panel of biomarkers could also be profiled in future *in-vitro* models of antibiotic-induced dysbiosis, exploring how these biomarkers are conserved across different antimicrobial therapies.
3. **Expand metabolomic approach to identify further markers and verify those identified.** The significance of SCFA within metabolomic research led to a drive to better identify these compounds within complex biological matrices (Saha et al., 2021; Weng et al., 2024). As biomarker panels become better defined, optimised pipelines for detection and quantification of identified markers may be beneficial.

4. **apply the metabolomic methods developed in this thesis to broader areas of gastrointestinal disease.** The metabolomic methods outline in this thesis could also be applied to the wider areas of gastrointestinal disease such as IBD, colorectal cancer and other gastrointestinal infections.

Bibliography

(ChEBI), C.E.o.B.I. 2024. [Online]. Available from: https://www.ebi.ac.uk/chebi/

(NICE), N.I.f.H.a.C.E. 2021. *Clostridioides difficile infection: antimicrobial prescribing* [Online]. [Accessed 06/06/2022].

[NICE], N.I.f.H.a.C.E. 2021. *NICE guideline [NG199] : Clostridioides difficile infection: antimicrobial prescribing.*

[NICE], N.I.f.H.a.C.E. 2022. *Medical technologies guidance [MTG71] Faecal microbiota transplant for recurrent Clostridioides difficile infection.* https://www.nice.org.uk/: https://www.nice.org.uk/.

Adewiah, S., Abubakar, A. and Yusuf, F. 2018. IDDF2018-ABS-0023 Butyrate acid as a potential marker for the diversity of gut microbiota in colorectal cancer patients. *Gut.* **67**(Suppl 2), pp.A34-A34.

Ahmed, S., Macfarlane George, T., Fite, A., McBain Andrew, J., Gilbert, P. and Macfarlane, S. 2007. Mucosa-Associated Bacterial Diversity in Relation to Human Terminal Ileum and Colonic Biopsy Samples. *Applied and Environmental Microbiology.* **73**(22), pp.7435-7442.

Ajami, N.J., Cope, J.L., Wong, M.C., Petrosino, J.F. and Chesnel, L. 2018. Impact of Oral Fidaxomicin Administration on the Intestinal Microbiota and Susceptibility to Clostridium difficile Colonization in Mice. *Antimicrobial Agents and Chemotherapy.* **62**(5), pp.e02112-02117.

Akare, S., Jean-Louis, S., Chen, W., Wood, D.J., Powell, A.A. and Martinez, J.D. 2006. Ursodeoxycholic acid modulates histone acetylation and induces differentiation and senescence. *International Journal of Cancer.* **119**(12), pp.2958-2969.

Alam, M.S., Gangiredla, J., Hasan, N.A., Barnaba, T. and Tartera, C. 2021. Aging-Induced Dysbiosis of Gut Microbiota as a Risk Factor for Increased Listeria monocytogenes Infection. *Front Immunol.* **12**, p672353.

Albenberg, L., Esipova, T.V., Judge, C.P., Bittinger, K., Chen, J., Laughlin, A., Grunberg, S., Baldassano, R.N., Lewis, J.D., Li, H., Thom, S.R., Bushman, F.D., Vinogradov, S.A. and Wu, G.D. 2014. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology.* **147**(5), pp.1055-1063.e1058.

Anand, A. and Glatt, A.E. 1993. Clostridium difficile infection associated with antineoplastic chemotherapy: a review. *Clin Infect Dis.* **17**(1), pp.109-113.

Anders, S. and Huber, W. 2010. Differential expression analysis for sequence count data. *Genome Biology.* **11**(10), pR106.

Andriatahina, T., Randrianirina, F., Hariniana, E.R., Talarmin, A., Raobijaona, H., Buisson, Y. and Richard, V. 2010. High prevalence of fecal carriage of extended-spectrum β-lactamase-producing Escherichia coli and Klebsiella pneumoniae in a pediatric unit in Madagascar. *BMC Infectious Diseases.* **10**(1), p204.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang, J., Guarner, F., Pedersen, O., de Vos, W.M., Brunak, S., Doré, J., Antolín, M., Artiguenave, F., Blottiere, H.M., Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariaz, G., Dervyn, R., Foerstner, K.U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le Roux, K., Maguin, E., Mérieux, A., Melo Minardi, R., M'Rini, C., Muller, J., Oozeer, R., Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K., Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S.D. and Bork, P. 2011. Enterotypes of the human gut microbiome. *Nature.* **473**(7346), pp.174-180.

Asaduzzaman, S.M. and Sonomoto, K. 2009. Lantibiotics: Diverse activities and unique modes of action. *Journal of Bioscience and Bioengineering.* **107**(5), pp.475-487.

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., Taniguchi, T., Takeda, K., Hori, S., Ivanov, II, Umesaki, Y., Itoh, K. and Honda, K. 2011. Induction of colonic regulatory T cells by indigenous Clostridium species. *Science.* **331**(6015), pp.337-341.

Avuthu, N. and Guda, C. 2022. Meta-Analysis of Altered Gut Microbiota Reveals Microbial and Metabolic Biomarkers for Colorectal Cancer. *Microbiol Spectr.* **10**(4), pe0001322.

Babu, A.F., Koistinen, V.M., Turunen, S., Solano-Aguilar, G., Urban, J.F., Jr., Zarei, I. and Hanhineva, K. 2022. Identification and Distribution of Sterols, Bile Acids, and Acylcarnitines by LC-MS/MS in Humans, Mice, and Pigs-A Qualitative Analysis. *Metabolites.* **12**(1), p49.

Bağcı, C., Patz, S. and Huson, D.H. 2021. DIAMOND+MEGAN: Fast and Easy Taxonomic and Functional Analysis of Short and Long Microbiome Sequences. *Current Protocols.* **1**(3), pe59.

Baines, S.D., Freeman, J. and Wilcox, M.H. 2005. Effects of piperacillin/tazobactam on Clostridium difficile growth and toxin production in a human gut model. *J Antimicrob Chemother.* **55**(6), pp.974-982.

Baines, S.D., O'Connor, R., Huscroft, G., Saxton, K., Freeman, J. and Wilcox, M.H. 2009. Mecillinam: a low-risk antimicrobial agent for induction of Clostridium difficile infection in an in vitro human gut model. *Journal of Antimicrobial Chemotherapy.* **63**(4), pp.838-839.

Bana, B. and Cabreiro, F. 2019. The Microbiome and Aging. *Annu Rev Genet.* **53**, pp.239-261.

Barker-Tejeda, T.C., Zubeldia-Varela, E., Macías-Camero, A., Alonso, L., Martín-Antoniano, I.A., Rey-Stolle, M.F., Mera-Berriatua, L., Bazire, R., Cabrera-Freitag, P., Shanmuganathan, M., Britz-McKibbin, P., Ubeda, C., Francino, M.P., Barber, D., Ibáñez-Sandín, M.D., Barbas, C., Pérez-Gordo, M. and Villaseñor, A. 2024. Comparative characterization of the infant gut microbiome and their maternal lineage by a multi-omics approach. *Nature Communications.* **15**(1), p3004.

Bartlett, J.G. 1996. Management of Clostridium difficile infection and other antibiotic-associated diarrhoeas. *Eur J Gastroenterol Hepatol.* **8**(11), pp.1054-1061.

Bartlett, J.G., Onderdonk, A.B., Cisneros, R.L. and Kasper, D.L. 1977. Clindamycin-associated colitis due to a toxin-producing species of Clostridium in hamsters. *J Infect Dis.* **136**(5), pp.701-705.

Bartoloni, A., Mantella, A., Goldstein, B.P., Dei, R., Benedetti, M., Sbaragli, S. and Paradisi, F. 2004. In-Vitro Activity of Nisin Against Clinical Isolates of Clostridium difficile. *Journal of Chemotherapy.* **16**(2), pp.119-121.

Berkell, M., Mysara, M., Xavier, B.B., van Werkhoven, C.H., Monsieurs, P., Lammens, C., Ducher, A., Vehreschild, M., Goossens, H., de Gunzburg, J., Bonten, M.J.M. and Malhotra-Kumar, S. 2021. Microbiota-based markers predictive of development of Clostridioides difficile infection. *Nat Commun.* **12**(1), p2241.

Biomarkers Definitions Working Group. 2001. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* **69**(3), pp.89-95.

Bolger, A.M., Lohse, M. and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* **30**(15), pp.2114-2120.

Bosnjak, M., Karpe, A.V., Van, T.T.H., Kotsanas, D., Jenkin, G.A., Costello, S.P., Johanesen, P., Moore, R.J., Beale, D.J., Srikhanta, Y.N., Palombo, E.A., Larcombe, S. and Lyras, D. 2023. Multi-omics analysis of hospital-acquired diarrhoeal patients reveals biomarkers of enterococcal proliferation and Clostridioides difficile infection. *Nat Commun.* **14**(1), p7737.

Breukink, E., Wiedemann, I., Kraaij, C.v., Kuipers, O.P., Sahl, H.G. and de Kruijff, B. 1999. Use of the Cell Wall Precursor Lipid II by a Pore-Forming Peptide Antibiotic. *Science.* **286**(5448), p2361.

Buchfink, B., Xie, C. and Huson, D.H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods.* **12**(1), pp.59-60.

Buckley, A.M., Moura, I.B., Altringham, J., Ewin, D., Clark, E., Bentley, K., Wilkinson, V., Spittal, W., Davis, G. and Wilcox, M.H. 2021. The use of first-generation cephalosporin antibiotics, cefalexin and cefradine, is not associated with induction of simulated Clostridioides difficile infection. *J Antimicrob Chemother.* **77**(1), pp.148-154.

Caballero-Flores, G., Pickard, J.M. and Núñez, G. 2023. Microbiota-mediated colonization resistance: mechanisms and regulation. *Nature Reviews Microbiology.* **21**(6), pp.347-360.

Caballero, S., Kim, S., Carter, R.A., Leiner, I.M., Sušac, B., Miller, L., Kim, G.J., Ling, L. and Pamer, E.G. 2017. Cooperating Commensals Restore Colonization Resistance to Vancomycin-Resistant Enterococcus faecium. *Cell Host Microbe.* **21**(5), pp.592-602.e594.

Cammarota, G., Ianiro, G., Tilg, H., Rajilić-Stojanović, M., Kump, P., Satokari, R., Sokol, H., Arkkila, P., Pintus, C., Hart, A., Segal, J., Aloi, M., Masucci, L., Molinaro, A., Scaldaferri, F., Gasbarrini, G., Lopez-Sanroman, A., Link, A., de Groot, P., de Vos, W.M., Högenauer, C., Malfertheiner, P., Mattila, E., Milosavljević, T., Nieuwdorp, M., Sanguinetti, M., Simren, M. and Gasbarrini, A. 2017. European consensus conference on faecal microbiota transplantation in clinical practice. *Gut.* **66**(4), pp.569-580.

Canfora, E.E., Jocken, J.W. and Blaak, E.E. 2015. Short-chain fatty acids in control of body weight and insulin sensitivity. *Nature Reviews Endocrinology.* **11**(10), pp.577-591.

Cascales, E., Buchanan, S.K., Duché, D., Kleanthous, C., Lloubès, R., Postle, K., Riley, M., Slatin, S. and Cavard, D. 2007. Colicin Biology. *Microbiology and Molecular Biology Reviews.* **71**(1), p158.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D.S., Weerasinghe, D., Zhang, P. and Karp, P.D. 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research.* **42**(D1), pp.D459-D471.

Catterson, J.H., Khericha, M., Dyson, M.C., Vincent, A.J., Callard, R., Haveron, S.M., Rajasingam, A., Ahmad, M. and Partridge, L. 2018. Short-Term, Intermittent Fasting Induces Long-Lasting Gut Health and TOR-Independent Lifespan Extension. *Curr Biol.* **28**(11), pp.1714-1724.e1714.

Cech, N.B. and Enke, C.G. 2001. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrometry Reviews.* **20**(6), pp.362-387.

Centers for Disease Control and Prevention, Infectious Diseases Society of America and American Society of Blood and Marrow Transplantation. 2000. Guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients. *Biol Blood Marrow Transplant.* **6**(6a), pp.7-83.

Chang Kai, C., Nagarajan, N. and Gan, Y.-H. 2024. Short-chain fatty acids of various lengths differentially inhibit Klebsiella pneumoniae and Enterobacteriaceae species. *mSphere.* **9**(2), pp.e00781-00723.

Chen, C., Yang, X. and Shen, X. 2019. Confirmed and Potential Roles of Bacterial T6SSs in the Intestinal Ecosystem. *Frontiers in Microbiology.* **10**(1484).

Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Segal, E., Weersma, R.K., Wijmenga, C., Zhernakova, A. and Fu, J. 2021. The long-term genetic stability and individual specificity of the human gut microbiome. *Cell.* **184**(9), pp.2302-2315.e2312.

Chilton, C.H., Crowther, G.S., Todhunter, S.L., Nicholson, S., Freeman, J., Chesnel, L. and Wilcox, M.H. 2014. Efficacy of surotomycin in an in vitro gut model of Clostridium difficile infection. *J Antimicrob Chemother.* **69**(9), pp.2426-2433.

Chilton, C.H., Freeman, J., Crowther, G.S., Todhunter, S.L., Nicholson, S. and Wilcox, M.H. 2012. Co-amoxiclav induces proliferation and cytotoxin production of Clostridium difficile ribotype 027 in a human gut model. *Journal of Antimicrobial Chemotherapy.* **67**(4), pp.951-954.

Chong, P.P. and Koh, A.Y. 2020. The gut microbiota in transplant patients. *Blood Reviews.* **39**, p100614.

Christensson, B., Nilsson-Ehle, I., Ljungberg, B., Nömm, I., Oscarsson, G., Nordström, L., Goscinsky, G., Löwdin, E., Linglöf, T., Nordström, B., Denstedt-Stigzelius, I., Lindhagen, K., Edlund, C. and Nord, C.E. 1991. A randomized multicenter trial to compare the influence of cefaclor and amoxycillin on the colonization resistance of the digestive tract in patients with lower respiratory tract infection. *Infection.* **19**(4), pp.208-215.

Cox, Laura M., Yamanishi, S., Sohn, J., Alekseyenko, Alexander V., Leung, Jacqueline M., Cho, I., Kim, Sungheon G., Li, H., Gao, Z., Mahana, D., Zárate Rodriguez, Jorge G., Rogers, Arlin B., Robine, N., Loke, P.n. and Blaser, Martin J. 2014. Altering the Intestinal Microbiota during a Critical Developmental Window Has Lasting Metabolic Consequences. *Cell.* **158**(4), pp.705-721.

Crowther, G.S., Chilton, C.H., Todhunter, S.L., Nicholson, S., Freeman, J., Baines, S.D. and Wilcox, M.H. 2014. Development and validation of a chemostat gut model to study both planktonic and biofilm modes of growth of Clostridium difficile and human microbiota. *PLoS One.* **9**(2), pe88396.

Crowther, G.S., Wilcox, M.H. and Chilton, C.H. 2016. An In Vitro Model of the Human Colon: Studies of Intestinal Biofilms and Clostridium difficile Infection. *Methods Mol Biol.* **1476**, pp.223-234.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., Marks, K.M., Prins, R.M., Ward, P.S., Yen, K.E., Liau, L.M., Rabinowitz, J.D., Cantley, L.C., Thompson, C.B., Vander Heiden, M.G. and Su, S.M. 2009. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature.* **462**(7274), pp.739-744.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., Biddinger, S.B., Dutton, R.J. and Turnbaugh, P.J. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* **505**(7484), pp.559-563.

Davis Birch, W.A., Moura, I.B., Ewin, D.J., Wilcox, M.H., Buckley, A.M., Culmer, P.R. and Kapur, N. 2023. MiGut: A scalable in vitro platform for simulating the human gut microbiome-Development, validation and simulation of antibiotic-induced dysbiosis. *Microb Biotechnol.* **16**(6), pp.1312-1324.

de Jonge, N.F., Mildau, K., Meijer, D., Louwen, J.J.R., Bueschl, C., Huber, F. and van der Hooft, J.J.J. 2022. Good practices and recommendations for using and benchmarking computational metabolomics metabolite annotation tools. *Metabolomics.* **18**(12), p103.

De Vuyst, L. and Leroy, F. 2011. Cross-feeding between bifidobacteria and butyrate-producing colon bacteria explains bifdobacterial competitiveness, butyrate production, and gas production. *International Journal of Food Microbiology.* **149**(1), pp.73-80.

De Vuyst, L., Moens, F., Selak, M., Rivière, A. and Leroy, F. 2014. Summer Meeting 2013: growth and physiology of bifidobacteria. *J Appl Microbiol.* **116**(3), pp.477-491.

Debast, S.B., Bauer, M.P. and Kuijper, E.J. 2014. European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for Clostridium difficile infection. *Clin Microbiol Infect.* **20 Suppl 2**, pp.1-26.

Dekkers, K.F., Sayols-Baixeras, S., Baldanzi, G., Nowak, C., Hammar, U., Nguyen, D., Varotsis, G., Brunkwall, L., Nielsen, N., Eklund, A.C., Bak Holm, J., Nielsen, H.B., Ottosson, F., Lin, Y.-T., Ahmad, S., Lind, L., Sundström, J., Engström, G., Smith, J.G., Ärnlöv, J., Orho-Melander, M. and Fall, T. 2022. An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nature Communications.* **13**(1), p5370.

Deriu, E., Liu, J.Z., Pezeshki, M., Edwards, R.A., Ochoa, R.J., Contreras, H., Libby, S.J., Fang, F.C. and Raffatellu, M. 2013. Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host Microbe.* **14**(1), pp.26-37.

Dong, Y., Han, M., Fei, T., Liu, H. and Gai, Z. 2024. Utilization of diverse oligosaccharides for growth by Bifidobacterium and Lactobacillus species and their in vitro co-cultivation characteristics. *International Microbiology.* **27**(3), pp.941-952.

Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K. and Bultman, S.J. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell metabolism.* **13**(5), pp.517-526.

Ducarmon, Q.R., Terveer, E.M., Nooij, S., Bloem, M.N., Vendrik, K.E.W., Caljouw, M.A.A., Sanders, I.M.J.G., van Dorp, S.M., Wong, M.C., Zwittink, R.D.

and Kuijper, E.J. 2021. Microbiota-associated risk factors for asymptomatic gut colonisation with multi-drug-resistant organisms in a Dutch nursing home. *Genome Medicine.* **13**(1), p54.

Duncan, S.H., Hold, G.L., Harmsen, H.J.M., Stewart, C.S. and Flint, H.J. 2002. Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as Faecalibacterium prausnitzii gen. nov., comb. nov. *Int J Syst Evol Microbiol.* **52**(Pt 6), pp.2141-2146.

Dysangco, A.T., Pierce, T.M. and Bravata, D.M. 2022. Risk of antibiotics associated with Clostridioides difficile infection: Antibiotic stewardship in action. *Antimicrobial Stewardship & Healthcare Epidemiology.* **2**(1), pe146 article no: e146 [no pagination].

Elvers, K.T., Wilson, V.J., Hammond, A., Duncan, L., Huntley, A.L., Hay, A.D. and van der Werf, E.T. 2020. Antibiotic-induced changes in the human gut microbiota for the most commonly prescribed antibiotics in primary care in the UK: a systematic review. *BMJ Open.* **10**(9), pe035677.

Erben, V., Poschet, G., Schrotz-King, P. and Brenner, H. 2021. Evaluation of different stool extraction methods for metabolomics measurements in human faecal samples. *BMJ Nutrition, Prevention & Health.* **4**(2), p374.

Ewels, P., Magnusson, M., Lundin, S. and Käller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* **32**(19), pp.3047-3048.

Eze, P., Balsells, E., Kyaw, M.H. and Nair, H. 2017. Risk factors for Clostridium difficile infections - an overview of the evidence base and challenges in data synthesis. *Journal of global health.* **7**(1), pp.010417-010417.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., Rosenbaum, M. and Gordon, J.I. 2013. The Long-Term Stability of the Human Gut Microbiota. *Science.* **341**(6141), p1237439.

Falony, G., Calmeyn, T., Leroy, F. and De Vuyst, L. 2009. Coculture Fermentations of Bifidobacterium Species and Bacteroides thetaiotaomicron Reveal a Mechanistic Insight into the Prebiotic Effect of Inulin-Type Fructans. *Applied and Environmental Microbiology.* **75**(8), pp.2312-2319.

Falony, G., Vlachou, A., Verbrugghe, K. and De Vuyst, L. 2006. Cross-feeding between Bifidobacterium longum BB536 and acetate-converting, butyrate-producing colon bacteria during growth on oligofructose. *Appl Environ Microbiol.* **72**(12), pp.7835-7841.

FDA. 2019. *Fecal Microbiota for Transplantation: Safety Communication- Risk of Serious Adverse Reactions Due to Transmission of Multi-Drug Resistant Organisms.* [Online]. [Accessed 12/2024]. Available from: https://www.fda.gov/safety/medical-product-safety-information/fecal-microbiota-transplantation-safety-communication-risk-serious-adverse-reactions-due

Feuerstadt, P., Chopra, T., Knapple, W., Van Hise, N.W., Dubberke, E.R., Baggott, B., Guthmueller, B., Bancke, L., Gamborg, M., Steiner, T.S., Van Handel, D. and Khanna, S. 2024. PUNCH CD3-OLS: A Phase 3 Prospective Observational Cohort Study to Evaluate the Safety and Efficacy of Fecal Microbiota, Live-jslm (REBYOTA) in Adults With Recurrent Clostridioides difficile Infection. *Clinical Infectious Diseases.* pciae437.

Feuerstadt, P., Louie, T.J., Lashner, B., Wang, E.E.L., Diao, L., Bryant, J.A., Sims, M., Kraft, C.S., Cohen, S.H., Berenson, C.S., Korman, L.Y., Ford, C.B., Litcofsky, K.D., Lombardo, M.J., Wortman, J.R., Wu, H., Auniņš, J.G., McChalicher, C.W.J., Winkler, J.A., McGovern, B.H., Trucksis, M., Henn, M.R. and von Moltke, L. 2022. SER-109, an Oral Microbiome Therapy for Recurrent Clostridioides difficile Infection. *N Engl J Med.* **386**(3), pp.220-229.

Foskolou, I.P., Bunse, L. and Van den Bossche, J. 2023. 2-hydroxyglutarate rides the cancer-immunity cycle. *Curr Opin Biotechnol.* **83**, p102976.

Franzosa, E.A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H.J., Reinker, S., Vatanen, T., Hall, A.B., Mallick, H., McIver, L.J., Sauk, J.S., Wilson, R.G., Stevens, B.W., Scott, J.M., Pierce, K., Deik, A.A., Bullock, K., Imhann, F., Porter, J.A., Zhernakova, A., Fu, J., Weersma, R.K., Wijmenga, C., Clish, C.B., Vlamakis, H., Huttenhower, C. and Xavier, R.J. 2019. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol.* **4**(2), pp.293-305.

Freeman, J., Baines, S.D., Jabes, D. and Wilcox, M.H. 2005. Comparison of the efficacy of ramoplanin and vancomycin in both in vitro and in vivo models of clindamycin-induced Clostridium difficile infection. *J Antimicrob Chemother.* **56**(4), pp.717-725.

Freeman J, M.I., Buckley A, Spittal W, Bentley K, Morgan A, Clark E, Cross R, Ewin D, Wang D, Mitra S, Wilcox M. 2021. Guess Poo: modelling Antibiotic-mediated Dysbiosis. In: *Anaerobe 21*, *Cardiff, UK*.

Freeman, J., O'Neill, F.J. and Wilcox, M.H. 2003. Effects of cefotaxime and desacetylcefotaxime upon Clostridium difficile proliferation and toxin production in a triple-stage chemostat model of the human gut. *J Antimicrob Chemother.* **52**(1), pp.96-102.

Frolova, M.S., Suvorova, I.A., Iablokov, S.N., Petrov, S.N. and Rodionov, D.A. 2022. Genomic reconstruction of short-chain fatty acid production by the human gut microbiota. *Frontiers in Molecular Biosciences.* **9**.

Gerassy-Vainberg, S., Blatt, A., Danin-Poleg, Y., Gershovich, K., Sabo, E., Nevelsky, A., Daniel, S., Dahan, A., Ziv, O., Dheer, R., Abreu, M.T., Koren, O., Kashi, Y. and Chowers, Y. 2018. Radiation induces proinflammatory dysbiosis: transmission of inflammatory susceptibility by host cytokine induction. *Gut.* **67**(1), pp.97-107.

German, J.B. and Dillard, C.J. 2004. Saturated fats: what dietary intake? *Am J Clin Nutr.* **80**(3), pp.550-559.

Gianetto-Hill, C.M., Vancuren, S.J., Daisley, B., Renwick, S., Wilde, J., Schroeter, K., Daigneault, M.C. and Allen-Vercoe, E. 2023. The Robogut: A Bioreactor Model of the Human Colon for Evaluation of Gut Microbial Community Ecology and Function. *Curr Protoc.* **3**(4), pe737.

Gillis, C.C., Hughes, E.R., Spiga, L., Winter, M.G., Zhu, W., Furtado de Carvalho, T., Chanin, R.B., Behrendt, C.L., Hooper, L.V., Santos, R.L. and Winter, S.E. 2018. Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. *Cell Host & Microbe.* **23**(1), pp.54-64.e56.

Graham, D.B. and Xavier, R.J. 2023. Conditioning of the immune system by the microbiome. *Trends Immunol.* **44**(7), pp.499-511.

Grönlund, M.M., Arvilommi, H., Kero, P., Lehtonen, O.P. and Isolauri, E. 2000. Importance of intestinal colonisation in the maturation of humoral immunity in early infancy: a prospective follow up study of healthy infants aged 0-6 months. *Archives of disease in childhood. Fetal and neonatal edition.* **83**(3), pp.F186-F192.

Gupta, V.K., Kim, M., Bakshi, U., Cunningham, K.Y., Davis, J.M., Lazaridis, K.N., Nelson, H., Chia, N. and Sung, J. 2020. A predictive index for health status using species-level gut microbiome profiling. *Nature Communications.* **11**(1), p4635.

Gutierrez, D., Weinstock, A., Antharam, V.C., Gu, H., Jasbi, P., Shi, X., Dirks, B., Krajmalnik-Brown, R., Maldonado, J., Guinan, J. and Thangamani, S. 2020. Antibiotic-induced gut metabolome and microbiome alterations increase the susceptibility to Candida albicans colonization in the gastrointestinal tract. *FEMS Microbiol Ecol.* **96**(1).

Harris, H.C., Buckley, A.M., Spittal, W., Ewin, D., Clark, E., Altringham, J., Bentley, K., Moura, I.B., Wilcox, M.H., Woodford, N., Davies, K. and Chilton, C.H. 2021. The effect of intestinal microbiota dysbiosis on growth and detection of carbapenemase-producing Enterobacterales within an in vitro gut model. *J Hosp Infect.* **113**, pp.1-9.

Hasper, H.E., Kramer, N.E., Smith, J.L., Hillman, J.D., Zachariah, C., Kuipers, O.P., de Kruijff, B. and Breukink, E. 2006. An Alternative Bactericidal Mechanism of Action for Lantibiotic Peptides That Target Lipid II. *Science.* **313**(5793), p1636.

Hayashi, A., Nagao-Kitamoto, H., Kitamoto, S., Kim, C.H. and Kamada, N. 2021. The Butyrate-Producing Bacterium Clostridium butyricum Suppresses Clostridioides difficile Infection via Neutrophil- and Antimicrobial Cytokine–Dependent but GPR43/109a-Independent Mechanisms. *The Journal of Immunology.* **206**(7), pp.1576-1585.

Hayashi, H., Takahashi, R., Nishi, T., Sakamoto, M. and Benno, Y. 2005. Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *J Med Microbiol.* **54**(Pt 11), pp.1093-1101.

Head, S.R., Komori, H.K., LaMere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R. and Ordoukhanian, P. 2014. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques.* **56**(2), pp.61-64, 66, 68, passim.

Hirano, S., Nakama, R., Tamaki, M., Masuda, N. and Oda, H. 1981. Isolation and characterization of thirteen intestinal microorganisms capable of 7 alpha-dehydroxylating bile acids. *Applied and environmental microbiology.* **41**(3), pp.737-745.

Hofmann, J.D., Otto, A., Berges, M., Biedendieck, R., Michel, A.M., Becher, D., Jahn, D. and Neumann-Schaal, M. 2018. Metabolic Reprogramming of Clostridioides difficile During the Stationary Phase With the Induction of Toxin Production. *Front Microbiol.* **9**, p1970.

Holmes, A.J., Chew, Y.V., Colakoglu, F., Cliff, J.B., Klaassens, E., Read, M.N., Solon-Biet, S.M., McMahon, A.C., Cogger, V.C., Ruohonen, K., Raubenheimer, D., Le Couteur, D.G. and Simpson, S.J. 2017. Diet-Microbiome Interactions in Health Are Controlled by Intestinal Nitrogen Source Constraints. *Cell Metab.* **25**(1), pp.140-151.

Hooper, L.V., Stappenbeck, T.S., Hong, C.V. and Gordon, J.I. 2003. Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat Immunol.* **4**(3), pp.269-273.

Hou, K., Wu, Z.-X., Chen, X.-Y., Wang, J.-Q., Zhang, D., Xiao, C., Zhu, D., Koya, J.B., Wei, L., Li, J. and Chen, Z.-S. 2022. Microbiota in health and diseases. *Signal Transduction and Targeted Therapy.* **7**(1), p135.

Human Microbiomes Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature.* **486**(7402), pp.207-214.

Iino, T., Mori, K., Tanaka, K., Suzuki, K.-i. and Harayama, S. 2007. Oscillibacter valericigenes gen. nov., sp. nov., a valerate-producing anaerobic bacterium isolated from the alimentary canal of a Japanese corbicula clam. *International Journal of Systematic and Evolutionary Microbiology.* **57**(8), pp.1840-1845.

Ilhan, Z.E., Marcus, A.K., Kang, D.W., Rittmann, B.E. and Krajmalnik-Brown, R. 2017. pH-Mediated Microbial and Metabolic Interactions in Fecal Enrichment Cultures. *mSphere.* **2**(3).

Imhann, F., Bonder, M.J., Vich Vila, A., Fu, J., Mujagic, Z., Vork, L., Tigchelaar, E.F., Jankipersadsing, S.A., Cenit, M.C., Harmsen, H.J.M., Dijkstra, G., Franke, L., Xavier, R.J., Jonkers, D., Wijmenga, C., Weersma, R.K. and Zhernakova, A. 2016. Proton pump inhibitors affect the gut microbiome. *Gut.* **65**(5), pp.740-748.

Iskandar, K., Molinier, L., Hallit, S., Sartelli, M., Hardcastle, T.C., Haque, M., Lugova, H., Dhingra, S., Sharma, P., Islam, S., Mohammed, I., Naina Mohamed, I., Hanna, P.A., Hajj, S.E., Jamaluddin, N.A.H., Salameh, P. and Roques, C. 2021. Surveillance of antimicrobial resistance in low- and middle-income countries: a scattered picture. *Antimicrob Resist Infect Control.* **10**(1), p63.

Jandera, P. 2013. LIQUID CHROMATOGRAPHY | Normal Phase☆. In: Worsfold, P., Poole, C., Townshend, A. and Miró, M. eds. *Encyclopedia of Analytical Science (Third Edition).* Oxford: Academic Press, pp.162-173.

Jeffery, I.B., Lynch, D.B. and O'Toole, P.W. 2016. Composition and temporal stability of the gut microbiota in older persons. *The ISME Journal.* **10**(1), pp.170-182.

Jernberg, C., Löfmark, S., Edlund, C. and Jansson, J.K. 2007. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *The ISME Journal.* **1**(1), pp.56-66.

Kaakoush, N.O. 2020. Sutterella Species, IgA-degrading Bacteria in Ulcerative Colitis. *Trends in Microbiology.* **28**(7), pp.519-522.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**(D1), pp.D587-d592.

Kanehisa, M. and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), pp.27-30.

Kang, J.D., Myers, C.J., Harris, S.C., Kakiyama, G., Lee, I.K., Yun, B.S., Matsuzaki, K., Furukawa, M., Min, H.K., Bajaj, J.S., Zhou, H. and Hylemon, P.B. 2019. Bile Acid 7α-Dehydroxylating Gut Bacteria Secrete Antibiotics that Inhibit Clostridium difficile: Role of Secondary Bile Acids. *Cell Chem Biol.* **26**(1), pp.27-34.e24.

Kappel, B.A., De Angelis, L., Heiser, M., Ballanti, M., Stoehr, R., Goettsch, C., Mavilio, M., Artati, A., Paoluzi, O.A., Adamski, J., Mingrone, G., Staels, B., Burcelin, R., Monteleone, G., Menghini, R., Marx, N. and Federici, M. 2020. Cross-omics analysis revealed gut microbiome-related metabolic pathways underlying atherosclerosis development after antibiotics treatment. *Mol Metab.* **36**, p100976.

KEGG. 2019. *Carbohydrate digestion and absorption.* [Online].

Kelly, C.R., Khoruts, A., Staley, C., Sadowsky, M.J., Abd, M., Alani, M., Bakow, B., Curran, P., McKenney, J., Tisch, A., Reinert, S.E., Machan, J.T. and Brandt, L.J. 2016. Effect of Fecal Microbiota Transplantation on Recurrence in Multiply Recurrent Clostridium difficile Infection: A Randomized Trial. *Ann Intern Med.* **165**(9), pp.609-616.

Kelly, P.E., Ng, H.J., Farrell, G., McKirdy, S., Russell, R.K., Hansen, R., Rattray, Z., Gerasimidis, K. and Rattray, N.J.W. 2022. An Optimised Monophasic Faecal Extraction Method for LC-MS Analysis and Its Application in Gastrointestinal Disease. *Metabolites.* **12**(11).

Keohane, D.M., Ghosh, T.S., Jeffery, I.B., Molloy, M.G., O'Toole, P.W. and Shanahan, F. 2020. Microbiome and health implications for ethnic minorities after enforced lifestyle changes. *Nature medicine.* **26**(7), pp.1089-1095.

Kiddee, A., Assawatheptawee, K., Na-udom, A., Boonsawang, P., Treebupachatsakul, P., Walsh, T.R. and Niumsup, P.R. 2019. Risk Factors for Extended-Spectrum β-Lactamase-Producing Enterobacteriaceae Carriage in Patients Admitted to Intensive Care Unit in a Tertiary Care Hospital in Thailand. *Microbial Drug Resistance.* **25**(8), pp.1182-1190.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., Zaslavsky, L., Zhang, J. and Bolton, E.E. 2023. PubChem 2023 update. *Nucleic Acids Research.* **51**(D1), pp.D1373-D1380.

Kim, S., Covington, A. and Pamer, E.G. 2017. The intestinal microbiota: Antibiotics, colonization resistance, and enteric pathogens. *Immunol Rev.* **279**(1), pp.90-105.

Kishino, S., Takeuchi, M., Park, S.-B., Hirata, A., Kitamura, N., Kunisawa, J., Kiyono, H., Iwamoto, R., Isobe, Y., Arita, M., Arai, H., Ueda, K., Shima, J., Takahashi, S., Yokozeki, K., Shimizu, S. and Ogawa, J. 2013. Polyunsaturated fatty acid saturation by gut lactic acid bacteria affecting host lipid composition. *Proceedings of the National Academy of Sciences.* **110**(44), pp.17808-17813.

Klein, E.Y., Van Boeckel, T.P., Martinez, E.M., Pant, S., Gandra, S., Levin, S.A., Goossens, H. and Laxminarayan, R. 2018. Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proceedings of the National Academy of Sciences.* **115**(15), pp.E3463-E3470.

Langfeld, L.Q., Du, K., Bereswill, S. and Heimesaat, M.M. 2021. A review of the antimicrobial and immune-modulatory properties of the gut microbiota-derived short chain fatty acid propionate - What is new? *Eur J Microbiol Immunol (Bp).* **11**(2), pp.50-56.

LeBlanc, J.G., Chain, F., Martín, R., Bermúdez-Humarán, L.G., Courau, S. and Langella, P. 2017. Beneficial effects on host energy metabolism of short-chain fatty acids and vitamins produced by commensal and probiotic bacteria. *Microb Cell Fact.* **16**(1), p79.

Lee, J., d'Aigle, J., Atadja, L., Quaicoe, V., Honarpisheh, P., Ganesh, B.P., Hassan, A., Graf, J., Petrosino, J., Putluri, N., Zhu, L., Durgan, D.J., Bryan, R.M., Jr., McCullough, L.D. and Venna, V.R. 2020. Gut Microbiota-Derived Short-Chain Fatty Acids Promote Poststroke Recovery in Aged Mice. *Circ Res.* **127**(4), pp.453-465.

Lee, S.M., Donaldson, G.P., Mikulski, Z., Boyajian, S., Ley, K. and Mazmanian, S.K. 2013. Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature.* **501**(7467), pp.426-429.

Leimena, M.M., Ramiro-Garcia, J., Davids, M., van den Bogert, B., Smidt, H., Smid, E.J., Boekhorst, J., Zoetendal, E.G., Schaap, P.J. and Kleerebezem, M. 2013. A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genomics.* **14**(1), p530.

Lekang, K., Shekhar, S., Berild, D., Petersen, F.C. and Winther-Larsen, H.C. 2022. Effects of different amoxicillin treatment durations on microbiome diversity and composition in the gut. *PLOS ONE.* **17**(10), pe0275737.

Leverrier, P., Dimova, D., Pichereau, V., Auffray, Y., Boyaval, P. and Jan, G. 2003. Susceptibility and adaptive response to bile salts in Propionibacterium freudenreichii: physiological and proteomic analysis. *Appl Environ Microbiol.* **69**(7), pp.3809-3818.

Lewis, B.B., Buffie, C.G., Carter, R.A., Leiner, I., Toussaint, N.C., Miller, L.C., Gobourne, A., Ling, L. and Pamer, E.G. 2015. Loss of Microbiota-Mediated Colonization Resistance to Clostridium difficile Infection With Oral Vancomycin Compared With Metronidazole. *J Infect Dis.* **212**(10), pp.1656-1665.

Lewis, G., Wang, B., Shafiei Jahani, P., Hurrell, B.P., Banie, H., Aleman Muench, G.R., Maazi, H., Helou, D.G., Howard, E., Galle-Treger, L., Lo, R., Santosh, S., Baltus, A., Bongers, G., San-Mateo, L., Gilliland, F.D., Rehan, V.K., Soroosh, P. and Akbari, O. 2019. Dietary Fiber-Induced Microbial Short Chain Fatty Acids Suppress ILC2-Dependent Airway Inflammation. *Front Immunol.* **10**, p2051.

Liu, C., Du, M.-X., Abuduaini, R., Yu, H.-Y., Li, D.-H., Wang, Y.-J., Zhou, N., Jiang, M.-Z., Niu, P.-X., Han, S.-S., Chen, H.-H., Shi, W.-Y., Wu, L., Xin, Y.-H., Ma, J., Zhou, Y., Jiang, C.-Y., Liu, H.-W. and Liu, S.-J. 2021. Enlightening the taxonomy darkness of human gut microbiomes with a cultured biobank. *Microbiome.* **9**(1), p119.

Liu, X., Qiu, X., Yang, Y., Wang, J., Wang, Q., Liu, J., Yang, F., Liu, Z. and Qi, R. 2023. Alteration of gut microbiome and metabolome by Clostridium butyricum can repair the intestinal dysbiosis caused by antibiotics in mice. *iScience.* **26**(3), p106190.

Louie, T., Golan, Y., Khanna, S., Bobilev, D., Erpelding, N., Fratazzi, C., Carini, M., Menon, R., Ruisi, M., Norman, J.M., Faith, J.J., Olle, B., Li, M., Silber, J.L. and Pardi, D.S. 2023. VE303, a Defined Bacterial Consortium, for Prevention of Recurrent Clostridioides difficile Infection: A Randomized Clinical Trial. *Jama.* **329**(16), pp.1356-1366.

Louis, P., Duncan, S.H., Sheridan, P.O., Walker, A.W. and Flint, H.J. 2022. Microbial lactate utilisation and the stability of the gut microbiome. *Gut Microbiome.* **3**, pe3 article no: e3 [no pagination].

Lu, W., Su, X., Klein, M.S., Lewis, I.A., Fiehn, O. and Rabinowitz, J.D. 2017. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annu Rev Biochem.* **86**, pp.277-304.

Luan, Z., Sun, G., Huang, Y., Yang, Y., Yang, R., Li, C., Wang, T., Tan, D., Qi, S., Jun, C., Wang, C., Wang, S., Zhao, Y. and Jing, Y. 2020. Metagenomics Study Reveals Changes in Gut Microbiota in Centenarians: A Cohort Study of Hainan Centenarians. *Frontiers in Microbiology.* **11**(1474).

Macfarlane, G.T., Macfarlane, S. and Gibson, G.R. 1998. Validation of a Three-Stage Compound Continuous Culture System for Investigating the Effect of Retention Time on the Ecology and Metabolism of Bacteria in the Human Colon. *Microb Ecol.* **35**(2), pp.180-187.

Macke, L., Schulz, C., Koletzko, L. and Malfertheiner, P. 2020. Systematic review: the effects of proton pump inhibitors on the microbiome of the digestive tract-evidence from next-generation sequencing studies. *Aliment Pharmacol Ther.* **51**(5), pp.505-526.

Maestri, O. and Joset, F. 2000. Regulation by external pH and stationary growth phase of the acetolactate synthase from Synechocystis PCC6803. *Molecular Microbiology.* **37**(4), pp.828-838.

Magnúsdóttir, S., Ravcheev, D., de Crécy-Lagard, V. and Thiele, I. 2015. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Frontiers in Genetics.* **6**.

Mahieu, R., Cassisa, V., Hilliquin, D., Coron, N., Pailhoriès, H., Kempf, M., Joly-Guillou, M.L. and Eveillard, M. 2017. Impact of faecal microbiota transplantation on mouse digestive colonization with two extensively resistant bacteria. *J Infect.* **75**(1), pp.75-77.

Mallott, E.K. and Amato, K.R. 2022. Butyrate Production Pathway Abundances Are Similar in Human and Nonhuman Primate Gut Microbiomes. *Molecular Biology and Evolution.* **39**(1), pmsab279.

Malm, L., Palm, E., Souihi, A., Plassmann, M., Liigand, J. and Kruve, A. 2021. Guide to Semi-Quantitative Non-Targeted Screening Using LC/ESI/HRMS. *Molecules.* [Online]. **26**(12). Available from: https://doi.org/10.3390/molecules26123524

McDonald, D., Hyde, E., Debelius, J.W., Morton, J.T., Gonzalez, A., Ackermann, G., Aksenov, A.A., Behsaz, B., Brennan, C., Chen, Y., DeRight Goldasich, L., Dorrestein, P.C., Dunn, R.R., Fahimipour, A.K., Gaffney, J., Gilbert, J.A., Gogul, G., Green, J.L., Hugenholtz, P., Humphrey, G., Huttenhower, C., Jackson, M.A., Janssen, S., Jeste, D.V., Jiang, L., Kelley, S.T., Knights, D., Kosciolek, T., Ladau, J., Leach, J., Marotz, C., Meleshko, D., Melnik, A.V., Metcalf, J.L., Mohimani, H., Montassier, E., Navas-Molina, J., Nguyen, T.T., Peddada, S., Pevzner, P., Pollard, K.S., Rahnavard, G., Robbins-Pianka, A., Sangwan, N., Shorenstein, J., Smarr, L., Song, S.J., Spector, T., Swafford, A.D., Thackray, V.G., Thompson, L.R., Tripathi, A., Vázquez-Baeza, Y., Vrbanac, A., Wischmeyer, P., Wolfe, E., Zhu, Q. and Knight, R. 2018. American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems.* **3**(3), pp.e00031-00018.

McDonald, L.C., Gerding, D.N., Johnson, S., Bakken, J.S., Carroll, K.C., Coffin, S.E., Dubberke, E.R., Garey, K.W., Gould, C.V., Kelly, C., Loo, V., Shaklee Sammons, J., Sandora, T.J. and Wilcox, M.H. 2018. Clinical Practice Guidelines for Clostridium difficile Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). *Clinical Infectious Diseases.* **66**(7), pp.e1-e48.

McOrist, A.L., Miller, R.B., Bird, A.R., Keogh, J.B., Noakes, M., Topping, D.L. and Conlon, M.A. 2011. Fecal butyrate levels vary widely among individuals but are usually increased by a diet high in resistant starch. *J Nutr.* **141**(5), pp.883-889.

Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., DuLong, C., Song, M., Nguyen, L.H., Mallick, H., Rimm, E.B., Izard, J., Huttenhower, C. and Chan, A.T. 2018. Stability of the human faecal microbiome in a cohort of adult men. *Nature Microbiology.* **3**(3), pp.347-355.

Metzendorf, N.G., Lange, L.M., Lainer, N., Schlüter, R., Dittmann, S., Paul, L.S., Troitzsch, D. and Sievers, S. 2022. Destination and Specific Impact of Different Bile Acids in the Intestinal Pathogen Clostridioides difficile. [Online]. Available from: https://doi.org/10.3389/fmicb.2022.814692

Miller, A.C., Arakkal, A.T., Sewell, D.K., Segre, A.M., Tholany, J., Polgreen, P.M. and Group, C.D.C.M.-H. 2023. Comparison of Different Antibiotics and the Risk for Community-Associated Clostridioides difficile Infection: A Case–Control Study. *Open Forum Infectious Diseases.* **10**(8), pofad413.

Miquel, S., Martín, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., Thomas, M., Wells, J.M. and Langella, P. 2013. Faecalibacterium prausnitzii and human intestinal health. *Current Opinion in Microbiology.* **16**(3), pp.255-261.

Miyamoto, J., Igarashi, M., Watanabe, K., Karaki, S.-i., Mukouyama, H., Kishino, S., Li, X., Ichimura, A., Irie, J., Sugimoto, Y., Mizutani, T., Sugawara, T., Miki, T., Ogawa, J., Drucker, D.J., Arita, M., Itoh, H. and Kimura, I. 2019. Gut microbiota confers host resistance to obesity by metabolizing dietary polyunsaturated fatty acids. *Nature Communications.* **10**(1), p4007.

Momose, Y., Hirayama, K. and Itoh, K. 2008. Competition for proline between indigenous Escherichia coli and E. coli O157:H7 in gnotobiotic mice associated with infant intestinal microbiota and its contribution to the colonization resistance against E. coli O157:H7. *Antonie Van Leeuwenhoek.* **94**(2), pp.165-171.

Moura, I.B., Grada, A., Spittal, W., Clark, E., Ewin, D., Altringham, J., Fumero, E., Wilcox, M.H. and Buckley, A.M. 2022. Profiling the Effects of Systemic Antibiotics for Acne, Including the Narrow-Spectrum Antibiotic Sarecycline, on the Human Gut Microbiota. *Front Microbiol.* **13**, p901911.

Moura, I.B., Normington, C., Ewin, D., Clark, E., Wilcox, M.H., Buckley, A.M. and Chilton, C.H. 2020. Method comparison for the direct enumeration of bacterial species using a chemostat model of the human colon. *BMC Microbiol.* **20**(1), p2.

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R. and Gordon, J.I. 2011. Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans. *Science.* **332**(6032), pp.970-974.

Muller, E., Algavi, Y.M. and Borenstein, E. 2022. The gut microbiome-metabolome dataset collection: a curated resource for integrative meta-analysis. *npj Biofilms and Microbiomes.* **8**(1), p79.

Mullish, B.H., McDonald, J.A.K., Pechlivanis, A., Allegretti, J.R., Kao, D., Barker, G.F., Kapila, D., Petrof, E.O., Joyce, S.A., Gahan, C.G.M., Glegola-Madejska, I., Williams, H.R.T., Holmes, E., Clarke, T.B., Thursz, M.R. and Marchesi, J.R. 2019. Microbial bile salt hydrolases mediate the efficacy of faecal microbiota transplant in the treatment of recurrent Clostridioides difficile infection. *Gut.* **68**(10), p1791.

Murata, K., Fukuda, Y., Simosaka, M., Watanabe, K., Saikusa, T. and Kimura, A. 1985. Metabolism of 2-oxoaldehyde in yeasts. *European Journal of Biochemistry.* **151**(3), pp.631-636.

Nel Van Zyl, K., Matukane, S.R., Hamman, B.L., Whitelaw, A.C. and Newton-Foot, M. 2022. Effect of antibiotics on the human microbiome: a systematic review. *International Journal of Antimicrobial Agents.* **59**(2), p106502.

Ng, K.M., Aranda-Díaz, A., Tropini, C., Frankel, M.R., Van Treuren, W., O'Loughlin, C.T., Merrill, B.D., Yu, F.B., Pruss, K.M., Oliveira, R.A., Higginbottom, S.K., Neff, N.F., Fischbach, M.A., Xavier, K.B., Sonnenburg, J.L. and Huang, K.C. 2019. Recovery of the Gut Microbiota after Antibiotics Depends on Host Diet, Community Context, and Environmental Reservoirs. *Cell host & microbe.* **26**(5), pp.650-665.e654.

Nirenberg, M.W. and Jakoby, W.B. 1960. Enzymatic utilization of gamma-hydroxybutyric acid. *J Biol Chem.* **235**, pp.954-960.

Noh, D.O. and Gilliland, S.E. 1993. Influence of bile on cellular integrity and beta-galactosidase activity of Lactobacillus acidophilus. *J Dairy Sci.* **76**(5), pp.1253-1259.

Normington, C., Moura, I.B., Bryant, J.A., Ewin, D.J., Clark, E.V., Kettle, M.J., Harris, H.C., Spittal, W., Davis, G., Henn, M.R., Ford, C.B., Wilcox, M.H. and Buckley, A.M. 2021. Biofilms harbour Clostridioides difficile, serving as a reservoir for recurrent infection. *npj Biofilms and Microbiomes.* **7**(1), p16.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C.M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D. and Pruitt, K.D. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**(D1), pp.D733-745.

O'Neill, J. 2014. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. Review on antimicrobial resistance. *Review on Antimicrobial Resistance, London, United Kingdom: https://amr-review. org/sites/default/files/AMR% 20Review% 20Paper.*

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J.-z., Abe, F. and Osawa, R. 2016. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiology.* **16**(1), p90.

Pearson, W.R. 2013. An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinformatics.* **Chapter 3**, pp.3.1.1-3.1.8.

Pensinger, D.A., Fisher, A.T., Dobrila, H.A., Van Treuren, W., Gardner, J.O., Higginbottom, S.K., Carter, M.M., Schumann, B., Bertozzi, C.R., Anikst, V., Martin, C., Robilotti, E.V., Chow, J.M., Buck, R.H., Tompkins, L.S., Sonnenburg, J.L. and Hryckowian, A.J. 2023. Butyrate Differentiates Permissiveness to Clostridioides difficile Infection and Influences Growth of Diverse C. difficile Isolates. *Infect Immun.* **91**(2), pe0057022.

Pitt, J.J. 2009. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev.* **30**(1), pp.19-34.

Prieto, A.I., Ramos-Morales, F. and Casadesús, J. 2004. Bile-induced DNA damage in Salmonella enterica. *Genetics.* **168**(4), pp.1787-1794.

Procházková, N., Laursen, M.F., La Barbera, G., Tsekitsidi, E., Jørgensen, M.S., Rasmussen, M.A., Raes, J., Licht, T.R., Dragsted, L.O. and Roager, H.M. 2024. Gut physiology and environment explain variations in human gut microbiome composition and metabolism. *Nature Microbiology.* **9**(12), pp.3210-3225.

Purcell, R.V., Visnovska, M., Biggs, P.J., Schmeier, S. and Frizelle, F.A. 2017. Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer. *Scientific reports.* **7**(1), pp.11590-11512.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Meta, H.I.T.C., Bork, P., Ehrlich, S.D. and Wang, J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature.* **464**(7285), pp.59-65.

Qureshi, S., Maria, N., Zeeshan, M., Irfan, S. and Qamar, F.N. 2021. Prevalence and risk factors associated with multi-drug resistant organisms (MDRO) carriage among pediatric patients at the time of admission in a tertiary care hospital of a developing country. A cross-sectional study. *BMC Infectious Diseases.* **21**(1), p547.

Rea, M.C., Dobson, A., O'Sullivan, O., Crispie, F., Fouhy, F., Cotter, P.D., Shanahan, F., Kiely, B., Hill, C. and Ross, R.P. 2011. Effect of broad- and narrow-spectrum antimicrobials on Clostridium difficile and microbial diversity in a model of the distal colon. *Proc Natl Acad Sci U S A.* **108 Suppl 1**(Suppl 1), pp.4639-4644.

Reichardt, N., Duncan, S.H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K.P., Flint, H.J. and Louis, P. 2014. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *Isme j.* **8**(6), pp.1323-1335.

Reinhold, D., Pielke-Lombardo, H., Jacobson, S., Ghosh, D. and Kechris, K. 2019. Pre-analytic Considerations for Mass Spectrometry-Based Untargeted Metabolomics Data. *Methods Mol Biol.* **1978**, pp.323-340.

Reisdorph, N.A., Walmsley, S. and Reisdorph, R. 2019. A Perspective and Framework for Developing Sample Type Specific Databases for LC/MS-Based Clinical Metabolomics. *Metabolites.* **10**(1).

Ríos-Covián, D., Ruas-Madiedo, P., Margolles, A., Gueimonde, M., de los Reyes-Gavilán, C.G. and Salazar, N. 2016. Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health. *Frontiers in Microbiology.* **7**.

Roberts, A.K., Harris, H.C., Smith, M., Giles, J., Polak, O., Buckley, A.M., Clark, E., Ewin, D., Moura, I.B., Spitall, W., Shone, C.C., Wilcox, M., Chilton, C. and Donev, R. 2020. A Novel, Orally Delivered Antibody Therapy and Its Potential to Prevent Clostridioides difficile Infection in Pre-clinical Models. *Front Microbiol.* **11**, p578903.

Roediger, W.E. 1980. Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man. *Gut.* **21**(9), p793.

Rooney, C.M., Sheppard, A.E., Clark, E., Davies, K., Hubbard, A.T.M., Sebra, R., Crook, D.W., Walker, A.S., Wilcox, M.H. and Chilton, C.H. 2019. Dissemination of multiple carbapenem resistance genes in an in vitro gut model simulating the human colon. *J Antimicrob Chemother.* **74**(7), pp.1876-1883.

Roupar, D., Berni, P., Martins, J.T., Caetano, A.C., Teixeira, J.A. and Nobre, C. 2021. Bioengineering approaches to simulate human colon microbiome ecosystem. *Trends in Food Science & Technology.* **112**, pp.808-822.

Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B., Chou, S., Gonen, T., Goodlett, D.R., Goodman, A.L. and Mougous, J.D. 2014. A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host Microbe.* **16**(2), pp.227-236.

Sadygov, R.G., Maroto, F.M. and Hühmer, A.F. 2006. ChromAlign: A two-step algorithmic procedure for time alignment of three-dimensional LC-MS chromatographic surfaces. *Anal Chem.* **78**(24), pp.8207-8217.

Saha, S., Day-Walsh, P., Shehata, E. and Kroon, P.A. 2021. Development and Validation of a LC-MS/MS Technique for the Analysis of Short Chain Fatty Acids in Tissues and Biological Fluids without Derivatisation Using Isotope Labelled Internal Standards. *Molecules.* [Online]. **26**(21). Available from: https://doi.org/10.3390/molecules26216444

Sakai, A., Katayama, K., Katsuragi, T. and Tani, Y. 2001. Glycolaldehyde-forming route in Bacillus subtilis in relation to vitamin B6 biosynthesis. *Journal of Bioscience and Bioengineering.* **91**(2), pp.147-152.

Santaus, T.M., Li, S., Saha, L., Chen, W.H., Bhagat, S., Stine, O.C. and Geddes, C.D. 2019. A comparison of Lyse-It to other cellular sample preparation, bacterial lysing, and DNA fragmentation technologies. *PLOS ONE.* **14**(7), pe0220102.

Sanyal, A.J., Shiffman, M.L., Hirsch, J.I. and Moore, E.W. 1991. Premicellar taurocholate enhances ferrous iron uptake from all regions of rat small intestine. *Gastroenterology.* **101**(2), pp.382-389.

Sassone-Corsi, M., Nuccio, S.-P., Liu, H., Hernandez, D., Vu, C.T., Takahashi, A.A., Edwards, R.A. and Raffatellu, M. 2016. Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature.* **540**(7632), pp.280-283.

Sato, Y., Atarashi, K., Plichta, D.R., Arai, Y., Sasajima, S., Kearney, S.M., Suda, W., Takeshita, K., Sasaki, T., Okamoto, S., Skelly, A.N., Okamura, Y., Vlamakis, H., Li, Y., Tanoue, T., Takei, H., Nittono, H., Narushima, S., Irie, J., Itoh, H., Moriya, K., Sugiura, Y., Suematsu, M., Moritoki, N., Shibata, S., Littman, D.R., Fischbach, M.A., Uwamino, Y., Inoue, T., Honda, A., Hattori, M., Murai, T., Xavier, R.J., Hirose, N. and Honda, K. 2021. Novel bile acid biosynthetic pathways are enriched in the microbiome of centenarians. *Nature.* **599**(7885), pp.458-464.

Saxton, K., Baines, S.D., Freeman, J., O'Connor, R. and Wilcox, M.H. 2009. Effects of exposure of Clostridium difficile PCR ribotypes 027 and 001 to fluoroquinolones in a human gut model. *Antimicrob Agents Chemother.* **53**(2), pp.412-420.

Schouten, M.A., Voss, A. and Hoogkamp-Korstanje, J.A.A. 1999. Antimicrobial Susceptibility Patterns of Enterococci Causing Infections in Europe. *Antimicrobial Agents and Chemotherapy.* **43**(10), pp.2542-2546.

Schrimpe-Rutledge, A.C., Codreanu, S.G., Sherrod, S.D. and McLean, J.A. 2016. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J Am Soc Mass Spectrom.* **27**(12), pp.1897-1905.

Scott, K.P., Martin, J.C., Duncan, S.H. and Flint, H.J. 2014. Prebiotic stimulation of human colonic butyrate-producing bacteria and bifidobacteria, in vitro. *FEMS Microbiol Ecol.* **87**(1), pp.30-40.

Sender, R., Fuchs, S. and Milo, R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology.* **14**(8), pe1002533.

Shogbesan, O., Poudel, D.R., Victor, S., Jehangir, A., Fadahunsi, O., Shogbesan, G. and Donato, A. 2018. A Systematic Review of the Efficacy and Safety of Fecal Microbiota Transplant for Clostridium difficile Infection in Immunocompromised Patients. *Can J Gastroenterol Hepatol.* **2018**, p1394379.

Shono, Y. and van den Brink, M.R.M. 2017. Empiric antibiotic use in allogeneic hematopoietic cell transplantation: should we avoid anaerobe coverage? *Blood Advances.* **1**(25), pp.2325-2328.

Simopoulos, A.P. 2016. An Increase in the Omega-6/Omega-3 Fatty Acid Ratio Increases the Risk for Obesity. *Nutrients.* **8**(3), p128.

Singh, R.K., Chang, H.-W., Yan, D., Lee, K.M., Ucmak, D., Wong, K., Abrouk, M., Farahnik, B., Nakamura, M., Zhu, T.H., Bhutani, T. and Liao, W. 2017. Influence of diet on the gut microbiome and implications for human health. *Journal of Translational Medicine.* **15**(1), p73.

Singhal, R., Donde, H., Ghare, S., Stocke, K., Zhang, J., Vadhanam, M., Reddy, S., Gobejishvili, L., Chilton, P., Joshi-Barve, S., Feng, W., McClain, C., Hoffman, K., Petrosino, J., Vital, M. and Barve, S. 2021. Decrease in acetyl-CoA pathway utilizing butyrate-producing bacteria is a key pathogenic feature of alcohol-induced functional gut microbial dysbiosis and development of liver disease in mice. *Gut Microbes.* **13**(1), p1946367.

Slimings, C. and Riley, T.V. 2014. Antibiotics and hospital-acquired Clostridium difficile infection: update of systematic review and meta-analysis. *Journal of Antimicrobial Chemotherapy.* **69**(4), pp.881-891.

Smith, L., Villaret-Cazadamont, J., Claus, S.P., Canlet, C., Guillou, H., Cabaton, N.J. and Ellero-Simatos, S. 2020. Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites.* **10**(3).

So, D., Whelan, K., Rossi, M., Morrison, M., Holtmann, G., Kelly, J.T., Shanahan, E.R., Staudacher, H.M. and Campbell, K.L. 2018. Dietary fiber intervention on gut microbiota composition in healthy adults: a systematic review and meta-analysis. *Am J Clin Nutr.* **107**(6), pp.965-983.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L.G., Gratadoux, J.-J., Blugeon, S., Bridonneau, C., Furet, J.-P., Corthier, G., Grangette, C., Vasquez, N., Pochart, P., Trugnan, G., Thomas, G., Blottière, H.M., Doré, J., Marteau, P., Seksik, P. and Langella, P. 2008. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences.* **105**(43), pp.16731-16736.

Sonnenborn, U. 2016. Escherichia coli strain Nissle 1917—from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties. *FEMS Microbiology Letters.* **363**(19), pfnw212.

Sonowal, R., Swimm, A., Sahoo, A., Luo, L., Matsunaga, Y., Wu, Z., Bhingarde, J.A., Ejzak, E.A., Ranawade, A., Qadota, H., Powell, D.N., Capaldo, C.T., Flacker, J.M., Jones, R.M., Benian, G.M. and Kalman, D. 2017. Indoles from commensal bacteria extend healthspan. *Proceedings of the National Academy of Sciences.* **114**(36), pp.E7506-E7515.

Sorbara, M.T., Dubin, K., Littmann, E.R., Moody, T.U., Fontana, E., Seok, R., Leiner, I.M., Taur, Y., Peled, J.U., van den Brink, M.R.M., Litvak, Y., Bäumler,

A.J., Chaubard, J.L., Pickard, A.J., Cross, J.R. and Pamer, E.G. 2019. Inhibiting antibiotic-resistant Enterobacteriaceae by microbiota-mediated intracellular acidification. *J Exp Med.* **216**(1), pp.84-98.

Sorg, J.A. and Sonenshein, A.L. 2010. Inhibiting the initiation of Clostridium difficile spore germination using analogs of chenodeoxycholic acid, a bile acid. *J Bacteriol.* **192**(19), pp.4983-4990.

Sorg Joseph, A. and Sonenshein Abraham, L. 2008. Bile Salts and Glycine as Cogerminants for Clostridium difficile Spores. *Journal of Bacteriology.* **190**(7), pp.2505-2512.

Souihi, A., Mohai, M.P., Martin, J.W. and Kruve, A. 2023. Mobile phase and column chemistry selection for high sensitivity non-targeted LC/ESI/HRMS screening of water. *Analytica Chimica Acta.* **1274**, p341573.

Starke, R., Jehmlich, N., Alfaro, T., Dohnalkova, A., Capek, P., Bell, S.L. and Hofmockel, K.S. 2019. Incomplete cell disruption of resistant microbes. *Scientific Reports.* **9**(1), p5618.

Stevens, V., Dumyati, G., Fine, L.S., Fisher, S.G. and van Wijngaarden, E. 2011. Cumulative antibiotic exposures over time and the risk of Clostridium difficile infection. *Clin Infect Dis.* **53**(1), pp.42-48.

Stevens, V., Dumyati, G., Fine, L.S., Fisher, S.G. and van Wijngaarden, E. 2011. Cumulative Antibiotic Exposures Over Time and the Risk of Clostridium difficile Infection. *Clinical Infectious Diseases.* **53**(1), pp.42-48.

Strimbu, K. and Tavel, J.A. 2010. What are biomarkers? *Curr Opin HIV AIDS.* **5**(6), pp.463-466.

Sulis, G., Sayood, S., Katukoori, S., Bollam, N., George, I., Yaeger, L.H., Chavez, M.A., Tetteh, E., Yarrabelli, S., Pulcini, C., Harbarth, S., Mertz, D., Sharland, M., Moja, L., Huttner, B. and Gandra, S. 2022. Exposure to World Health Organization's AWaRe antibiotics and isolation of multidrug resistant bacteria: a systematic review and meta-analysis. *Clinical Microbiology and Infection.* **28**(9), pp.1193-1202.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J. and Viant, M.R. 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics.* **3**(3), pp.211-221.

Sun, L., Zhang, X., Zhang, Y., Zheng, K., Xiang, Q., Chen, N., Chen, Z., Zhang, N., Zhu, J. and He, Q. 2019. Antibiotic-Induced Disruption of Gut Microbiota Alters Local Metabolomes and Immune Responses. *Frontiers in Cellular and Infection Microbiology.* **9**.

Tailford, L.E., Crost, E.H., Kavanaugh, D. and Juge, N. 2015. Mucin glycan foraging in the human gut microbiome. *Front Genet.* **6**, p81.

Talasniemi, J.P., Pennanen, S., Savolainen, H., Niskanen, L. and Liesivuori, J. 2008. Analytical investigation: assay of D-lactate in diabetic plasma and urine. *Clin Biochem.* **41**(13), pp.1099-1103.

Taur, Y., Coyte, K., Schluter, J., Robilotti, E., Figueroa, C., Gjonbalaj, M., Littmann, E.R., Ling, L., Miller, L., Gyaltshen, Y., Fontana, E., Morjaria, S., Gyurkocza, B., Perales, M.-A., Castro-Malaspina, H., Tamari, R., Ponce, D., Koehne, G., Barker, J., Jakubowski, A., Papadopoulos, E., Dahi, P., Sauter, C., Shaffer, B., Young, J.W., Peled, J., Meagher, R.C., Jenq, R.R., van den Brink, M.R.M., Giralt, S.A., Pamer, E.G. and Xavier, J.B. 2018. Reconstitution of the gut microbiota of antibiotic-treated patients by autologous fecal microbiota transplant. *Science Translational Medicine.* **10**(460), peaap9489.

Taur, Y., Jenq, R.R., Perales, M.-A., Littmann, E.R., Morjaria, S., Ling, L., No, D., Gobourne, A., Viale, A., Dahi, P.B., Ponce, D.M., Barker, J.N., Giralt, S., van den Brink, M. and Pamer, E.G. 2014. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood.* **124**(7), pp.1174-1182.

Thananimit, S., Pahumunto, N. and Teanpaisan, R. 2022. Characterization of Short Chain Fatty Acids Produced by Selected Potential Probiotic Lactobacillus Strains. *Biomolecules.* [Online]. **12**(12). Available from: https://doi.org/10.3390/biom12121829

Thanissery, R., Winston, J.A. and Theriot, C.M. 2017. Inhibition of spore germination, growth, and toxin activity of clinically relevant C. difficile strains by gut microbiota derived secondary bile acids. *Anaerobe.* **45**, pp.86-100.

Theriot, C.M., Koenigsknecht, M.J., Carlson, P.E., Hatton, G.E., Nelson, A.M., Li, B., Huffnagle, G.B., Z. Li, J. and Young, V.B. 2014. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nature Communications.* **5**(1), p3114.

Tigchelaar, E.F., Bonder, M.J., Jankipersadsing, S.A., Fu, J., Wijmenga, C. and Zhernakova, A. 2016. Gut microbiota composition associated with stool consistency. *Gut.* **65**(3), pp.540-542.

Toto, F., Marangelo, C., Scanu, M., De Angelis, P., Isoldi, S., Abreu, M.T., Cucchiara, S., Stronati, L., Del Chierico, F. and Putignani, L. 2024. A Novel Microbial Dysbiosis Index and Intestinal Microbiota-Associated Markers as Tools of Precision Medicine in Inflammatory Bowel Disease Paediatric Patients. *International Journal of Molecular Sciences.* [Online]. **25**(17). Available from: https://doi.org/10.3390/ijms25179618

Tresman, R. and Goldenberg, S.D. 2018. Healthcare resource use and attributable cost of Clostridium difficile infection: a micro-costing analysis comparing first and recurrent episodes. *Journal of Antimicrobial Chemotherapy.* **73**(10), pp.2851-2855.

Tsukuda, N., Yahagi, K., Hara, T., Watanabe, Y., Matsumoto, H., Mori, H., Higashi, K., Tsuji, H., Matsumoto, S., Kurokawa, K. and Matsuki, T. 2021. Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. *Isme j.* **15**(9), pp.2574-2590.

Human Tissue Act 2004 2004. https://www.legislation.gov.uk/:

UKHSA. 2024. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR).*

Urdaneta, V. and Casadesús, J. 2017. Interactions between Bacteria and Bile Salts in the Gastrointestinal and Hepatobiliary Tracts. *Frontiers in medicine.* **4**, pp.163-163.

Usta-Gorgun, B. and Yilmaz-Ersan, L. 2020. Short-chain fatty acids production by Bifidobacterium species in the presence of salep. *Electronic Journal of Biotechnology.* **47**, pp.29-35.

Valo, E., Colombo, M., Sandholm, N., McGurnaghan, S.J., Blackbourn, L.A.K., Dunger, D.B., McKeigue, P.M., Forsblom, C., Groop, P.-H., Colhoun, H.M., Turner, C. and Dalton, R.N. 2022. Effect of serum sample storage temperature on metabolomic and proteomic biomarkers. *Scientific Reports.* **12**(1), p4571.

van Burik, J.A.H., Schreckhise, R.W., White, T.C., Bowden, R.A. and Myerson, D. 1998. Comparison of six extraction techniques for isolation of DNA from filamentous fungi. *Medical Mycology.* **36**(5), pp.299-303.

Van de Wiele, T., Van den Abbeele, P., Ossieur, W., Possemiers, S. and Marzorati, M. 2015. The Simulator of the Human Intestinal Microbial Ecosystem (SHIME(®)). In: Verhoeckx, K., Cotter, P., López-Expósito, I., Kleiveland, C., Lea, T., Mackie, A., Requena, T., Swiatecka, D. and Wichers, H. eds. *The Impact of Food Bioactives on Health: in vitro and ex vivo models.* Cham (CH): Springer

Copyright 2015, The Author(s). pp.305-317.

Venturini, C., Bowring, B., Fajardo-Lubian, A., Devine, C. and Iredell, J. 2021. Effects of Antibiotic Treatment with Piperacillin/Tazobactam versus Ceftriaxone on the Composition of the Murine Gut Microbiota. *Antimicrob Agents Chemother.* **65**(2).

Vincent, J.-L., Bassetti, M., François, B., Karam, G., Chastre, J., Torres, A., Roberts, J.A., Taccone, F.S., Rello, J., Calandra, T., De Backer, D., Welte, T. and Antonelli, M. 2016. Advances in antibiotic therapy in the critically ill. *Critical Care.* **20**(1), p133.

Visekruna, A. and Luu, M. 2021. The Role of Short-Chain Fatty Acids and Bile Acids in Intestinal and Liver Function, Inflammation, and Carcinogenesis. *Frontiers in Cell and Developmental Biology.* **9**.

Vital, M., Howe Adina, C. and Tiedje James, M. 2014. Revealing the Bacterial Butyrate Synthesis Pathways by Analyzing (Meta)genomic Data. *mBio.* **5**(2), pp.10.1128/mbio.00889-00814.

Vrieze, A., Van Nood, E., Holleman, F., Salojärvi, J., Kootte, R.S., Bartelsman, J.F., Dallinga-Thie, G.M., Ackermans, M.T., Serlie, M.J., Oozeer, R., Derrien, M., Druesne, A., Van Hylckama Vlieg, J.E., Bloks, V.W., Groen, A.K., Heilig, H.G., Zoetendal, E.G., Stroes, E.S., de Vos, W.M., Hoekstra, J.B. and Nieuwdorp, M. 2012. Transfer of intestinal microbiota from lean donors

increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology.* **143**(4), pp.913-916.e917.

Wang, J., Tan, G., Liu, X. and Huang, Y. 2015. Expanding the Coverage of Metabolome Using Multiple Liquid Chromatography Modes. In: *ASMS 2015*.

Wang, R., Fu, Y. and Lai, L. 1997. A New Atom-Additive Method for Calculating Partition Coefficients. *Journal of Chemical Information and Computer Sciences.* **37**(3), pp.615-621.

Wang, X., Heazlewood, S.P., Krause, D.O. and Florin, T.H. 2003. Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. *J Appl Microbiol.* **95**(3), pp.508-520.

Wang, Z., Zolnik, C.P., Qiu, Y., Usyk, M., Wang, T., Strickler, H.D., Isasi, C.R., Kaplan, R.C., Kurland, I.J., Qi, Q. and Burk, R.D. 2018. Comparison of Fecal Collection Methods for Microbiome and Metabolomics Studies. *Frontiers in Cellular and Infection Microbiology.* **8**.

Ward, J.B.J., Lajczak, N.K., Kelly, O.B., O'Dwyer, A.M., Giddam, A.K., Ní Gabhann, J., Franco, P., Tambuwala, M.M., Jefferies, C.A., Keely, S., Roda, A. and Keely, S.J. 2017. Ursodeoxycholic acid and lithocholic acid exert anti-inflammatory actions in the colon. *American Journal of Physiology-Gastrointestinal and Liver Physiology.* **312**(6), pp.G550-G558.

Watanabe, Y., Fujisaka, S., Ikeda, K., Ishikawa, M., Yamada, T., Nawaz, A., Kado, T., Kuwano, T., Nishimura, A., Bilal, M., Liu, J., Yagi, K., Hase, K. and Tobe, K. 2021. Gut microbiota, determined by dietary nutrients, drive modification of the plasma lipid profile and insulin resistance. *iScience.* **24**(5).

Weng, C.-Y.C., Suarez, C., Cheang, S.E., Couture, G., Goodson, M.L., Barboza, M., Kalanetra, K.M., Masarweh, C.F., Mills, D.A., Raybould, H.E. and Lebrilla, C.B. 2024. Quantifying Gut Microbial Short-Chain Fatty Acids and Their Isotopomers in Mechanistic Studies Using a Rapid, Readily Expandable LC–MS Platform. *Analytical Chemistry.* **96**(6), pp.2415-2424.

Wilson, K.H., Kennedy, M.J. and Fekety, F.R. 1982. Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. *Journal of Clinical Microbiology.* **15**(3), p443.

Wishart, D.S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B.L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V.W., Varshavi, D., Varshavi, D., Allen, D., Arndt, D., Khetarpal, N., Sivakumaran, A., Harford, K., Sanford, S., Yee, K., Cao, X., Budinski, Z., Liigand, J., Zhang, L., Zheng, J., Mandal, R., Karu, N., Dambrova, M., Schiöth, H.B., Greiner, R. and Gautam, V. 2022. HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res.* **50**(D1), pp.D622-d631.

Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M.A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P.,

Amegbey, G., Block, D., Hau, D.D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G.E., Macinnis, G.D., Weljie, A.M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B.D., Vogel, H.J. and Querengesser, L. 2007. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **35**(Database issue), pp.D521-526.

Wu, S., Yoon, S., Zhang, Y.-G., Lu, R., Xia, Y., Wan, J., Petrof, E.O., Claud, E.C., Chen, D. and Sun, J. 2015. Vitamin D receptor pathway is required for probiotic protection in colitis. *American Journal of Physiology-Gastrointestinal and Liver Physiology.* **309**(5), pp.G341-G349.

Xu, C., Zhu, H. and Qiu, P. 2019. Aging progression of human gut microbiota. *BMC Microbiology.* **19**(1), p236.

Yachida, S., Mizutani, S., Shiroma, H., Shiba, S., Nakajima, T., Sakamoto, T., Watanabe, H., Masuda, K., Nishimoto, Y., Kubo, M., Hosoda, F., Rokutan, H., Matsumoto, M., Takamaru, H., Yamada, M., Matsuda, T., Iwasaki, M., Yamaji, T., Yachida, T., Soga, T., Kurokawa, K., Toyoda, A., Ogura, Y., Hayashi, T., Hatakeyama, M., Nakagama, H., Saito, Y., Fukuda, S., Shibata, T. and Yamada, T. 2019. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nature Medicine.* **25**(6), pp.968-976.

Yamamoto, M., Yamaguchi, R., Munakata, K., Takashima, K., Nishiyama, M., Hioki, K., Ohnishi, Y., Nagasaki, M., Imoto, S., Miyano, S., Ishige, A. and Watanabe, K. 2012. A microarray analysis of gnotobiotic mice indicating that microbial exposure during the neonatal period plays an essential role in immune system development. *BMC genomics.* **13**(1), pp.335-335.

Yang, Q., Zhang, A.-h., Miao, J.-h., Sun, H., Han, Y., Yan, G.-l., Wu, F.-f. and Wang, X.-j. 2019. Metabolomics biotechnology, applications, and future trends: a systematic review. *RSC Advances.* **9**(64), pp.37245-37257.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R. and Gordon, J.I. 2012. Human gut microbiome viewed across age and geography. *Nature.* **486**(7402), pp.222-227.

Yu, M., Jia, H., Zhou, C., Yang, Y., Zhao, Y., Yang, M. and Zou, Z. 2017. Variations in gut microbiota and fecal metabolic phenotype associated with depression by 16S rRNA gene sequencing and LC/MS-based metabolomics. *Journal of Pharmaceutical and Biomedical Analysis.* **138**, pp.231-239.

Ze, X., Duncan, S.H., Louis, P. and Flint, H.J. 2012. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. *The ISME Journal.* **6**(8), pp.1535-1543.

Zhang, J., Kobert, K., Flouri, T. and Stamatakis, A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics.* **30**(5), pp.614-620.

Zhu, S., Li, H., Liang, J., Lv, C., Zhao, K., Niu, M., Li, Z., Zeng, L. and Xu, K. 2020. Assessment of oral ciprofloxacin impaired gut barrier integrity on gut bacteria in mice. *International Immunopharmacology.* **83**, p106460.

Zoued, A., Durand, E., Bebeacua, C., Brunet, Y.R., Douzi, B., Cambillau, C., Cascales, E. and Journet, L. 2013. TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. *The Journal of biological chemistry.* **288**(38), pp.27031-27041.

Appendix A

Supplementary Material: Metabolomic Methods Development

A.1 Equipment

Table A-1 Equipment used for LCMS Workflow

Equipment	Product Code/ID	Supplier
Mass Spectrometer	Exploris 240	Thermo Fisher
HPLC / Autosampler System	Vanquish	Thermo Fisher
Refrigerated Centrifuge	4-16KS	Sigma
Vacuum Evaporator	Concentrator Plus	Eppendorf
Sonicating Water bath	USC-T	VWR
Bead Beater	Fast-Prep 24 5g	MP Bio

Table A-2 LCMS Consumables

Equipment	Product Code/ID	Supplier
Microcentrifuge Tubes 1.5mL / 2mL	0030123328/ 0030123344	Eppendorf
HPLC Sample Vial	186002802	Waters

Table A-3 Fridges and Freezers

Equipment	Product Code/ID	Supplier
-80 Freezer	U410 HEF	New Brunswick
-20 Freezer	RLV1825	Labcold

A.2 Mass Spectrometry Metabolite Library of Standards

Mass Spectrometry Metabolite Library of Standards (MSMLS) were purchased from IROA Technologies

Table A-4 MSMLS Used in Validation and Library Generation

Product	Product Code/ID	Lot Number
Microbiome Metabolite Library of Standards	GUTMLS	230-08
Bile Acid/Carnitine/Sterol Metabolite Library of Standards	BACSMLS	205-36 2

Appendix B

Supplementary Material: In-Vitro Gut Models

B.1 Bacterial Culture Media

The Media used for the culture and enumeration of bacterial species was a combination of in-house produced from base agar powders and supplier-poured media.

B.1.1 Supplier Poured Media

Table B-1 Supplier Poured Media Used for Bacterial Enumeration

Media	Product Code/ID	Supplier
Nutrient Agar	PO0155A	Oxoid
MacConkey Agar with Salt	PO0149A	Oxoid
Fastidious Anaerobe Agar (F.A.A.) with Horse Blood	PB0225A	Oxoid
CHROMID CARBA SMART	41468520	Biomerieux

B.1.2 In-house Media Used for Bacterial Enumeration

B.1.2.1 Formula for In-house Media

Table B-2 Braziers CCEY + Lysozyme agar (CCEYL)

Constituent	g/L (Unless stated)
Clostridium Difficile Agar Base (Brazier's)	48
Cycloserine, and Cefoxitin Supplement X093	2 vials
Defibrinated Horse Blood[1]	20mL
20% Saponin	500µL
Lysozyme	5mg

[1]Horse blood lysed with 20% Saponin prior to addition

Table B-3 Braziers CCEY + Lysozyme, moxifloxacin, amphotericin B and colistin agar (MXF)

Constituent	g/L (Unless stated)
Clostridium Difficile Agar Base (Brazier's)	48
Cycl`oserine, and Cefoxitin Supplement X093	2 vials
Defibrinated Horse Blood[2]	20mL
20% Saponin	500µL
Lysozyme	5mg
Moxifloxacin	2mg/L
Amphotericin B	32mL
Colistin	10mg/L

[2]Horse blood lysed with 20% Saponin prior to addition

Table B-4 Bacteroides Bile Aesculin agar (BBE)

Constituent	g/L (Unless stated)
Clostridium Difficile Agar Base (Brazier's)	4.5
Haemin	20mL
Vitamin K1	20µL
Kanamycin	75 mg/L
Vancomycin	7.5 mg/L
Penicillin G	1 mg/L
Colistin	10 mg/L

Table B-5 Nutrient agar:

Constituent	g/L (Unless stated)
Nutrient agar base	28
Agar (Bacteriological)	5

Table B-6 Kanamycin Aesculin Azide, Nalidixic Acid, Aztreonam agar (KAA)

Constituent	g/L (Unless stated)
Kanamycin Aesculin Azide agar base	42.6
Agar Bacteriological	5
Kanamycin	20 mg/L
Aztreonam	10 mg/L
Nalidixic Acid	10 mg/L

Lincomycin	1 mg/L

Table B-7 Beerens agar

Constituent	g/L (Unless stated)
Columbia agar base	42.5
Glucose	5
Agar (Bacteriological)	5
L-cysteine HCL	0.5

Adjust pH to 5.0 with 10n NaOH and 10n HCL

Table B-8 LAMVAB

Constituent	g/L (Unless stated)
MRS Broth	42.5
L-cysteine HCL	0.5
Agar (Bacteriological)	20
Vancomycin	20 mg/L

B.1.2.2 Agar Base Powders

Table B-9 Agar Powders used in In-House Media

Media Base	Product Code/ID	Supplier
Kanamycin Aesculin Azide (KAA)	CM0591	Oxoid
MRS Broth	CM0359	Oxoid
Bacteroides Bile Esculin agar base	CM0888	Oxoid
Colombia agar base	CM0331B	Oxoid
Agar (Bacteriological)	LP0011B	Oxoid
Braziers CCEY Agar	7000004515	Neogen

B.1.2.3 Media Supplements & Antibiotics

Table B-10 Supplements Used in the Creation of In-house Media[1]

Supplements	Product Code/ID	Supplier
Arabinogalactan	Sc-210883	ChemCruz
Bile Salts	B8756-100G	Sigma-Aldrich
Calcium chloride	C5670-100g	Sigma-Aldrich
Chenodeoxycholic acid	218990	Fluorochem
Cysteine HCL	F093727	Fluorochem
D-Glucose	G/0450/53	Fisher Chemical
Di-potassium monohydrogen phosphate	795496-500G	Sigma
Haemin (porcine)	H9039-100G	Oxoid
Horse Blood (defibrinated)	HB035	TCS
Lithocholic acid	L6250-10G	Merck
Lysozyme	L6876-10 G	Sigma-Aldrich
Magnesium Sulphate	M5921-500G	Sigma-Aldrich
Mucin	M2378-500G	Sigma-Aldrich
PBS Tablets	18912-014	Gibco
Pectin	416860010	Thermo Scientific
Peptone Water	CM009B	Oxoid
Potassium dihydrogen phosphate	P0662-500G	Sigma-Aldrich
Propanoic Acid	81910-L	Sigma-Aldrich
Saponin	s7900-100g	Sigma-Aldrich
Sodium chloride	S/3160/60	Fischer Chemical
Sodium hydrogen carbonate	102399802	Sigma-Aldrich
Starch	S/7960/53	Fisher Chemical
Tween 80	P1754-500ML	Sigma-Aldrich
Vitamin K1	10575.06	Thermo Scientific
Yeast extract	LP0021B	Oxoid

[1]Supplements in solid agar for bacterial Isolation and Enumeration and In-*Vitro* Gut Model Growth media

B.1.3 Antimicrobials

Table B-11 Antimicrobials Used for In-vitro Study

Antibiotic	Product Code/ID	Supplier
Amoxicillin	A8523-1g	Sigma-Aldrich
Amphotericin B	A2942-100ML	Sigma-Aldrich
Aztreonam		
Ciprofloxacin	C031-25g	Toku-E
Colistin sulphate	BIC0118	Apollo scientific
Cycloserine/Cefoxitin	X093	Neogen
Kanamycin	11815-024	Apollo Scientific
Lincomycin	J61251.03	Alpha Aesar
Moxifloxacin	PHR1542-1g	Sigma-Aldrich
Nalidixic acid	BIN0134	Apollo Scientific
Piperacillin-Tazobactam	TS/DRUGS/07/2016	Aurobindo
Vancomycin	BIV0155	Apollo Scientific

B.2 Strains & Cell Lines

B.2.1 Bacterial Strains

Bacterial Strains used for Media QC and Cabinet QC; all strains purchased from UK HAS Culture collections

Table B-12 Reference Bacterial Strains with Catalogue Numbers

Strain	Catalogue No	Strain
Pseudomonas aeruginosa	NCTC 12903	Reference Strain
Clostridioides difficile	NCTC 13366	Reference Strain
Staphylococcus aureus	NCTC 12973	Reference Strain
Kocuria rhizophila	NCTC 8340	Reference Strain
Escherichia coli	NCTC 12241	Reference Strain
Bacteroides fragilis	NCTC 9343	Type Strain
Lactobacillus rhamnosus	NCTC 12953	Type Strain
Bifidobacterium longum	NCTC 11818	Type Strain
Clostridium perfringens	NCTC 8237	Type Strain
Enterococcus faecalis	NCTC 12697	Reference Strain

B.2.2 Eukaryotic Cell lines

Table B-13 Eukaryotic Cell Lines Used for Cell Cytotoxicity Neutralisation Assays.

Cell Line	Catalogue No	Supplier
Vero (African Green Monkey kidney cells)	84113001	Merck

B.3 Equipment

Table B-14 Anaerobic Workstation

Equipment	Product Code/ID	Supplier
Whitley A95 Workstation	A95	Don Whitley Scientific
Nitrogen (Oxygen Free)	44-x	BOC
10% Carbon Dioxide, 5% Hydrogen/Nitrogen (Anaerobic) Cylinder	290563-L	BOC

Table B-15 Aerobic Culture

Equipment	Product Code/ID	Supplier
37°C Incubator	Heratherm	Thermo

Table B-16 CO_2 Culture

Equipment	Product ID	Supplier
CO_2 Incubator	MCO 5 AC-PE	Panasonic

Table B-17 Balances

Equipment	Product Code/ID	Supplier
Analytical Balance	ac210p	Sartorius
Precision Balance	PNS 300-2	Kern

Table B-18 Autoclave & Sterilisation Cycle

Equipment	Product Code/ID	Supplier
Benchtop Autoclave		Priorclave

Equipment Sterilisation cycle: Carried out at 121ºC for 15 minutes with a temperature probe placed in the centre of load

Media Sterilisation cycle: Carried out at 123ºC for 26 minutes with a temperature probe placed in liquid balance of media volume

Table B-19 In-Vitro Models (Leeds-HGM & MiGut)

Equipment	Product Code/ID	Supplier
Fermenter Controller	Microlab 0 pH control unit	Brighton-systems
Glassware	NA	Soham Scientific
MiGut	NA	William A Davis Birch
Nitrogen generator	NG5-1	SLS Lab
Peristaltic pump	MasterFlex ls Easyload3	MasterFlex MasterFlex
Tubing	Marprene #16 Tygon adk00007	Watson Marlow Saint Gobain
Water Bath (Circulating)	tc120	Grant

Table B-20 Other Equipment

Equipment	Product Code/ID	Supplier
Inverted Microscope	DM IL	Leica

Appendix C

Supplementary Material: Metagenomic Data and Analysis for In-vitro Intestinal Microbiome Investigation

3.1 NGS Data Analysis and QC

Sequences were supplied as FASTQ files and downloaded from the ARC3 platform. After retrieval, file integrity was verified using md5 checksum.

3.1.1 Raw Sequence General Statistics

Pre-QC review of raw NGS FASTQ files shows an average read length of 151 base pairs across all samples. For MiGut samples, the average sequence count of 24million, with a duplication rate of 50%. Similarly, Leeds HGM samples possessed an average read length of 27 million sequences per file and a duplication rate of 45%.

Table C-1 General Statistics for NGS data for miGUT and Leeds

	Reactor Stage	Phase	Read 1			Read 2		
			% Dups	% GC	Million Seqs	% Dups	% GC	Million Seqs
MiGut	Proximal	SS	49.00%	53%	26.1	47.90%	53%	26.1
	Distal	SS	33.30%	54%	26.9	32.80%	54%	26.9
	Proximal	AMX	53.00%	46%	21.7	51.80%	46%	21.7
	Distal	AMX	31.50%	53%	21.5	31.30%	53%	21.5
	Proximal	CIP	66.80%	39%	25.1	64.30%	39%	25.1
	Proximal	P–Abx	73.00%	39%	29.3	70.90%	39%	29.3
	Proximal	P-Path	69.80%	38%	28.3	68.70%	39%	28.3
	Distal	P-Path	32.90%	47%	24.4	32.30%	47%	24.4
Leeds HGM	Proximal	SS	51.40%	44%	27.1	50.40%	44%	27.1
	Distal	SS	25.40%	48%	22.8	24.90%	48%	22.8
	Proximal	AMX	66.90%	50%	38.2	65.70%	50%	38.2

	Distal	AMX	30.10%	47%	23.6	29.80%	47%	23.6
	Proximal	CIP	59.60%	50%	32.2	57.60%	50%	32.2
	Distal	CIP	54.50%	50%	28.3	52.60%	51%	28.3
	Proximal	TZP	62.10%	51%	26.6	61.30%	51%	26.6
	Distal	TZP	42.80%	43%	25	42.20%	43%	25
	Proximal	P–Abx	52.50%	42%	23.5	51.20%	42%	23.5
	Distal	P–Abx	27.10%	47%	23.2	26.60%	47%	23.2
	Proximal	P-Path	48.80%	43%	23.5	47.60%	44%	23.5
	Distal	P-Path	26.40%	49%	33.1	26.10%	49%	33.1

Raw Sequence Mean Quality Score

Mean quality scores remain at acceptable levels throughout the sequencing run presented in Figure C-1 Raw Sample Mean Quality Scores.

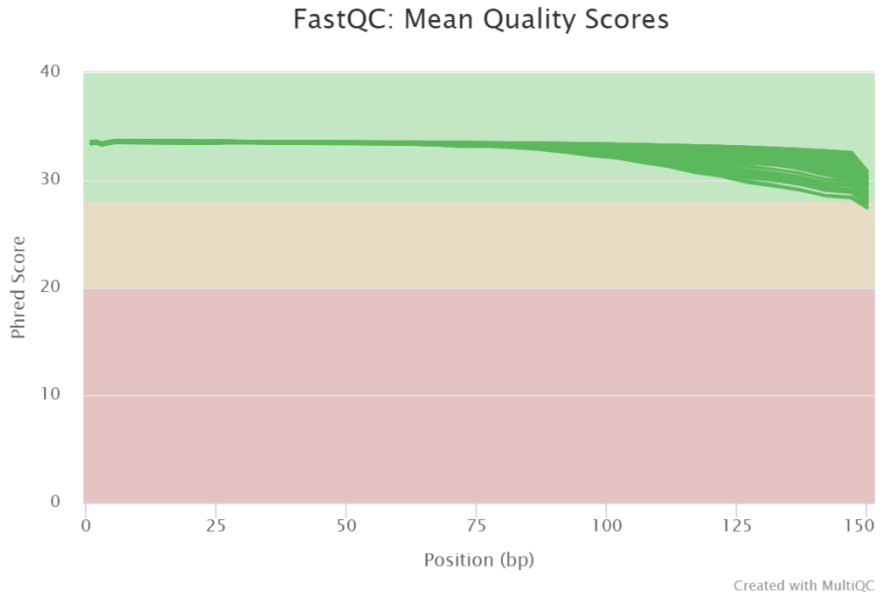

Figure C-1 Raw Sample Mean Quality Scores.

Raw Sequence Adapter Content

Significant adapter read-through was observed across all samples (Figure C-2 Proportion of Raw Sequences with Illumina Universal Adapter Contamination prior to Trimming), it remains unclear whether the adapters were trimmed prior to being supplied. Notably samples believed to experience issues during library creation experienced the least adapter read-through, potentially due to antifoam

causing sheering problems resulting in larger fragment seizes within the libraries.

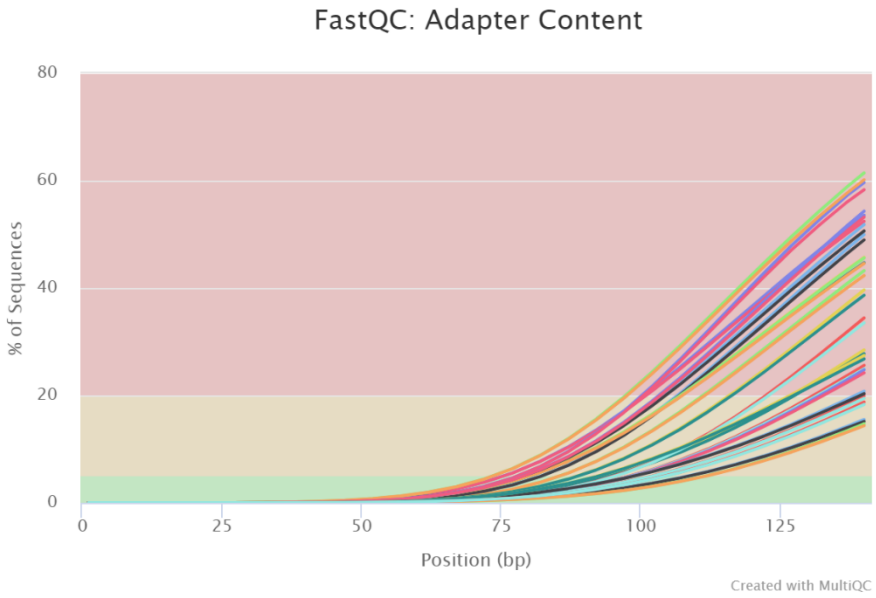

Figure C-2 Proportion of Raw Sequences with Illumina Universal Adapter Contamination prior to Trimming

3.1.2 Trimmomatic

Trimmomatic was run in paired-end mode with default parameters for trimming Illumina data. Summary statistics for Trimmomatic output presented in Table C-2 Trimmomatic Summary Data for Pair-End Trimming**.** Reverse Only Surviving and Dropped reads are both excluded from the table at 0% across all samples. Trimmomatic reduced Illumina adapter content significantly, falling below 1% of sequences across all samples post-processing.

The default behaviour for palindromic adapter removal drops reverse reads after detecting adapter read-through in forwarding reads. The high proportion of read-through in some sample's accounts for the low "Both Surviving" percentage for some samples.

Table C-2 Trimmomatic Summary Data for Pair-End Trimming

	Reactor Stage	Phase	Both Surviving		Forward Only Surviving	
			Million Seqs	%	Million Seqs	%
MiGut	Proximal	SS	18.6	71.41	7.4	28.59
	Distal	SS	19.4	72.25	7.47	27.75
	Proximal	AMX	17.9	82.47	3.8	17.53
	Distal	AMX	9.8	45.79	11.6	54.21
	Proximal	CIP	10.7	42.76	14.3	57.24
	Proximal	P–Abx	15.5	53.12	13.7	46.88
	Proximal	P-Path	14.3	50.80	13.9	49.20
	Distal	P-Path	12.4	50.85	11.9	49.15
Leeds HGM	Proximal	SS	18.5	68.35	8.5	31.65
	Distal	SS	17.9	78.89	4.8	21.11
	Proximal	AMX	21.5	56.42	16.6	43.58
	Distal	AMX	14.5	61.60	9.0	38.40
	Proximal	CIP	4.2	44.17	17.9	55.83
	Distal	CIP	9.7	34.52	18.5	65.48
	Proximal	TZP	9.7	36.53	16.8	63.47
	Distal	TZP	17.4	69.85	7.5	30.15
	Proximal	P–Abx	18.2	77.69	5.2	22.31
	Distal	P–Abx	17.8	76.88	5.3	23.12
	Proximal	P-Path	19.5	83.26	3.9	16.74
	Distal	P-Path	13.6	41.35	19.4	58.65

3.1.3 Paired-End Read Merger (PEAR)

Surviving paired reads were merged with PEAR (0.9.6) with default parameters. 88% of surviving reads were merged across all samples (90.5% MiGut , 86% Leeds HGM) . Following assembly of Unpaired sequences and Unassembled Sequences from Read 1 were concatenated to utilise as many reads as possible in further analysis.

Table C-3 Summary of Assembled Reads in Millions (M Seq) and Percentage of Total Paired Reads

	Reactor Stage	Phase	Assembled Reads	
			Million Seq	%
MiGut	Proximal	SS	16.4	88
	Distal	SS	17.3	89
	Proximal	AMX	13.8	77
	Distal	AMX	8.8	89
	Proximal	CIP	10.3	96
	Proximal	P–Abx	14.8	95
	Proximal	P-Path	13.7	95
	Distal	P-Path	11.7	95
Leeds HGM	Proximal	SS	17.1	92
	Distal	SS	14.4	80
	Proximal	AMX	19.5	90
	Distal	AMX	13.4	93
	Proximal	CIP	13.7	96
	Distal	CIP	9.3	95
	Proximal	TZP	8.1	83
	Distal	TZP	13.9	80
	Proximal	P–Abx	14.7	80
	Distal	P–Abx	14.5	81
	Proximal	P-Path	15.0	76
	Distal	P-Path	13.0	94

Appendix D

Supplementary Material: Metabolomics

4.1 Log-Transformed Intensity Distributions in Negative and Positive Modes (Steady State)

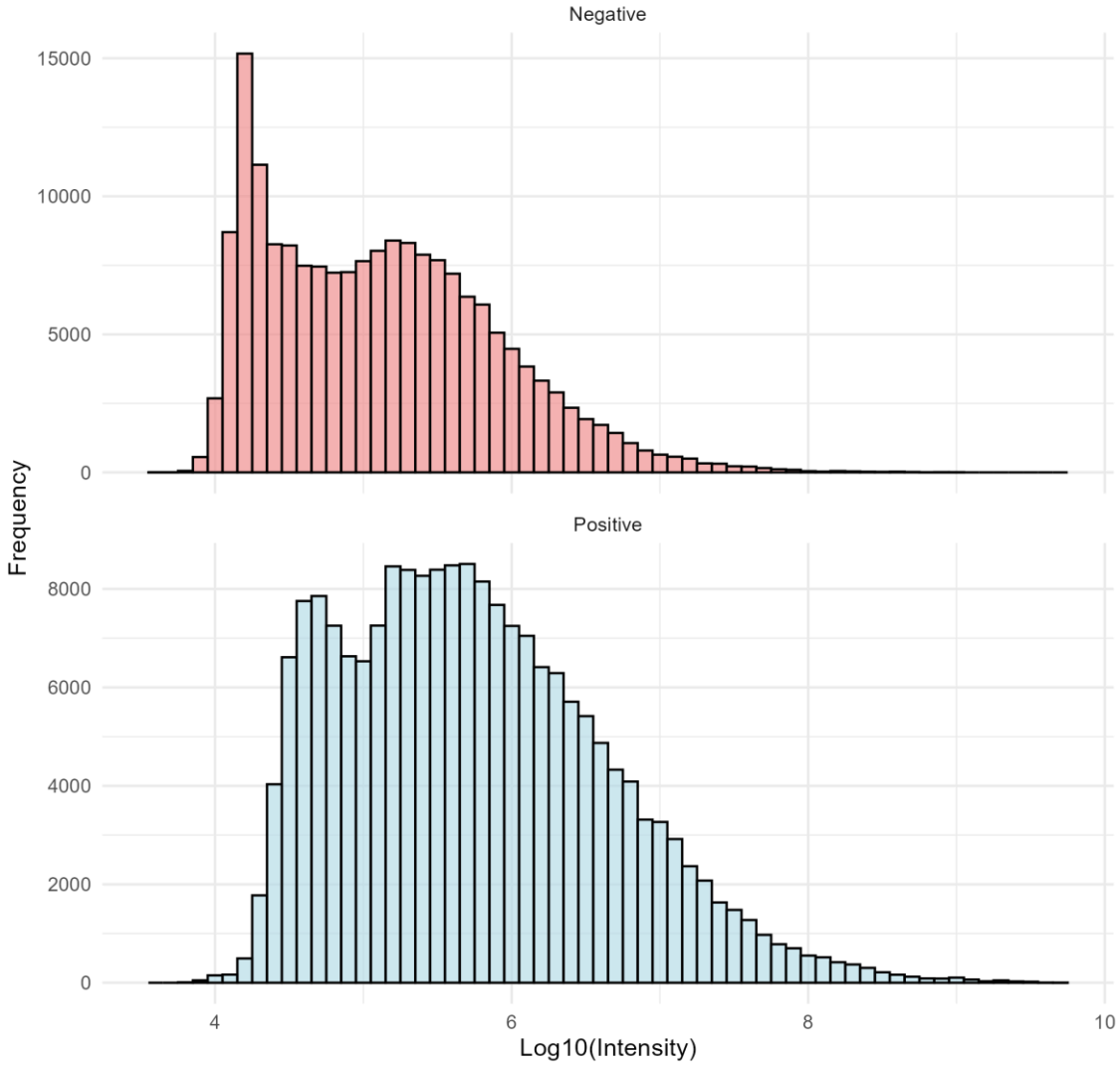

Figure D-1 Log-Transformed Intensity Distributions for Negative and Positive ionisation modes (Steady State). The histograms Illustrate a potentially skewed distribution in both ionisation modes. Data truncation may be attributable may be accountable for the filtering of low-intensity compounds during the processing pipeline.