



**University of
Sheffield**

**Understanding the metabolic impact of a multidrug resistant
plasmid, pLL35, on diverse *Escherichia coli* lineages.**

Kathryn Clare Billane

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Science
School of Biosciences

December 2023

Abstract

Antibiotic resistant infections are predicted to claim 10 million lives annually by 2050. Multi-drug resistant bacterial strains cause untreatable infections or necessitate last-line antibiotics. Conjugative plasmids carry resistance genes and facilitate the evolution of multidrug resistance by horizontal transfer of genetic material. Plasmids have complex interactions with their bacterial hosts that are not fully understood.

Chapter 2 describes an analysis pipeline for untargeted metabolomics that was developed and published to make this complex methodology more accessible, then used throughout the thesis.

This thesis explores the metabolic relationship between the multi-drug resistant plasmid pLL35 in diverse *Escherichia coli* strains across key stages: plasmid acquisition, subsequent coevolution and exposure to antibiotics.

Chapter 3: Upon acquisition of the plasmid, metabolomics reveals strain specific alterations in functions such as ubiquinone biosynthesis, central energy production and amino acid biosynthesis.

Chapter 4: 3 strains that had been experimentally evolved in previous work were compared to their ancestors and displayed metabolic alterations associated with adaptation to lab conditions; amino acid biosynthesis, glycolysis and pyrimidine biosynthesis. The strains had specific metabolic alterations according to evolutionary treatment (plasmid carrying with and without cefotaxime selection) with a singular exception: the global regulator cyclic AMP, which was consistently upregulated in plasmid carriers in response to antibiotic selection. Functions affected by coevolution and selection included amino acid biosynthesis, central energy production and stress responses.

Chapter 5: 4 plasmid carrying strains are exposed to kanamycin, cefotaxime and ciprofloxacin. Common stress and resistance responses are observed, but responses at a pathway and metabolite level are strain specific in extent and direction of change. Pathways affected include amino acid and nucleotide metabolism, energy production, biosynthesis of cell wall and biofilm components and redox stress management, providing insights into bacterial stress responses and resistance mechanisms beyond plasmid-encoded genes. Thus, *Escherichia coli* adapt to plasmid carriage and antibiotic exposure by subtle alterations of its metabolism on a network wide scale.

Acknowledgment

I would like to thank my supervisors Ellie Harrison, Duncan Cameron and Michael Brockhurst for making this project possible. Thank you for all your support, time and guidance through the many changes of the past few years.

Thanks to Michael particularly for guiding the project through the murky waters of the COVID-19 pandemic, and for your writing expertise.

Thanks to Duncan for helping me navigate the challenges of metabolomics and for your moral support.

Thanks to Ellie for your guidance from lab work through to and especially writing up. Thank you for your support through the ups and downs of this PhD.

I am also grateful to the BBSRC for funding this project through the White Rose doctoral training program.

To Laura Carrilero, thank you for your patience and teaching in the lab, and continued support through the years.

To Heather Walker, thank you for your help with mass spectrometry.

To Lizzy Parker, thank you for sharing in the trials of metabolomic analysis so we could create our guide!

To Harry Wright, thank you for your help with statistics.

To the Harrison Lab group, particularly Grace, Mary and Charly, thank you for making our office friendly and encouraging.

To Hero, for standing on my keyboard, sitting on my notes and keeping me company at all hours.

To my parents, Michael and Maria, thank you for always believing in me and your constant support. I am honoured to be your daughter. Thank you to the rest of my family for always being in my corner and for reminding me, 'not a jot'.

Lastly and most importantly, thank you Alice and Robin, without whom I could not have completed this PhD. Your unfailing support, belief and encouragement means the world to me. Thank you for always being there when I needed you.

Declaration

I, the author, confirm that the Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means. This work has not been previously been presented for an award at this, or any other university.

The following publications have arisen from this thesis for which I am first author, or joint first author:

- Billane, K. *et al.* (2022) 'Why do plasmids manipulate the expression of bacterial phenotypes?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1842), pp. 1–9. doi:10.1098/rstb.2020.0461.
- Parker, E.J. *et al.* (2023) 'Untangling the complexities of processing and analysis for un-2 targeted LC-MS data using open-source tools', *Metabolites*, 13(4), pp. 436–448. doi:<https://doi.org/10.3390/metabo13040463>.

Billane *et al.*, 2022 is used as part of the introduction to the thesis.

The methodology used throughout this thesis is published as Parker *et al.*, 2023

Table of Contents

Abstract.....	3
Acknowledgment.....	4
Declaration.....	6
Table of Contents	7
List of tables	9
List of figures	10
Chapter 1 : Main Introduction	13
1.1 Introduction	13
1.2 Review Why do plasmids manipulate the expression of bacterial phenotypes?.....	16
1.3 Thesis Aims.....	25
Chapter 2 : Metabolomics methodology development	28
2.1 Statement of contribution	29
2.2 Laboratory Protocol	33
2.3 Article : Untangling the Complexities of Processing and Analysis for Untargeted LC-MS Data Using Open-Source Tools	37
Chapter 3 : Strain specific metabolic responses of diverse E.coli lineages to the acquisition of a multi-drug resistant plasmid.	48
3.1 Abstract.....	48
3.2 Introduction	48
3.3 Methods.....	53
3.4 Results	55
3.5 Discussion	68
Chapter 4 : Strain specific/limited parallelism responses to coevolution and antibiotic selection of diverse E.coli lineages with a multidrug resistant plasmid.	73
4.1 Abstract.....	73
4.2 Introduction	73
4.3 Methods.....	78
4.4 Results	80
4.5 Discussion	109
Chapter 5 : Metabolic responses of E.coli carrying a multidrug-resistant plasmid to antibiotic selection.....	114
5.1 Abstract.....	114
5.2 Introduction	115

5.3 Methods.....	119
5.4 Results	121
5.5 Discussion	134
Chapter 6 : General Discussion	142
6.1 Overview.....	142
6.2 Chapter 2 Metabolomics methodology development	143
6.3 Chapter 3 Strain specific metabolic responses of diverse <i>E.coli</i> lineages to the acquisition of a multi-drug resistant plasmid.	144
6.4 Chapter 4 Strain specific responses to coevolution and antibiotic selection of diverse <i>E.coli</i> lineages with a multidrug resistant plasmid.....	144
6.5 Chapter 5 Shared metabolic responses to antibiotic exposure in <i>Escherichia coli</i> carrying a multidrug-resistant plasmid	145
6.6 Common themes across chapters	146
6.7 Hypothesis Generation.....	149
6.8 Conclusions	151
References	153
Supplementary Materials.....	171
Supplementary Chapter 3	171
S3 – Data Tables.....	171
S3 – Principle Component Analysis	175
S3 - Mass Spectrometry Parameters	176
S3 - XCMS online Parameters.....	182
S3 - Normalisation of Data.....	183
Supplementary Chapter 4	185
S4 – Data Tables.....	185
S4 – Principle Component Analysis	193
S4 – Normalisation of Data.....	194
S4 –Mass Spectrometry Parameters	195
Supplementary Chapter 5	202
S5 – Normalisation of Data.....	202
S5 – Principle Component Analysis	206
S5 – Data Tables.....	218
S5 – Mass Spectrometry Parameters	256

List of tables

3.1	<i>E.coli</i> strain types	53
3.2	Significant fold changes – plasmid acquisition	66
4.1	<i>E.coli</i> strain types	78
4.2	Metabolite identifications for random forest results strain F022	85
4.3	Metabolite identifications for random forest results strain ELU39	89
4.4	Metabolite identifications for random forest results strain MG1655	93
4.5	Metabolite identifications for significant metabolites plasmid carrying vs plasmid free	101
4.6	Metabolite identifications for significant metabolites antibiotic selection vs plasmid free	103
4.7	Metabolite identifications for significant metabolites antibiotic selection vs plasmid carrying	104
5.1	Summary of resistance genes of pLL35	118
5.2	Proportion of metabolome affected by treatment	127
S3.1	Significant fold changes – plasmid acquisition	169
S4.1	Significant fold changes – plasmid carrying vs plasmid free	183
S4.2	Significant fold changes – antibiotic selection vs plasmid free	186
S4.3	Significant fold changes – antibiotic selection vs plasmid carrying	187
S5.1	Metabolite identifications for the Kanamycin treatment	218

S5.2	Metabolite identifications for the cefotaxime treatment	224
S5.3	Metabolite identifications for the ciprofloxacin treatment	234

List of figures

2.1	Metabolomic workflow	32
3.1	Plasmid Schematic	52
3.2	Random forest results for plasmid acquisition	58
3.3	Heatmap of plasmid acquisition metabolome profiles	61
3.4	Plasmid acquisition PCA	62
3.5	Fold Change maps for plasmid acquisition	63
3.6	KEGG metabolism map for plasmid acquisition	67
4.1	PCA comparing ancestral to evolved metabolomes	80
4.2	Random forest results for F022 compared to the ancestor	83
4.3	Random forest results for ELU39 compared to the ancestor	87
4.4	Random forest results for MG1655 compared to the ancestor	92
4.5	Heatmap of post evolution metabolomes	94

4.6	Fold change maps for post-evolution data	99
4.7	KEGG metabolism map for plasmid carrying vs plasmid free	106
4.8	KEGG metabolism map for antibiotic selection vs plasmid free	107
4.9	KEGG metabolism map for antibiotic selection vs plasmid carrying	108
5.1	Fold change maps strain F022	128
5.2	Fold change maps strain F054	129
5.3	Fold change maps strain ELU39	130
5.4	Fold change maps strain MG1655	131
5.5	Bar chart of commonly affected metabolic functions in the ciprofloxacin treatment	132
5.6	Bar chart of commonly affected metabolic functions in the cefotaxime treatment	132
5.7	Bar chart of commonly affected metabolic functions in the kanamycin treatment	133
6.1	Edit of published figure in Chapter 1	147
S3.1	PCA of plasmid acquisition by strain	173
S3.2	Chapter 3 data normalisation by strain	181
S3.3	Chapter 3 data normalisation	182
S4.1	PCA by evolution treatment	191
S4.2	Chapter 4 data normalisation	192

S5.1	Chapter 5 data normalisation strain F022	200
S5.2	Chapter 5 data normalisation strain F054	201
S5.3	Chapter 5 data normalisation strain ELU39	202
S5.4	Chapter 5 data normalisation strain MG1655	203
S5.5	PCAs for strain F022	204
S5.6	PCAs for strain F054	207
S5.7	PCAs for strain ELU39	210
S5.8	PCAs for strain MG1655	213

Chapter 1 : Main Introduction

1.1 Introduction

WHO has declared antimicrobial resistance a top 10 global health threat. The worldwide spread of resistance genes is central to this crisis, and hospitals are facing an increasing number of untreatable infections (*Antimicrobial Resistance*, 2021). Resistance increases mortality rates and hospitalisation duration of initial admission infections and iatrogenic infections, and also increases risk factors for routine medical procedures like surgery and treatment of immuno-compromised patients such as those undergoing cancer treatment (Mathers, Peirano and Pitout, 2015).

One of the primary mechanisms for the dissemination of resistance genes is horizontal gene transfer (HGT) facilitated by conjugative plasmids, semi autonomous mobile genetic elements (MGEs) (Carattoli et al., 2013). Plasmids are circular pieces of DNA that are formed by a functionally necessary backbone which carries genes for their replication, maintenance and conjugation machinery, enabling DNA transfer through a pilus that can circumvent species barriers (Hall, Brockhurst and Harrison, 2017). In addition to this, plasmids encode accessory genes which often confer a fitness benefit to bacteria by aiding survival of environmental stressors (Norman, Hansen and Sørensen, 2009; Smillie *et al.*, 2010).

Plasmids are a mechanism for rapid adaptation because they can mobilise genes and thus transfer new functions horizontally (Hall, Brockhurst and Harrison, 2017). This facilitates evolution because the new functions can then be maintained and transferred vertically, or incorporated into bacterial chromosomal DNA (Harrison and Brockhurst, 2012; Rodríguez-Beltrán *et al.*, 2021). Horizontal transfer of resistance genes is an effective consequence to the selection pressure exerted by antibiotics (Hall, Brockhurst and Harrison, 2017). The success of plasmids harbouring resistance genes means that the antibiotic resistance crisis is a moving target that is incredibly difficult to contain or treat (San Millan, 2018).

Successful plasmid transfer depends on a balance of costs and benefits to the bacteria upon acquisition. Plasmids can be advantageous in the right environmental contexts, for example resistance genes in the presence of antibiotics, or heavy metals (Norman,

Hansen and Sørensen, 2009). However, plasmids are also frequently associated with fitness costs. Costs can include translational load of additional genetic material, induction of SOS responses, energy usage for pilus constructions and subsequent exposure to phage, creation of cytotoxic gene products and manipulation of regulatory networks (Hall, Brockhurst and Harrison, 2017; San Millan and MacLean, 2017; Dunn *et al.*, 2021; Billane *et al.*, 2022). In some cases costs may be severe enough to impair plasmid persistence in the absence of selection, as with pNUK73 in *P. aeruginosa* (San Millan *et al.*, 2014). Even when plasmids are costly, the source of these costs are not clear, as demonstrated by a study challenging the prevailing hypothesis that plasmids impose a burden on their hosts through the demand on bacterial translation machinery. When ribosomes were hindered in *E.coli* K-12 MG1655 containing one of a series of diverse, clinically relevant plasmids, costs did not increase (Rodríguez-Beltrán *et al.*, 2022).

The severity of plasmid impact on a given bacterial host is dependent on compatibility and hosts may have their fitness affected differently even by genetically similar plasmids (Enne *et al.*, 2004; De Gelder *et al.*, 2007; Humphrey *et al.*, 2012). A summary of fitness changes associated with plasmid carriage cited a range of change in fitness negatively from 1.1% to 27% and positively up to 5% (Carroll and Wong, 2018). Comparing the impacts of an ecologically compatible plasmid pOXA-48_K8 to 25 isolates of *E.coli* and 25 of *K.pneumoniae* demonstrated a small overall phenotypic trend of decreased growth. However when competition assays were conducted for each isolate against plasmid free counterparts there was a spectrum of relative fitness effects from the plasmid ranging from detrimental to beneficial, for both *E. coli* and *K. pneumoniae*, demonstrating the variability and specificity of plasmid-bacterial relationships (Alonso-del Valle *et al.*, 2021).

Plasmid carriage is a series of cost and benefits trade-offs between accessory gene advantage, fitness impact and environmental context. Evolutionary theory therefore suggests that plasmids ought not to be as prevalent as they are because their loss from bacterial populations should be selected for. In the absence of positive selection the plasmid would be expected to be lost due to purifying selection, and even in environmental contexts where accessory genes confer advantages, the movement of plasmid genes into the bacterial chromosome and subsequent lysis or ejection of the plasmid would be selected for (Carroll and Wong, 2018). However, plasmids are stable

for long periods of time without positive selection (Santos-Lopez *et al.*, 2017; Dunn *et al.*, 2021) and are able to transfer within and between bacterial species despite fitness costs (Benz *et al.*, 2021; Sezmis *et al.*, 2023).

This discord between theory and observation has been termed the 'plasmid paradox', for which answers have been found by looking at the relationship between plasmids and their bacterial hosts through the lens of evolution, and by acknowledging that plasmids play an active role in these relationships (San Millan *et al.*, 2014; MacLean and San Millan, 2015; Santos-Lopez *et al.*, 2017; Brockhurst and Harrison, 2022; Sezmis *et al.*, 2023).

Plasmid maintenance in bacterial populations can be explained by ecological and evolutionary mechanisms (Brockhurst and Harrison, 2022). This includes the horizontal transfer of conjugative plasmids which keeps replication and dissemination high enough for maintenance, a process sometimes co-opted by other mobile genetic elements (Peña-Miller *et al.*, 2015; Lopatkin *et al.*, 2017; Che *et al.*, 2021). This is aided by plasmid induced permissiveness of a bacterial host to the carriage of further plasmids (Dionisio, Zilhão and Gama, 2019). The variability of plasmid impact allows for their maintenance in a community and bacterial populations that can stably carry a plasmid sometimes become a reservoir population from which plasmids can continually transfer out of (Hall *et al.*, 2016; Dunn *et al.*, 2021). An evolutionary mechanism for plasmid maintenance is genetic amelioration of costs where the bacteria, plasmid or both mutate to resolve genetic conflicts (Harrison *et al.*, 2015; Hall *et al.*, 2021; Carrilero, Dunn and Moran, 2023).

There is increasing evidence that plasmids also have an impact on the transcriptome and affect bacterial phenotype more extensively than the provision of accessory gene function.

This is explored in the following review.

1.2 Review Why do plasmids manipulate the expression of bacterial phenotypes?

PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Opinion piece



Cite this article: Billane K, Harrison E, Cameron D, Brockhurst MA. 2021 Why do plasmids manipulate the expression of bacterial phenotypes? *Phil. Trans. R. Soc. B* **377**: 20200461.
<https://doi.org/10.1098/rstb.2020.0461>

Received: 17 March 2021

Accepted: 9 July 2021

One contribution of 18 to a theme issue 'The secret lives of microbial mobile genetic elements'.

Subject Areas:

evolution, microbiology

Keywords:

horizontal gene transfer, plasmid, parasitism, mutualism

Author for correspondence:

Michael A. Brockhurst

e-mail: michael.brockhurst@manchester.ac.uk

Why do plasmids manipulate the expression of bacterial phenotypes?

Kathryn Billane¹, Ellie Harrison¹, Duncan Cameron¹ and Michael A. Brockhurst²

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

²Division of Evolution and Genomic Sciences, School of Biological Sciences, University of Manchester, Manchester M13 9PT, UK

EH, 0000-0002-2050-4631; MAB, 0000-0003-0362-820X

Conjugative plasmids play an important role in bacterial evolution by transferring niche-adaptive traits between lineages, thus driving adaptation and genome diversification. It is increasingly clear, however, that in addition to this evolutionary role, plasmids also manipulate the expression of a broad range of bacterial phenotypes. In this review, we argue that the effects that plasmids have on the expression of bacterial phenotypes may often represent plasmid adaptations, rather than mere deleterious side effects. We begin by summarizing findings from untargeted omics analyses, which give a picture of the global effects of plasmid acquisition on host cells. Thereafter, because many plasmids are capable of both vertical and horizontal transmission, we distinguish plasmid-mediated phenotypic effects into two main classes based upon their potential fitness benefit to plasmids: (i) those that promote the competitiveness of the host cell in a given niche and thereby increase plasmid vertical transmission, and (ii) those that promote plasmid conjugation and thereby increase plasmid horizontal transmission. Far from being mere vehicles for gene exchange, we propose that plasmids often act as sophisticated genetic parasites capable of manipulating their bacterial hosts for their own benefit.

This article is part of the theme issue 'The secret lives of microbial mobile genetic elements'.

1. Introduction

Plasmids are semi-autonomous, self-replicating, non-chromosomal DNA elements that are commonly present in bacterial genomes [1]. Many bacterial genomes contain multiple plasmid replicons [2,3], and plasmids have been discovered in the genomes of diverse bacterial taxa from a wide variety of ecological niches, including environmental and clinical settings [4,5]. Plasmid genes can be divided into those encoding either backbone or accessory functions [1,6]. The backbone genes encode plasmid functions, including replication and maintenance, whereas the accessory genes encode non-plasmid functions of potential utility to the bacterial host cell [1,7].

Some plasmids enable the transfer of accessory genes between bacterial strains and species, even between phylogenetically distant lineages [8]. Horizontal gene transfer (HGT) is thus a major driving force in the evolution of bacteria and has contributed significantly to the genomic and ecological diversification of bacterial taxa [9–12]. Plasmid accessory genes encode a wide range of ecological functions, including resistance to toxins, metabolic and catabolic capabilities, and production of virulence factors and anticompensator toxins [13,14]. Plasmids thus enable their bacterial hosts to adapt to environmental stresses, such as antibiotics and toxic metals, or to colonize new niches, for example, through the exploitation of novel substrates or new hosts [8,15,16]. The huge number and diversity of accessory genes creates a vast pool of genetic

variation, enabling bacteria to undergo rapid evolutionary innovation [8,17]. Given this important role in HGT, it is understandable, therefore, that most studies of the ecological and evolutionary impact of plasmids have focused on these accessory functions.

It is becoming increasingly clear, however, that besides the accessory gene functions they encode, plasmid acquisition alters the expression of a wide range of bacterial phenotypes [11,16,18]. These effects of plasmid carriage have typically been studied as the underlying causes of fitness costs because, at least in the laboratory, plasmid acquisition is frequently associated with reduced growth of plasmid-bearers compared with plasmid-free cells [11]. Costly side effects of plasmid carriage are thought to include: induction of SOS responses, cytotoxic gene products, disruption of cellular homeostasis, and the energetic burden of replicating, transcribing and translating new genetic material [11,19].

Nonetheless, plasmids have also been shown to cause differential expression of chromosomal genes, altering the expression of a wide variety of bacterial traits in ways that do not always appear straightforwardly maladaptive. Indeed, there is growing evidence to suggest that, in some cases, these plasmid-mediated alterations to the bacterial phenotype may have niche-adaptive fitness consequences that may well be missed in highly simplified laboratory environments [18]. Plasmid manipulation of bacterial gene regulation could, therefore, play an important role in the relationship between plasmids and their bacterial hosts and, moreover, could mediate the fitness effects of plasmid acquisition.

In this review, we argue that the effects that plasmids have on the expression of bacterial phenotypes may often represent plasmid adaptations, rather than mere deleterious side effects. As self-replicating biological entities, plasmids are capable of evolving adaptations to increase their own fitness. A plasmid's fitness can be defined as the sum of its vertical and horizontal replication (i.e. at bacterial cell division and plasmid conjugation events, respectively). As such, the fitness interests of plasmids need not necessarily always be aligned to those of the bacterial host cell. We begin by summarizing findings from untargeted omics analyses, which give a picture of the global effects of plasmid acquisition on host cells. Thereafter, because many plasmids are capable of both vertical and horizontal transmission, we distinguish plasmid-mediated phenotypic effects into two main classes of potential fitness benefit: (i) those that promote the competitiveness of the host cell in a given niche and thereby increase plasmid replication through vertical transmission, and (ii) those that promote plasmid conjugation and thereby increase plasmid replication through horizontal transmission.

2. What is the 'omic' footprint of plasmid acquisition upon the host cell?

Omics methods can provide an untargeted global view of the impact of plasmid acquisition on the bacterial cell. Transcriptomics, proteomics and metabolomics have each been used to compare plasmid-carrying cells with plasmid-free cells. These studies reveal extensive variation between plasmid–host pairings, in terms of both the degree of alteration caused by the plasmid and the range of cellular functions that are affected (table 1). Whereas some plasmids affect the expression or translation of several hundreds of genes and

many diverse functions, other plasmids have much more limited effects upon their host cell [15,20,21].

In transcriptomic studies, the percentage of differentially expressed chromosomal genes ranges from 0.59 to 20% across diverse plasmid–host interactions [15,20]. This typically includes both up- and downregulation, and where very large numbers of chromosomal genes are affected, is often linked to the plasmid altering the expression of chromosomal regulators. For example, Coulson *et al.* [15] demonstrated that two plasmid-encoded transcriptional regulators affected expression of 18% of the bacterial genome by altering expression of 31 chromosomal regulatory genes, including transcriptional regulators, sigma factors and an anti-termination regulator [15]. Similarly, Shintani *et al.* [22] showed that the acquisition of pCAR1 affected host transcriptional regulators. In a related study, pCAR1 affected the expression of 463 (8.08%) conserved open reading frames (ORFs) in *Pseudomonas putida* KT2440, several of which are involved in translation, transcription and DNA replication cellular processes [21]. Plasmid acquisition can also lead to very large fold-changes in the expression of specific chromosomal genes. For example, in *P. putida* KT2440, acquisition of the plasmid pCAR1 led to 100–200-fold upregulation of the chromosomal gene encoding the efflux system MexEF-OprN (161.8-fold change for MexE, 186.5-fold change for MexF and 113.0-fold change for OprN) [21,22] resulting in a 70-fold increase in the concentration of the MexF protein in the cell (PP_3426) [23].

Chromosomal genes differentially expressed upon plasmid acquisition are involved in a wide variety of bacterial cellular functions. These most commonly include metabolism, respiration, secretion systems, signalling, translation and transcription, motility, the tricarboxylic acid (TCA) cycle and iron acquisition (table 1). While these differentially expressed functions may be common across diverse bacterium–plasmid pairings, the specific genes affected tend to differ. Metabolic pathways altered by plasmid acquisition include amino acid and nucleotide metabolism, and metabolism of energy sources, carbohydrates, nitrogen and lipids [20–24,26]. The direction of the effect of plasmid acquisition upon the expression of secretion systems tends to vary by secretion system, such that Type-III (T3SS) and Type-IV (T4SS) secretion systems are usually upregulated, whereas Type-VI (T6SS) secretion systems are usually downregulated in plasmid carriers, though not exclusively [20,23–26]. All of these secretion systems can have a variety of functions, but generally T3SS and T4SS contribute to bacterial virulence, with an added functional role in conjugation for T4SS [27]. By contrast, T6SS secretion is involved in bacterium–bacterium communication and interaction, including toxin-mediated killing of competitors [27]. Downregulation of genes required for the flagellar complex may account for observed reduction in motility for plasmid-bearers in some cases [21,23]. Other notable bacterial functions affected by plasmid acquisition include surface polysaccharides (e.g. PNAG) and adhesion-related functions involved in biofilm formation, which, for example, in the case of *Acinetobacter baumannii* and *Salmonella enterica*, were downregulated in plasmid-bearers [24,25].

Comparative studies where the same plasmid is introduced into diverse bacterial strains or species reveal that a given plasmid can have very different transcriptional effects in different host backgrounds. For example, the A/C2

Table 1. Bacterial cellular functions differently expressed following plasmid acquisition, compiled from untargeted proteomic, transcriptomics and metabolomics studies.

3

function	bacteria	plasmid	reference
metabolism	<i>Escherichia coli</i> DH10B, <i>Escherichia coli</i> AR060302, <i>Salmonella enterica</i> SL317, <i>Salmonella enterica</i> SL486, <i>Salmonella enterica</i> MH16125, <i>Shewanella oneidensis</i> MR-1	A/C2	[24]
amino acid metabolism	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
nucleotide metabolism	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
energy metabolism	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[21]
	<i>Escherichia coli</i> DH10B, <i>Salmonella enterica</i> SL317, <i>Pseudomonas aeruginosa</i>	A/C2	[24]
carbohydrate metabolism	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas putida</i> KT2440	pCAR1	[22]
	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
nitrogen metabolism	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
lipid metabolism	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
respiration	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[21]
	<i>Salmonella enterica</i> MH16125, <i>Shewanella oneidensis</i> MR-1	A/C2	[24]
secretion systems	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
Type-III	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Salmonella enterica</i> SL317, <i>Salmonella enterica</i> SL486, <i>Salmonella enterica</i> MH16125	A/C2	[24]
Type-VI	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Acinetobacter baumannii</i>	pAB5	[25]
	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
signalling	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
translation and transcription	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PA01, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]

(Continued.)

Table 1. (Continued.)

function	bacteria	plasmid	reference
motility	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PAO1, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[21]
	<i>Salmonella enterica</i> SL317, <i>Salmonella enterica</i> SL486, <i>Salmonella enterica</i> MH16125	A/C2	[24]
biofilm formation and adherence	<i>Acinetobacter baumannii</i>	pAB5	[25]
	<i>Salmonella enterica</i> SL317, <i>Salmonella enterica</i> SL486, <i>Salmonella enterica</i> MH16125	A/C2	[24]
TCA cycle	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PAO1, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[21]
	<i>Escherichia coli</i> DH10B, <i>Shewanella oneidensis</i> MR-1	A/C2	[24]
iron acquisition	<i>Acinetobacter baumannii</i>	pAB5	[25]
	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PAO1, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
	<i>Salmonella enterica</i> SL486, <i>Salmonella enterica</i> MH16125, <i>Shewanella oneidensis</i> MR-1	A/C2	[24]
transporters	<i>Acinetobacter baumannii</i>	pAB5	[25]
	<i>Pseudomonas aeruginosa</i>	pBS228, Rms149, pAKD1, pAMBL1, pAMBL2, pNUK73	[20]
	<i>Pseudomonas putida</i> KT2440	pCAR1	[23]
	<i>Pseudomonas putida</i> KT2440, <i>Pseudomonas aeruginosa</i> PAO1, <i>Pseudomonas fluorescens</i> Pf0-1	pCAR1	[22]
	<i>Salmonella enterica</i> MH16125	A/C2	[24]

plasmid causes downregulation of pathogenicity islands in *Salmonella* hosts, but primarily affects metabolism in *Escherichia coli* strains and *Shewanella oneidensis*. Metabolic functions affected in *E. coli* included: upregulation of 2-carbon and fatty acid metabolism, glycolate metabolism and glycoylate cycle, amino acid degradation and downregulation of amino acid biosynthesis [24]. Very few functions were affected consistently by A/C2 acquisition across all bacterial hosts. Upregulation of genes involved in oxidation/reduction reactions, cellular metabolism and metal cofactor binding occurred in all hosts, while only two genes were universally downregulated, *qacEΔ1* for a quaternary ammonium compound-resistance protein and *sul1* a sulfonamide-resistance dihydropteroate synthase [24]. A comparative study of the pCAR1 plasmid in three different *Pseudomonas* host species (*P. putida* KT2440, *Pseudomonas aeruginosa* PAO1 and *Pseudomonas fluorescens* Pf0-1) showed large differences in the extent of differential expression across species: 15.3% of KT2440 genes, 2.7% of PAO1 genes and 0.7% of Pf0-1 chromosomal genes [21]. Only four genes were affected by plasmid acquisition in all three host species, including one involved in iron acquisition, and two possibly involved in acetate metabolism that were in the same operon [21,22]. Interestingly, the effect of pCAR1 carriage on transcription was most similar

between KT2440 and PAO1, despite KT2440 being more closely related to Pf0-1 phylogenetically, suggesting that transcriptional effects do not scale straightforwardly with genetic similarity of the host in this case.

Alternatively, changes in gene regulation have been quantified for a given bacterial host carrying different plasmids: in *P. aeruginosa* PAO1, a variety of plasmids altered regulation of a few common functional groups, most prominently metabolism (of amino acid, energy production and nitrogen) and secretion systems (Type-III and Type-VI) [20]. Furthermore, 38 chromosomal genes were consistently differentially transcribed in plasmid-bearers carrying different plasmids [20]. The rest of the transcriptional profile varied, indicating that despite these similarities, each plasmid also affected the expression of distinct sets of host functions.

Metabolic analysis has shown that plasmid acquisition can alter metabolic pathways such as glycolysis, the TCA cycle and the pentose phosphate pathway in *E. coli*, corresponding to transcriptomic data from other studies [27]. Untargeted metabolic analysis using mass spectrometry showed the abundances of a large number of compounds were affected in the same way by diverse plasmids in *P. aeruginosa* PAO1. Out of the 5000 compounds that were detected, the levels of 462 compounds were altered by

plasmid acquisition across the sample set, of which the abundance of 11 compounds was significantly different in plasmid carriers for four of the six plasmids, which is much higher than would be expected by chance [20]. Of particular note were altered nucleotide abundances, particularly down-regulated RNA nucleotides and upregulated (or unaltered) deoxynucleotides [20]. However, relatively few compounds could be identified (1.94%), so while metabolic analysis using mass spectrometry appears promising, more studies that cover a greater diversity of bacterial species and plasmids will be needed to identify robust patterns.

The existing omics studies discussed here have some limitations. First, it is rarely confirmed that the plasmid-carrying transconjugants have not acquired other chromosomal mutations that may alter chromosomal transcription independently of the plasmid. This could be determined by curing the plasmid and confirming that transcription returns to wild-type levels, or by genome sequencing the transconjugant to confirm no additional mutations are present [28]. More studies with these additional controls would be valuable. The studies discussed here also almost exclusively focus on gamma-proteobacterial hosts, and it would be useful for future studies to investigate the impact of plasmids in a broader taxonomic range of bacterial hosts outside of this well-studied clade.

The diversity of plasmids is such that it may be difficult or impossible to predict *a priori* how plasmid-encoded genes interact with bacterial regulatory networks [11,29]. We might expect that adaptive plasmid manipulation would cause relatively consistent transcriptional effects across multiple host genotypes encountered in the plasmid's recent evolutionary history. By contrast, among the few existing comparative studies, it would appear that each bacterium and plasmid pairing has a different, unique differential expression profile. However, such studies typically use a few bacterial strains isolated from different locations and habitats; meanwhile, the natural host of the plasmid is often unknown. Future studies are required, therefore, that compare the transcriptional effects of plasmids upon hosts that they coexisted with in nature within ecologically coherent communities, and thus are likely to represent the recent selective environment for the plasmid. In the studies highlighted above, although the specific genes affected may vary, groups of cellular functions commonly affected by plasmid carriage do begin to emerge, for example, bacterial metabolism appears to be the most frequently affected of these functions. While this could represent adaptive manipulation by the plasmid, an alternative hypothesis is that this could instead be a generic response of bacteria to the acquisition of plasmids, and future studies should attempt to distinguish between these competing hypotheses. In future, it will also be valuable to study how the expression of bacterial functions is affected by plasmid acquisition within the context of relevant environmental niches to better understand how plasmids shape the host bacterial phenotype and fitness in nature.

3. Linking altered expression of bacterial functions to plasmid fitness

Understanding the evolutionary impact of plasmid manipulation of the expression of bacterial phenotypes requires an understanding of how these different bacterial phenotypes are linked to plasmid fitness. Plasmid fitness has two main

components, first, replication by vertical transmission to daughter cells, and second, replication by horizontal transmission through cell-to-cell conjugation. In the following sections, we suggest ways in which plasmid manipulation of the expression of chromosomally encoded bacterial traits could potentially affect these plasmid fitness components.

(a) Bacterial phenotypes likely to affect plasmid vertical transmission

Increased vertical plasmid transmission can result from enhanced survival and/or growth of the host bacterium in a given niche. We make the distinction between plasmid fitness benefits deriving from the accessory genes encoded by the plasmid and those caused by differential expression of chromosomally encoded bacterial genes, and focus here only on the latter. To illustrate this idea, we highlight bacterial phenotypes where plasmid-induced changes in expression of chromosomal genes could cause niche-adaptive alterations benefiting both the bacterium and the plasmid. We suggest that this evolutionary strategy could be evident in plasmid manipulation of bacterial traits, including virulence, resistance to antimicrobials and metabolism, that allow bacterial cells to survive stressors or colonize new niches (figure 1).

(i) Increased resistance to antimicrobials

Although many plasmids encode antibiotic resistance genes, in a number of cases, plasmid acquisition has been shown to alter the expression of chromosomally encoded resistance determinants. For example, acquisition of the pCAR1 plasmid causes massive upregulation of the MexEF–OprN efflux system in a number of *Pseudomonas* host species [20]. The MexEF–OprN efflux system provides resistance to a range of antibiotics, including some quinolones, sulfonamides and chloramphenicols [21–23]. Carriage of pCAR1 is, therefore, likely to increase bacterial resistance to antibiotics without itself encoding antibiotic resistance genes, thus potentially enhancing the survival of plasmid-carrying bacterial cells (and thus the plasmid itself) in antibiotic-containing environments.

(ii) Alternative energy sources

The most common differentially regulated bacterial function affected by plasmid acquisition is metabolism. Often, multiple aspects of metabolism are altered (e.g. carbohydrate, energy, amino acid), with the direction of regulation often varying among bacterium–plasmid pairings, sometimes for the same functional group of genes [20]. An interesting example where a consistent effect is observed across diverse host strains is the plasmid pLL35, which causes the upregulation of bacterial anaerobic metabolism genes in phylogenetically diverse *E. coli* backgrounds [28]. Although the effect on bacterial growth is unknown, it is possible that by shifting the host cell from aerobic towards anaerobic metabolism, the plasmid may potentiate gut colonization, and thereby promote the fitness of both the bacterium and the plasmid in this niche.

(iii) Host colonization

Several plasmids have been shown to manipulate the expression of traits likely to enhance bacterial survival within eukaryotic hosts [30]. For example, certain extended spectrum beta-lactamase (ESBL) plasmids upregulate genes (*ompA*, *nha*, *dnaJ*, *arcA*) and outer membrane proteins that enhance survival

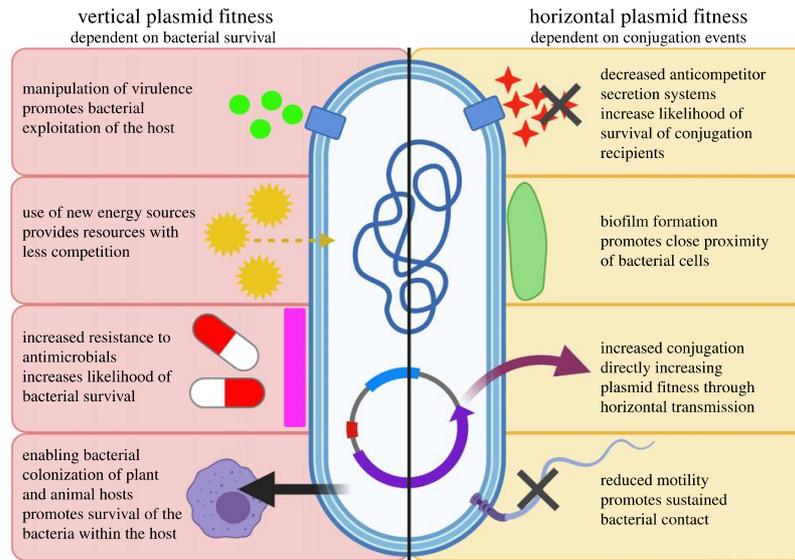


Figure 1. A schematic of how the bacterial phenotypes altered by plasmid acquisition could affect plasmid fitness (created in BioRender.com). We distinguish phenotypic effects according to their likely effects on the modes of plasmid inheritance, vertical from mother cell to daughter cell by replication, or horizontal from cell to cell by conjugation.

of extra intestinal pathogenic *E. coli* in host serum [31,32]. The plasmid pMAR2 upregulates expression of a chromosomal adhesin in enteropathogenic *E. coli*, thus enhancing host colonization by promoting the formation of attaching and effacing lesions in intestinal epithelial cells [33]. Finally, a *Rhodococcus equi* plasmid alters the expression of chromosomal virulence regulators promoting macrophage colonization [15] by arresting phagosomal maturation [34]. In each of these cases, by enhancing bacterial survival within the eukaryotic host, the plasmids may increase their own fitness as well as that of their bacterium in this niche.

(iv) Virulence

Plasmids can promote bacterial exploitation of eukaryotic hosts by altering the production of chromosomally encoded virulence factors. Several plasmids upregulate the bacterial T3SS [20,24], which delivers toxins to degrade eukaryotic cells, thus freeing up host resources for bacterial consumption. In *Chlamydia* species, a plasmid-encoded transcriptional regulator, Pgp4, controls expression of chromosomal genes required for the bacterium to exit the host infected cell in order to infect other cells, a fundamental stage in the *Chlamydia* infection cycle [35–38].

(b) Bacterial phenotypes likely to affect plasmid horizontal transmission

Many plasmids can transfer horizontally to new host cells by conjugation. Even non-conjugative plasmids sometimes undergo horizontal transfer by piggy-backing on the conjugation machinery of other coexisting plasmids, and this can be vital to ensure their survival in the population [39]. The rate of plasmid conjugation is usually plasmid-regulated in a manner that is responsive to conditions in the host cell, such as growth stage [40]. In addition, the rate of plasmid conjugation varies across environments and, for example, can be higher on surfaces that enable higher levels of cell-

cell contact than in planktonic culture [41,42]. In what follows, we highlight examples where plasmids induce changes in bacterial phenotypes that could enhance plasmid conjugation, promoting spread of the plasmid in the bacterial population or community. Because conjugation is energetically expensive to host cells and exposes them to killing by phages that bind the conjugation pilus, these phenotypic changes may be to the detriment of host cell fitness. Bacterial phenotypes that may potentially enhance plasmid horizontal transmission include manipulation of motility, biofilm formation, the T6SS and the DNA replication process (figure 1).

(i) Motility

Plasmid acquisition is often associated with reduced bacterial motility, sometimes caused by plasmid-mediated downregulation of the flagellar complex [21,23,24,43,44]. Cell-to-cell contact is vital for successful conjugation [41], and thus reduced motility may increase the likelihood that bacterial cells remain in contact long enough for the plasmid to undergo conjugation [45], thus potentially enhancing the horizontal transmission of the plasmid.

(ii) Biofilm formation

Increased biofilm formation has been reported in a range of bacterial taxa upon acquisition of conjugative plasmids [46–48]. In *Bacillus subtilis*, increased biofilm formation is mediated by a plasmid-encoded Rap protein (RapP), an intracellular response regulator involved in biofilm formation and sporulation, among other functions [48,49]. Similarly, in some strains of enteropathogenic *E. coli*, ESBL plasmid acquisition is associated with increased production of extracellular biofilm components [32]. Opportunities for plasmid conjugation are expected to be increased in spatially structured populations such as biofilms, presumably owing to increased cell-to-cell contacts, and, therefore, increasing biofilm production may well indirectly increase plasmid horizontal transmission.

(iii) Maintenance and transfer

Plasmid pCAR1 encodes three nucleoid-associated proteins (NAPs). NAPs are global regulators in transcriptional networks, affecting quorum-sensing systems, and bacterial metabolism [50–53]. Intriguingly, plasmids that encode NAPs are more likely to be conjugative [54], suggesting that plasmids may use NAPs to manipulate host cell regulatory networks in ways that promote horizontal plasmid fitness.

(iv) Altering bacterial competition

In *A. baumannii*, the plasmid pAB5 encodes a repressor that deactivates the bacterium's T6SS [55], which would otherwise kill non-kin cells by injecting them with toxins. By deactivating the host cell's T6SS, however, the plasmid ensures the survival of transconjugants, thus increasing the success of conjugation events [56] and thereby the plasmid's rate of horizontal transmission. Intriguingly, by leaving the original host cell unable to deploy its T6SS apparatus in competition with other bacteria, the plasmid may decrease its host's own fitness. This illustrates how plasmid fitness interests can conflict with the bacterial host's fitness interests. Such traits can be favoured provided that the resulting increase in horizontal plasmid replication outweighs the loss of vertical plasmid replication.

4. Future research directions

This review has highlighted some of the growing evidence that the relationship between plasmids and bacteria may be more subtle and manipulative than previously acknowledged. Plasmid manipulation of the expression of bacterial chromosomal genes demonstrates the breadth of parasitic and mutualistic evolutionary strategies plasmids use to maximize fitness. Future studies should consider the following directions:

- How does plasmid manipulation vary across environmental contexts? Laboratory conditions are unlikely to reveal the full extent of niche-specific phenotypic effects caused by plasmid manipulation. Some of the largest effects on bacterial functions have been seen in studies that assess fitness in macrophages or serum [15,31,32]. In macrophages the plasmid affected expression of 20% of bacterial chromosomal genes, including those that slowed phagosome maturation, a key virulence strategy for survival within the eukaryotic host. Future studies should be conducted under conditions more similar to those encountered by the bacteria in nature.
- How does plasmid manipulation vary across a broader taxonomic range of bacterial hosts? Most of the studies discussed in this review have focused on gammaproteobacterial hosts. In order to gain a fuller and more representative view of the impact of plasmids on the expression of bacterial phenotypes beyond this clade, future studies should test a far broader diversity of bacterial hosts and plasmids.
- How might integrated omics studies aid our understanding of how differential regulation leads to altered bacterial phenotypes? Untargeted omics approaches are an efficient way of obtaining the molecular underpinning of bacterial phenotype, and allow us to see nuanced effects of plasmid acquisition. There are many more metabolites than genes to encode their synthesis, and

metabolic pathways are complex and adaptable [57]. It is nearly impossible to predict effects on the metabolome from the wide array of genes that may be differentially expressed upon plasmid acquisition. Therefore, an integrated, multifaceted omics approach may reveal more of the story.

- How does plasmid manipulation of bacteria evolve? One obvious route for plasmid co-option of bacterial gene regulation would be through duplication of bacterial regulatory genes onto the plasmid, followed by divergence. Plasmids (and other mobile elements) frequently acquire bacterial genes through rearrangements [58]. However, it is unclear if such an evolutionary path would be likely. Genes heavily embedded into gene networks tend to be underrepresented on mobile elements [59]. This may be explained by highly connected genes causing far higher disruption to the cell regulatory network [59]. Duplication of bacterial regulatory genes may, therefore, face more significant fitness barriers to establishment than, for example, the acquisition of an accessory trait. Alternatively, plasmid manipulation may arise through non-specific disruption of regulatory networks. Plasmid acquisition can lead to widespread, subtle (and not so subtle) shifts in bacterial gene expression [20–24]. Where these shifts benefit the plasmid, they may be acted on by selection to further embed this function. Further work will be needed to determine what evolutionary trajectories lead to the origination of plasmid regulatory manipulation.
- What are the dynamics of plasmid manipulation traits in bacterial populations and communities? The inheritance of plasmid manipulation traits is likely to differ significantly from inheritance of accessory traits. Plasmid accessory traits are typically, perhaps necessarily [59], self-contained regulatory units whereas manipulation of bacterial gene regulation is likely to be dependent and contingent upon the regulatory network(s) present in the bacterial host. Following from this, we might predict that bacterial manipulation traits may only function in a narrow taxonomic range of hosts, explaining the high variability in the breadth and extent of regulatory effects across hosts, whereas by contrast accessory genes are expected to function similarly across taxonomically diverse hosts.

5. Conclusion

Plasmids can have effects on bacterial phenotypes that extend beyond those conferred by the accessory gene cargo that they encode, by manipulating the expression of genes encoded on the bacterial chromosome. We identify two possible ways that such manipulation could affect plasmid fitness: first, by increasing the growth of the bacterial cell in a particular niche and thus increasing the vertical transmission of the plasmid, or second, by altering the phenotype of bacterial cells in ways that increase the likelihood of conjugation of the plasmid, thus increasing its horizontal transmission. This dichotomy highlights the potential for plasmid manipulation of bacterial phenotypes to result in both mutualistic and parasitic interaction with the bacterial host. Identifying the mechanisms of plasmid manipulation is challenging (cf. [55]) but will be essential to better understand how and why plasmid manipulation has evolved and the role it plays in the evolutionary success of plasmids.

Data accessibility. This article has no additional data.

Authors' contributions. K.B. wrote the manuscript. E.H. commented on drafts of the manuscript. D.C. commented on drafts of the manuscript. M.A.B. commented on drafts of the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by a BBSRC White Rose DTP studentship to K.B. supervised by M.A.B., E.H. and D.C. and grants from NERC and BBSRC to M.A.B. and E.H. E.H. is supported by a NERC Independent Research Fellowship.

References

- Norman A, Hansen LH, Sørensen SJ. 2009 Conjugative plasmids: vessels of the communal gene pool. *Phil. Trans. R. Soc. B* **364**, 2275–2289. (doi:10.1098/rstb.2009.0037)
- Eberhard WG. 1990 Evolution in bacterial plasmids and levels of selection. *Q. Rev. Biol.* **65**, 3–22. (doi:10.1086/416582)
- San Millan A, Peña-Miller R, Toll-Riera M, Halbert ZV, McLean AR, Cooper BS, Maclean RC. 2014 Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. *Nat. Commun.* **5**, 5208. (doi:10.1038/ncomms6208)
- Smalla K *et al.* 2000 PCR-based detection of mobile genetic elements in total community DNA. *Microbiology* **146**, 1255–1258. (doi:10.1099/00221287-146-6-1256)
- Stoesser N *et al.* 2016 Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio* **7**, e02162-15. (doi:10.1128/mBio.02162-15)
- Frost LS, Leplae R, Summers AO, Toussaint A. 2005 Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* **3**, 722–732. (doi:10.1038/nrmicro1235)
- MacLean RC, San Millan A. 2015 Microbial evolution: towards resolving the plasmid paradox. *Curr. Biol.* **25**, R764–R767. (doi:10.1016/j.cub.2015.07.006)
- Hall JPI, Brockhurst MA, Harrison E. 2017 Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria. *Phil. Trans. R. Soc. B* **372**, 20160424. (doi:10.1098/rstb.2016.0424)
- Redondo-Salvo S, Fernández-López R, Ruiz R, Vielva L, de Toro M, Rocha EPC, Garcillán-Barcia MP, Cruz FD. 2020 Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat. Commun.* **11**, 3602. (doi:10.1038/s41467-020-17278-2)
- Heuer H, Smalla K. 2012 Plasmids foster diversification and adaptation of bacterial populations in soil. *FEMS Microbiol. Rev.* **36**, 1083–1104. (doi:10.1111/j.1574-6976.2012.00337.x)
- San Millan A, MacLean RC. 2017 Fitness costs of plasmids: a limit to plasmid transmission. *Microb. Transm.* **5**, 65–79. (doi:10.1128/9781555819743.ch4)
- Brockhurst MA, Harrison E, Hall JPI, Richards T, McNally A, Maclean RC. 2019 The ecology and evolution of pangenomes. *Curr. Biol.* **29**, R1094–R1103. (doi:10.1016/j.cub.2019.08.012)
- Thomas CM (ed.). 2000 *Horizontal gene pool: bacterial plasmids and gene spread*. Amsterdam, The Netherlands: Harwood Academic Publishers.
- Sen D, Van der Auwera GA, Rogers LM, Thomas CM, Brown CJ, Top EM. 2011 Broad-host-range plasmids from agricultural soils have IncP-1 backbones with diverse accessory genes. *Appl. Environ. Microbiol.* **77**, 7975–7983. (doi:10.1128/AEM.05439-11)
- Coulson GB, Miranda-CasoLuengo AA, Miranda-CasoLuengo R, Wang X, Oliver J, Willingham-Lane JM, Meijer WG, Hondalus MK. 2015 Transcriptome reprogramming by plasmid-encoded transcriptional regulators is required for host niche adaptation of a macrophage pathogen. *Infect. Immun.* **83**, 3137–3145. (doi:10.1128/IAI.00230-15)
- MacLean RC, San Millan A. 2019 The evolution of antibiotic resistance. *Science* **365**, 1082–1083. (doi:10.1126/science.aax3879)
- Hall JPI, Harrison E, Lilley AK, Paterson S, Spiers AJ, Brockhurst MA. 2015 Environmentally co-occurring mercury resistance plasmids are genetically and phenotypically diverse and confer variable context-dependent fitness effects. *Environ. Microbiol.* **17**, 5008–5022. (doi:10.1111/1462-2920.12901)
- Vial L, Hommais F. 2020 Plasmid-chromosome cross-talks. *Environ. Microbiol.* **22**, 540–556. (doi:10.1111/1462-2920.14880)
- Baltrus DA. 2013 Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* **28**, 489–495. (doi:10.1016/j.tree.2013.04.002)
- San Millan A, Toll-Riera M, Qi Q, Betts A, Hopkinson RJ, McCullagh J, Maclean RC. 2018 Integrative analysis of fitness and metabolic effects of plasmids in *Pseudomonas aeruginosa* PAO1. *ISME J.* **12**, 3014–3024. (doi:10.1038/s41396-018-0224-8)
- Takahashi Y *et al.* 2015 Modulation of primary cell function of host *Pseudomonas* bacteria by the conjugative plasmid pCAR1. *Environ. Microbiol.* **17**, 134–155. (doi:10.1111/1462-2920.12515)
- Shintani M *et al.* 2010 Response of the *Pseudomonas* host chromosomal transcriptome to carriage of the IncP-7 plasmid pCAR1. *Environ. Microbiol.* **12**, 1413–1426. (doi:10.1111/j.1462-2920.2009.02110.x)
- Vasileva D, Suzuki-Minakuchi C, Kosono S, Yoshida M, Okada K, Nojiri H. 2018 Proteome and acylome analyses of the functional interaction network between the carbazole-degradative plasmid pCAR1 and host *Pseudomonas putida* KT2440. *Environ. Microbiol. Rep.* **10**, 299–309. (doi:10.1111/1758-2229.12639)
- Lang KS, Johnson TJ. 2015 Transcriptome modulations due to A/C2 plasmid acquisition. *Plasmid* **80**, 83–89. (doi:10.1016/j.plasmid.2015.05.005)
- Di Venanzio G, Moon KH, Weber BS, Lopez J, Ly PM, Potter RF, Dantas G, Feldman MF. 2019 Multidrug-resistant plasmids repress chromosomally encoded T6SS to enable their dissemination. *Proc. Natl Acad. Sci. USA* **116**, 1378–1383. (doi:10.1073/pnas.1812557116)
- Lang J, Planamente S, Mondy S, Dessaux Y, Moréra S, Faure D. 2013 Concerted transfer of the virulence Ti plasmid and companion At plasmid in the *Agrobacterium tumefaciens*-induced plant tumour. *Mol. Microbiol.* **90**, 1178–1189. (doi:10.1111/mmi.12423)
- Wang Z, Xiang L, Shao J, Wegrzyn A, Wegrzyn G. 2006 Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism. *Microb. Cell Fact.* **5**, 34. (doi:10.1186/1475-2859-5-34)
- Dunn S, Carrilero L, Brockhurst M, McNally A. 2021 Limited and strain-specific transcriptional and growth responses to acquisition of a multidrug resistance plasmid in genetically diverse *Escherichia coli* lineages. *mSystems* **6**, e00083-21. (doi:10.1128/mSystems.00083-21)
- Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. 2010 Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74**, 434–452. (doi:10.1128/MMBR.00020-10)
- Ronin I, Katsowich N, Rosenshine I, Balaban NQ. 2017 A long-term epigenetic memory switch controls bacterial virulence bimodality. *eLife* **6**, e19599. (doi:10.7554/eLife.19599)
- Ranjan A *et al.* 2018 ESBL-plasmid carriage in *E. coli* enhances in vitro bacterial competition fitness and serum resistance in some strains of pandemic sequence types without overall fitness cost. *Gut Pathog.* **10**, 24. (doi:10.1186/s13099-018-0243-z)
- Schauffer K, Semmler T, Pickard DJ, de Toro M, de la Cruz F, Wieler LH, Ewers C, Guenther S. 2016 Carriage of extended-spectrum beta-lactamase-plasmids does not reduce fitness but enhances virulence in some strains of pandemic *E. coli* lineages. *Front. Microbiol.* **7**, 336. (doi:10.3389/fmicb.2016.00336)
- Gomez-Duarte OG, Kaper JB. 1995 A plasmid-encoded regulatory region activates chromosomal eaeA expression in enteropathogenic *Escherichia coli*. *Infect. Immun.* **63**, 1767–1776. (doi:10.1128/iai.63.5.1767-1776.1995)
- von Bargen K, Polidori M, Becken U, Huth G, Prescott JF, Haas A. 2009 *Rhodococcus equi* virulence-associated protein A is required for diversion of phagosome biogenesis but not for cytotoxicity. *Infect. Immun.* **77**, 5676–5681. (doi:10.1128/iai.00856-09)
- Song L *et al.* 2013 *Chlamydia trachomatis* plasmid-encoded Pgp4 is a transcriptional regulator of

- virulence-associated genes. *Infect. Immun.* **81**, 636–644. (doi:10.1128/IAI.01305-12)
36. Patton MJ *et al.* 2018 Plasmid negative regulation of CPAF expression is Pgp4 independent and restricted to invasive *Chlamydia trachomatis*. *MBio* **9**, 1–14. (doi:10.1128/mBio.02164-17)
 37. Yang C, Starr T, Song L, Carlson JH, Sturdevant GL, Beare PA, Whitmire WM, Caldwell HD. 2015 Chlamydial lytic exit from host cells is plasmid regulated. *MBio* **6**, e01648-15. (doi:10.1128/mBio.01648-15)
 38. Zhong G. 2017 Chlamydial plasmid-dependent pathogenicity. *Trends Microbiol.* **25**, 141–152. (doi:10.1016/j.tim.2016.09.006)
 39. Peña-Miller R, Rodríguez-González R, Madean RC, San Millan A. 2015 Evaluating the effect of horizontal transmission on the stability of plasmids under different selection regimes. *Mob. Genet. Elements* **5**, 29–33. (doi:10.1080/2159256X.2015.1045115)
 40. Sysoeva TA, Kim Y, Rodriguez J, Lopatkin AJ, You L. 2020 Growth-stage-dependent regulation of conjugation. *AIChE J.* **66**, e16848. (doi:10.1002/aic.16848)
 41. Levin BR, Stewart FM, Rice VA. 1979 The kinetics of conjugative plasmid transmission: fit of a simple mass action model. *Plasmid* **2**, 247–260. (doi:10.1016/0147-619X(79)90043-X)
 42. Hausner M, Wuertz S. 1999 High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. *Appl. Environ. Microbiol.* **65**, 3710–3713. (doi:10.1128/AEM.65.8.3710-3713.1999)
 43. Jiang X, Liu X, Law COK, Wang Y, Lo WU, Weng X, Chan TF, Ho PL, Lau TC. 2017 The CTX-M-14 plasmid pHK01 encodes novel small RNAs and influences host growth and motility. *FEMS Microbiol. Ecol.* **93**, fix090. (doi:10.1093/femsec/fix090)
 44. Huang C, Liu LZ, Kong HK, Law COK, Hoa PQ, Ho PL, Lau TCK. 2020 A novel incompatibility group X3 plasmid carrying *bla_{NDM-1}* encodes a small RNA that regulates host fucose metabolism and biofilm formation. *RNA Biol.* **17**, 1767–1776. (doi:10.1080/15476286.2020.1780040)
 45. Luque A, Paytubi S, Sánchez-Montejo J, Gibert M, Balsalobre C, Madrid C. 2019 Crosstalk between bacterial conjugation and motility is mediated by plasmid-borne regulators. *Environ. Microbiol. Rep.* **11**, 708–717. (doi:10.1111/1758-2229.12784)
 46. Matsumoto A, Izutsu H, Miyashita N, Ohuchi M. 1998 Plaque formation by and plaque cloning of *Chlamydia trachomatis* biovar trachoma. *J. Clin. Microbiol.* **36**, 3013–3019. (doi:10.1128/JCM.36.10.3013-3019.1998)
 47. Ghigo J-M. 2001 Natural conjugative plasmids induce bacterial biofilm development. *Nature* **412**, 442–445. (doi:10.1038/35086581)
 48. Parashar V, Konkol MA, Kearns DB, Neiditch MB. 2013 A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence. *J. Bacteriol.* **195**, 2437–2448. (doi:10.1128/JB.02030-12)
 49. Koetje EJ, Hajdo-Milasinovic A, Kiewiet R, Bron S, Tjalsma H. 2003 A plasmid-borne Rap–Phr system of *Bacillus subtilis* can mediate cell-density controlled production of extracellular proteases. *Microbiology* **149**, 19–28. (doi:10.1099/mic.0.25737-0)
 50. Shintani M, Suzuki-Minakuchi C, Nojiri H. 2015 Nucleoid-associated proteins encoded on plasmids: occurrence and mode of function. *Plasmid* **80**, 32–44. (doi:10.1016/j.plasmid.2015.04.008)
 51. Nieto Penalver CG, Cantet F, Morin D, Haras D, Vorholt JA. 2006 A plasmid-borne truncated *luxI* homolog controls quorum-sensing systems and extracellular carbohydrate production in *Methylobacterium extorquens* AM1. *J. Bacteriol.* **188**, 7321–7324. (doi:10.1128/JB.00649-06)
 52. Yun CS *et al.* 2010 Pmr, a histone-like protein H1 (H-NS) family protein encoded by the IncP-7 plasmid pCAR1, is a key global regulator that alters host function. *J. Bacteriol.* **192**, 4720–4731. (doi:10.1128/JB.00591-10)
 53. Yun CS, Takahashi Y, Shintani M, Takeda T, Suzuki-Minakuchi C, Okada K, Yamane H, Nojiri H. 2016 MvaT family proteins encoded on IncP-7 plasmid pCAR1 and the host chromosome regulate the host transcriptome cooperatively but differently. *Appl. Environ. Microbiol.* **82**, 832–842. (doi:10.1128/AEM.03071-15)
 54. Suzuki-Minakuchi C *et al.* 2015 Effects of three different nucleoid-associated proteins encoded on IncP-7 plasmid pCAR1 on host *Pseudomonas putida* KT2440. *Appl. Environ. Microbiol.* **81**, 2869–2880. (doi:10.1128/AEM.00023-15)
 55. Di Venanzo G *et al.* 2019 Urinary tract colonization is enhanced by a plasmid that regulates uropathogenic *Acinetobacter baumannii* chromosomal genes. *Nat. Commun.* **10**, 2763. (doi:10.1038/s41467-019-10706-y)
 56. Weber BS, Ly PM, Irwin JN, Pukatzki S, Feldman MF. 2015 A multidrug resistance plasmid contains the molecular switch for type VI secretion in *Acinetobacter baumannii*. *Proc. Natl Acad. Sci. USA* **112**, 9442–9447. (doi:10.1073/pnas.1502966112)
 57. Cotton TEA, Pétriacq P, Cameron DD, Meselmani MA, Schwarzenbacher R, Rolfe SA, Ton J. 2019 Metabolic regulation of the maize rhizobiome by benzoxazinoids. *ISME J.* **13**, 1647–1658. (doi:10.1038/s41396-019-0375-2)
 58. Che Y, Yang Y, Xu X, Břinda K, Polz MF, Hanage WP, Zhang T. 2021 Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proc. Natl Acad. Sci. USA* **118**, e2008731118. (doi:10.1073/pnas.2008731118)
 59. Cohen O, Gophna U, Pupko T. 2011 The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. *Mol. Biol. Evol.* **28**, 1481–1489. (doi:10.1093/molbev/msq333)

1.3 Thesis Aims

This thesis aims to understand how *E.coli* metabolism is changed by acquisition of a multi-drug resistant plasmid, subsequent co-evolution and exposure to antibiotic stress. The data in this thesis complements previous studies conducted on the *E.coli* strains and pLL35 plasmid which included phenotypic assays, a laboratory evolution experiment, genomics and transcriptomics. The addition of metabolomics from this research provides a more complete omics-based analysis of these strains, contributing to future research area 3 as highlighted in the earlier review: ‘how might integrated omics studies aid our understanding of how differential regulation leads to altered bacterial phenotypes?’

Chapter 3 analyses the impact of the acquisition of the MDR plasmid pLL35 on 9 diverse *E.coli* strains from a range of ecological backgrounds. This chapter begins to broaden the lineage diversity of strains studied by taking clinical and environmental examples. By comparing the strain without pLL35 to isolates of those strains that had pLL35 conjugated over recently (within 30 generations) this data can assess changes in the bacterial metabolism as a consequence of plasmid acquisition.

Chapter 4 analyses a selection of the strains taken from the populations at the endpoint of a 700-generation evolution experiment. To assess how the plasmid-bacterial relationship had evolved, metabolomics of the post-evolution treatment groups were compared to the ancestral state. Further to assess the metabolic differences caused by treatment in the evolution experiment, the groups ‘plasmid free’, ‘plasmid carrying’ and ‘plasmid carrying under antibiotic selection’ were compared to each other. These analyses aimed to demonstrate how the presence of a plasmid may alter the metabolic network of *E.coli* over evolutionary time, and secondly separate the impacts of plasmid carriage alone, and plasmid carriage with positive selection of the resistance genes. This chapter contributes to future research area 4, as highlighted in the earlier review: ‘how does plasmid manipulation of bacteria evolve?’

Chapter 5 assesses the effects of antibiotic stress on the metabolism of a selection of *E.coli* strains, including 2 of a clinical background, which had recently (within 30

generations) acquired pLL35, to aid our understanding of the resistance mechanisms of the bacteria and plasmid as a unit, and the metabolic changes seen in the previous chapters.

To obtain a comprehensive assay of intracellular *E.coli* metabolism, untargeted metabolomics was conducted throughout this thesis using liquid chromatography mass spectrometry. Metabolomics can be defined as a “systematic study of the unique chemical fingerprints that specific cellular processes leave behind” (Di Minno *et al.*, 2021). By producing a highly detailed, precise view of the molecular underpinning of cellular function, which is closely related to phenotype, metabolomics is an invaluable tool for microbiology and medical research (Xu *et al.*, 2014; Vincent *et al.*, 2016; Mardegan *et al.*, 2021; Wei *et al.*, 2021). Transcriptomics and genomics are commonly used to study interactions between plasmids and bacteria, and have done so extensively to great effect, but predicting metabolome effects solely from differentially expressed genes is not straightforward (Billane *et al.*, 2022). The abundance of regulatory effects revealed by transcriptomics suggests plasmid’s impact on their bacterial hosts is different at different biological levels (Billane *et al.*, 2022). Metabolomics may therefore help untangle the nuances of plasmid acquisition.

Untargeted metabolomics comprehensively describes the metabolome and is used in this study to avoid constraints on our hypothesis from previous assumptions, and instead facilitate the generation of new hypotheses (Schrimpe-Rutledge *et al.*, 2016; Di Minno *et al.*, 2021). This technique is particularly suited to identifying perturbations in biological systems, and therefore ideal to study the impact of a plasmid on bacteria (Want, Cravatt and Siuzdak, 2005; Vincent *et al.*, 2016; Allwood *et al.*, 2021; Di Minno *et al.*, 2021).

An understanding of how the bacterial metabolome changes upon plasmid acquisition will contribute to the explanation of resistance plasmid success and persistence. Together the data in this thesis will contribute to an information bank upon which further questions will be based to fully understand the facets of plasmid- bacteria relationships. This will facilitate the development of plasmid-targeting strategies to limit the

pervasiveness of antibiotic resistance genes, an area which is currently underdeveloped for in-vivo solutions (Buckner, Ciusa and Piddock, 2018).

Chapter 2 : Metabolomics methodology development

Metabolomics is the comprehensive assay of all the metabolites in a biological system, and is valuable for understanding biochemical underpinnings of phenotype as the metabolism is the biological level most closely associated with function (Goodacre *et al.*, 2004; Ryan and Robards, 2006).

In this thesis, untargeted liquid-chromatography mass spectrometry is chosen for its precision and broad detection range (Aldridge and Rhee, 2014). Untargeted metabolomics are ideal for identifying perturbations in biological systems, especially when effects are widespread, subtle or unexpected, making it a valuable tool for the generation of novel hypotheses. (Want, Cravatt and Siuzdak, 2005; Vincent *et al.*, 2016; Allwood *et al.*, 2021; Di Minno *et al.*, 2021). In addition, bacteria are well suited to metabolomics and have been used from first conception of the technique (Oliver *et al.*, 1998; Tweeddale, Notley-Mcrobb and Ferenci, 1998). This makes untargeted metabolomics the ideal tool to understand the dynamics of bacterial and plasmid relationships.

The challenges of processing untargeted metabolomic data are widely acknowledged (Allwood *et al.*, 2021). These challenges stem from the nature of the data itself, namely that data processing has a steep learning curve due to the decisions required to process and analyse a given dataset appropriately, such as peak picking and retention time correction settings, filtering, normalisation and potentially the pitfalls that come with overprocessing (Parker *et al.*, 2023). Furthermore, putative identification of metabolites from untargeted data is notoriously difficult (Xu *et al.*, 2014), resulting in thousands of unidentified signals in datasets (Allwood *et al.*, 2021). In part this is due to there being a huge number of metabolites compared to the number of genes (Aldridge and Rhee, 2014). The barriers to engaging in metabolomics are in part responsible for the lack of documentation to navigate these barriers, and together this perpetuates the underutilisation of the technique (Allwood *et al.*, 2021; Parker *et al.*, 2023).

COVID-19 restrictions necessitated the development of a data processing and analysis method that would function off-site and did not rely on proprietary software.

Furthermore, existing methods vary by mass spectrometer, type of mass spectrometry and experimental design and navigating this as a novice is difficult. There is a lack of documentation following a full data analysis process that explains how to make the decisions to tailor data processing and analysis to a given experiment (Schrimpe-Rutledge *et al.*, 2016; Blaženović *et al.*, 2018; Dudzik *et al.*, 2018; Misra, 2018; Chaleckis *et al.*, 2019).

Collaborating with other students and staff at the University of Sheffield, we published an untargeted metabolomics workflow that follows a worked example from laboratory protocols through data processing to analysis (Parker *et al.*, 2023).

The specific data analysis workflow used for this thesis, part of the published methodology, is as follows (Fig 2.1). After sample collection (described pages 28-32) the data was converted to the interoperable .mzml format using proteowizard. XCMS online (parameter programme 84500) was then used to convert mass spectrometer waves into singular peaks, detect and correct for retention time in the liquid chromatography column and re-align the data to produce a list of masses and intensities of the masses per sample. Formatting the XCMS output files followed the published script (Parker *et al.*, 2023). Data analysis was conducted using metaboanalyst, based in R, and included principle component analysis, fold change analysis and random forest analysis. Putative identifications for significant masses were then investigated in a manual manner using the databases Metlin, KEGG and ECMDB.

2.1 Statement of contribution

A successful funding application from the University of Sheffield enabled KCB, the author of this thesis, and EJP to transform method development concerning untargeted metabolomic data acquisition, processing and analysis into an open access guide. The content was developed and written by EJP and KCB equally, and reviewed by new users during workshops led by EJP and KCB.

The author of this thesis KCB has joint first authorship with EJP for the following publication <https://doi.org/10.3390/metabo13040463> which describes software that

was built based entirely on the aforementioned guide. KCB contributed to writing the original draft and creating figures and editing the manuscript during the review process.

The software <https://untargeted-metabolomics-workflow.netlify.app/> has the sections:

00 – Overview of workflow

01 – Metabolite extraction

02 – Data acquisition

03 – Converting data to open format

04 – Data pre-processing

05 – Extracting and formatting peak table and metadata

06 – Multivariate analysis

07 – Putative metabolite ID

08 – Data archiving and reporting

The author contributed to development and content creation throughout the writing of the software, but was particularly responsible for the following sections

1. Overview of workflow
2. Metabolite extraction.
 - Wet lab protocols to prepare *Escherichia coli* for this stage were developed solely by the author and were made available at the institute level but ultimately not included in the examples given in the software. The full protocol to follow.
3. Converting data to open format
 - Proteowizard
 - MSConvert LCMS
4. Data pre-processing
 - Processing LCMS
5. Extracting and formatting peak table and metadata
 - Peak Table LCMS
6. Multivariate analysis
 - Multivariate analysis
 - Metaboanalyst
7. Putative metabolite ID
 - What are my metabolites?
 - METLIN
 - KEGG

The author was involved in editing and refinement of all other sections and ensuring the workflow allowed reproducibility in line with FAIR (Findability, Accessibility, Interoperability, Reusability) principles, just as other authors reviewed the author's work.

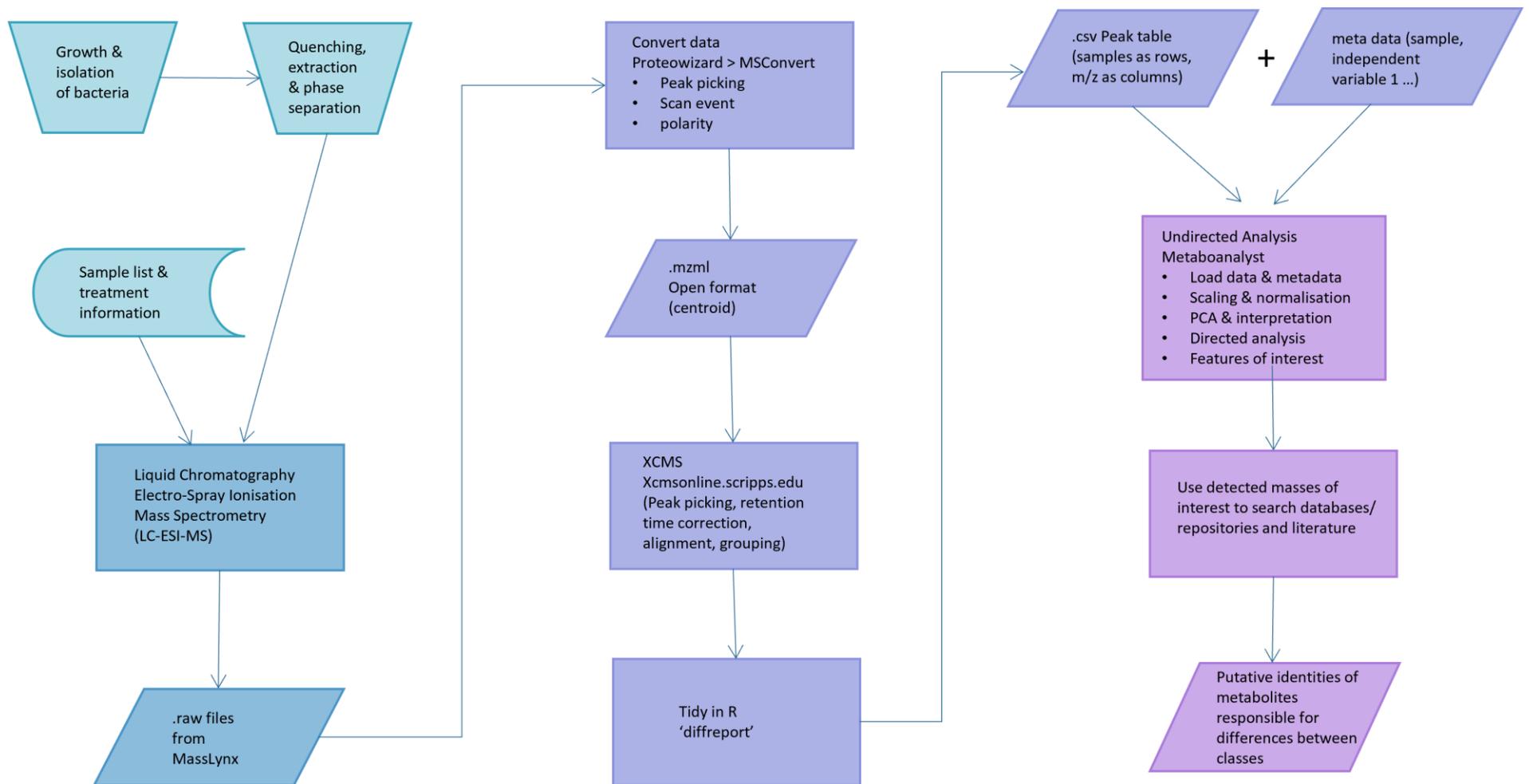


Figure 2.1. A representation of the workflow specific to the author spanning the wet lab preparation of samples, through the data processing to data analysis.

2.2 Laboratory Protocol

The following protocol was developed by the author and is the basis for all sample preparation for the data obtained throughout this thesis, with some modifications for later chapters. The bacterial isolation protocol was developed by the author and the metabolite extraction protocol was adapted from an in-house metabolomic sample preparation written by Dr. H J Walker and optimised for *Escherichia coli*.

Method Development for untargeted metabolomic analysis of diverse *Escherichia coli* strains.

This experiment used *Escherichia coli* isolate from the strains; ST131 derived from clinical bacteraemia, ST1122 and ST394 derived from river effluent and MG1655, a laboratory strain. The plasmid is pLL35, a 106kb incFII(K)-9 plasmid originating from *Klebsiella pneumoniae*. Cultures are grown in nutrient broth at 37°C in a shaking incubator.

Bacterial isolation protocol for metabolite extraction

Materials required per sample:

- 2 Microcosms
- 2 mL Safe-loc eppendorf
- 0.22µm filter
- 10 mL syringe
- 5 mL syringe x 2
- 7 mL M9 buffer

Other Materials

- Agar
- Nutrient broth
- 9cm petri dishes
- Sterile falcon tubes
- Inoculation loops
- Liquid nitrogen

Preparation

Autoclave microcosms containing 10ml nutrient broth.

Autoclave nutrient agar, pour and set plates.

Prepare a master x10 concentration of M9 buffer solution

- In 1L DH20:
- 128g Sodium phosphate dibasic (67.8g if anhydrous)
- 30g Monopotassium phosphate
- 5g Sodium chloride
- 10g Ammonium chloride

Make a x1 working solution.

Streak out *E.coli* strains from glycerol stocks into agar plates and incubate overnight at 37°C.

Inoculate microcosms with a single colony using a loop. Leave lids slightly loose to allow oxygen into the microcosm. Grow in a shaking incubator at 37°C and 180rpm for 24 hours.

Label 2nd set of microcosms and safe-loc eppendorfs.

Procedure

1. Inoculate fresh 10 mL nutrient broth microcosms with 1% (100µl) of 24 hour microcosms and grow in the shaking incubator at 37°C and 180rpm for 3 hours.
2. Gather 1x M9 working solution, syringes, filters, labelled eppendorfs, waste bottles and decant a small flask of liquid nitrogen.
3. Using the 10 mL syringe, take up all of the contents of the microcosm and push through the filter.
4. Follow this with 5 mL of M9 using the 5 mL syringe.
5. Flip filter, move to the 2 mL safe-loc eppendorf and with a fresh 5 mL syringe use 2 mL M9 to wash bacteria out of the filter and into the eppendorf.
6. Repeat steps 3-5 for each sample using sterile syringes and filters each time.
7. Spin down in a centrifuge at 12,000g for 5 minutes.

8. Remove supernatant.
9. Ensure lids are secure on the eppendorf and freeze in the liquid nitrogen.
10. After a short period of time carefully extract eppendorfs from the nitrogen and transfer to the -80°C freezer.

Metabolite extraction protocol for frozen bacterial samples

Materials required:

Sterile MilliQ H₂O (H₂O)

Chloroform (CHCl₃)

Methanol (MeOH)

Eppendorf tubes (2 per sample)

Ice (and dry ice for transportation)

Preparation:

- Pre-chill some eppendorf racks in the -80°C freezer.
- Get a container of ice.
- Pre-chill several mL of H₂O to 0°C.
- Pre-chill some pure CHCl₃ to 0°C.
- Make a several mL mixture (Mix A) of 2.5:1:1 MeOH:CHCl₃:H₂O and pre-chill to 0°C.
- Pre-chill a centrifuge to 0-4°C.

Procedure:

1. Take the eppendorf containing the frozen bacterial sample.
2. Add 180 µL Mix A to bacteria, vortex to resuspend and incubate on ice for 10 minutes.
3. Centrifuge at 12,000g for 5 minutes (0-4 °C).

4. Vortex to resuspend pellet, incubate on ice for 10 minutes.
5. Add 20 μL cold CHCl_3 , and add 30 μL of cold H_2O , shake and mix.
6. Centrifuge at 14,000g for 15 minutes (0-4 $^\circ\text{C}$).
7. You should have two clear phases, aqueous at the top and CHCl_3 at the bottom of the tube. Remove 50-100 μL of aqueous phase and put in a fresh pre-chilled tube. Quickly plunge the tip into CHCl_3 phase, remove ~ 50 μL and put in a fresh pre-chilled tube.
8. Freeze all samples at -80°C , until ready for analysis.
9. Keep the pellet in the eppendorf tubes and freeze.

Article

Untangling the Complexities of Processing and Analysis for Untargeted LC-MS Data Using Open-Source Tools

Elizabeth J. Parker ^{1,*}, Kathryn C. Billane ^{1,*}, Nichola Austen ², Anne Cotton ¹, Rachel M. George ³, David Hopkins ⁴, Janice A. Lake ⁴, James K. Pitman ¹, James N. Prout ¹, Heather J. Walker ³, Alex Williams ¹ and Duncan D. Cameron ⁴

¹ School of Biosciences, University of Sheffield, Sheffield S10 2TN, UK

² Department of Biology, University of Oxford, Oxford OX1 3RB, UK

³ biOMICS Mass Spectrometry Centre, University of Sheffield, Sheffield S10 2TN, UK

⁴ Department of Earth and Environmental Sciences, University of Manchester, Manchester M13 9PL, UK

* Correspondence: lizzyparkerpannell@gmail.com (E.J.P.); kcbillane1@sheffield.ac.uk (K.C.B.)

Abstract: Untargeted metabolomics is a powerful tool for measuring and understanding complex biological chemistries. However, employment, bioinformatics and downstream analysis of mass spectrometry (MS) data can be daunting for inexperienced users. Numerous open-source and free-to-use data processing and analysis tools exist for various untargeted MS approaches, including liquid chromatography (LC), but choosing the ‘correct’ pipeline isn’t straight-forward. This tutorial, in conjunction with a user-friendly online guide presents a workflow for connecting these tools to process, analyse and annotate various untargeted MS datasets. The workflow is intended to guide exploratory analysis in order to inform decision-making regarding costly and time-consuming downstream targeted MS approaches. We provide practical advice concerning experimental design, organisation of data and downstream analysis, and offer details on sharing and storing valuable MS data for posterity. The workflow is editable and modular, allowing flexibility for updated/changing methodologies and increased clarity and detail as user participation becomes more common. Hence, the authors welcome contributions and improvements to the workflow via the online repository. We believe that this workflow will streamline and condense complex mass-spectrometry approaches into easier, more manageable, analyses thereby generating opportunities for researchers previously discouraged by inaccessible and overly complicated software.

Keywords: metabolomics; untargeted; mass-spectrometry; open-source; bioinformatics



Citation: Parker, E.J.; Billane, K.C.; Austen, N.; Cotton, A.; George, R.M.; Hopkins, D.; Lake, J.A.; Pitman, J.K.; Prout, J.N.; Walker, H.J.; et al. Untangling the Complexities of Processing and Analysis for Untargeted LC-MS Data Using Open-Source Tools. *Metabolites* **2023**, *13*, 463. <https://doi.org/10.3390/metabo13040463>

Academic Editor: Seongho Kim

Received: 1 February 2023

Revised: 16 March 2023

Accepted: 20 March 2023

Published: 23 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Untargeted metabolomics is an increasingly popular tool for identifying perturbations within a metabolome and revealing phenotypic complexity in systems [1–4]. It is commonly the first part of a two-step research pipeline, where untargeted studies are used to gather information, identify the metabolome, and generate hypotheses. This is followed by targeted metabolomics which measures specific compounds and requires a priori knowledge of the whole metabolome [1,4,5]. Key to a metabolomics workflow are the data processing and handling steps, which take raw mass spectrometry data and convert them for use in a wide array of multivariate and statistical methods. Currently there is no one standardised pipeline for this step due to variation from sampling methods, instrumentation used, analytical methods employed and the deficit of standardised guidelines [6–11].

After over a decade of experience with proprietary software, the challenge was to address a number of issues with current common practices and embrace an open-source approach to metabolomics data processing and analysis that can have a future legacy. As well as navigating the plethora of analysis options available, with the advent of remote working, it became apparent that researchers conducting untargeted metabolomics analysis required resources to learn how to process mass spectrometry data remotely.

The objective of this work was to develop a guide focussed on processing and analysis of mass spectrometry data, collected to address untargeted metabolomics questions, primarily in the fields of environmental metabolomics and the study of complex plant stress responses. However, the tutorial and workflow have been applied in a range of experimental systems including *E. coli*, potato, barley, organic fertilisers, field soil samples, human cervical mucus, and *Chlorella*. The aim is that the guide will help to move towards standardised methodology and comparable research across the field of metabolomics.

The newly developed workflow presented here is designed to address the question: Which compounds might be responsible for the difference in metabolomic fingerprint between the classes (groups) of samples?

The workflow converts mass spectrometry data to open formats for experiments in which a wide array of compounds are compared between two or more classes of samples. The steps may not result in a definitive difference or unquestionable compound identification, rather the workflow will direct further research and highlight potential compounds to focus on for targeted analysis. This resource is aimed at non-experts, and early career researchers who may not have extensive coding or analytical knowledge. Users are introduced and guided through pre-processing options and data formatting steps which result in a peak table data frame. This peak table forms the basis of the next steps in the workflow, multivariate analysis and putative metabolite ID to give a list of potential compounds that are differentially expressed between groups of samples which can inform the hypothesis for downstream targeted analyses. Alongside some command-line interface, GUI software has also been utilised in the workflow, which can be simpler to learn and easier to operate for new and non-expert users of metabolomics data analysis software [12]. Notably, all software approaches discussed here are free, as the authors believe it is important that the discussed pipelines are accessible.

This collaborative and open-source workflow guide for untargeted metabolomics addresses the need for data-handling tutorials [1] with the key aims of widespread use and continuous improvement, ultimately encouraging integration with multi-omic workflows.

2. Materials and Methods

2.1. Overview and Workflow Diagram

This tutorial guides the user through the untargeted metabolomics workflow that has been developed with some explanation of what each stage achieves. Further details are available in step-by-step guides on the associated website (<https://untargeted-metabolomics-workflow.netlify.app/> accessed on (27 January 2023)), which includes links to relevant open-source tools, and our own interoperable code where appropriate. This tutorial covers the steps required to process LC-ESI-MS data, however detailed instructions for processing MALDI-ToF-MS and DI-ESI-MS using similar open-source tools are also available on the associated website.

An index of openly-available datasets is provided at https://untargeted-metabolomics-workflow.netlify.app/00_overview/06_demo-data/ (accessed on 9 March 2023). These example datasets can be used to demonstrate the workflow presented here.

The workflow has been divided into stages. The following number codes are used in the online guide as well as in the R [13] code and workflow diagram (for an abridged version of this diagram see Figure 1).

00. Overviews, workflow diagram & useful information
01. Metabolite extraction
02. Data acquisition (Mass Spectrometry)
03. Converting data to open format
04. Data pre-processing
05. Extracting & formatting peak table & metadata
06. Multivariate analysis (PCA) & further analysis (if applicable)
07. Putative metabolite identification
08. Archiving data & citing resources

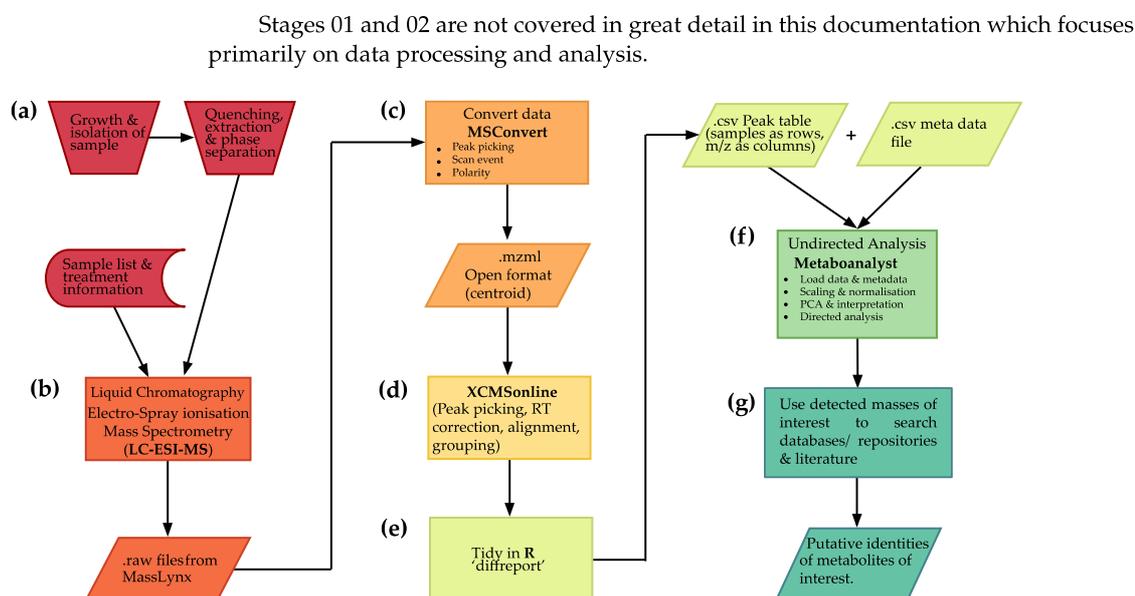


Figure 1. Workflow diagram for processing and analysis of untargeted LC-MS metabolomics data. (a) sample selection and preparation. (b) Mass spectrometry analysis of samples. (c) Conversion of data to open format. (d) Data pre-processing and (e) production of a feature matrix with experimental information included. (f) Statistical analysis for selection of features of interest and (g) identification of features of interest by comparison with literature and existing metabolite databases.

2.2. Experimental Design and Quality Control

Difficulties in analysis and/or workflows can arise from complexities in experimental structure. Many terms are used interchangeably in different contexts. Most tools for untargeted metabolomics are set up for one factor analysis with two or three levels e.g.,

- Case vs. control
- Wild-type vs. transgenic line
- Strain 1 vs. strain 2 vs. strain 3

However, more complex experimental designs are quite often implemented e.g.,

- Two factors with two or more levels in each such as +/– treatment for two strains
- Time course for one or two factors such as +/– treatment for two strains over three time points

To begin, the expectations of which groups of metabolite fingerprints may differ from one another must be considered, and to what extent.

- What are the biological replicates being analysed and are they independent of each other (or has the same organism/population been sampled multiple times)?
- Are there technical replicates (i.e., repeated runs of the same sample)?
- Are Quality Control (QC) samples required? Are analytical standards needed?
- What groupings are required to answer the research questions outlined?

Quality control (QC) can mean different things to researchers from different fields. There are a few simple quality control options for checking that there has not been subtle (or not so subtle) variation accumulating during the run. Decisions must be made on which one (or more) of these are necessary depending on the type of sample to be analysed and the MS techniques employed:

- Spike all prepared samples with a compound for which the m/z (and RT) is known and which is unlikely to be otherwise present in the experimental samples;

- Prepare a pooled QC sample from an aliquot of each of the samples and include this at regular intervals in the MS run;
- Include blanks and/or extraction blanks at regular intervals in the MS run;
- Use lock mass calibration (for Waters instruments).

There are some basic data quality control steps you can take to limit errors during processing and analysis:

- Check file sizes of .raw files across the MS run;
- Check file sizes of converted .mzML files—reconvert any that are unexpected;
- Compare spectra between technical replicates

2.3. Metabolite Extraction and Data Acquisition

Details of quenching, metabolite extraction or choice of mass spectrometry platform are not covered here, as they will likely be specific to the organism and/or tissue involved and the questions being addressed. Figure 2 provides a conceptual overview of metabolite extraction and data acquisition from plant tissues. See [14,15] for introductory guidance and [16] for a specific metabolite extraction method appropriate to plant tissues for this workflow.

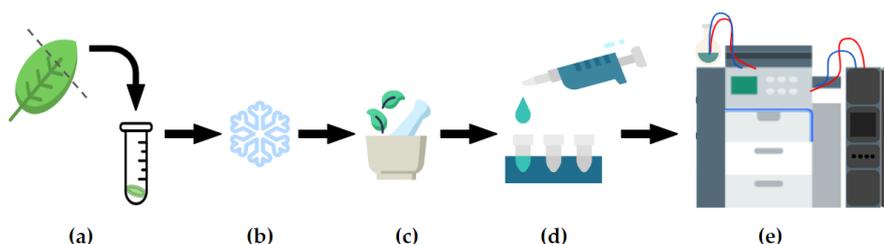


Figure 2. Conceptual diagram of an untargeted metabolomics workflow, from leaf to mass spectrometry analysis. After sample harvest (a), metabolic reactions in a sample tissue must be first quenched (b); i.e., via liquid nitrogen immersion), cell walls lysed and the sample homogenised (c) to permit extraction of compounds within the cells using a range of solvents (d). Extracts may then be diluted and submitted to mass spectrometry analysis (e); e.g., UPLC-ESI-MS).

2.4. Preparing Metadata for Analysis

To process and analyse data using our workflow, two .csv files are required (these can be created in excel, R, google sheets etc. depending on preference) as long as the order and headings of the columns follow the pattern detailed below.

For samplelist.csv the following columns are required:

- “Filename”: this is a list of the filenames of the .mzml files (the part before the .mzml)
- “Filetext”: this is the name that has been manually added to the metadata of that sample
- “MSFile” or an equivalent column that contains either “pos” or “neg” within it. Any other columns will be ignored in this file.

For treatments.csv at least two columns are required:

- “Filetext”: this must contain all the distinct values of “Filetext” from samplelist.csv
- “Variable1”: the naming of this column is left to the user. For example, in an MS run comparing a wild-type to a control, this column could be named “treatment” and filled with “WT” and “C” as appropriate
- “Variable2” etc: further variables. This may include batch identifiers (for example if many samples were run over multiple days), treatments or environmental variables

These are kept in a folder with the .mzml data files. Examples can be found on the website at https://untargeted-metabolomics-workflow.netlify.app/03_conversion-to-open-format/05_samples-treatments/ (accessed on 27 January 2023).

3. Results

3.1. Converting Data to Open Format Using Proteowizard

Converting proprietary data files (which contain a large amount of data and metadata about the run in separate files) to a more manageable format, such as .mzML (the standard open-data format for mass spectrometry [17]) is essential. We have developed this workflow using .RAW files, which are specific to Waters software and are not compatible with many open-source tools. To convert .RAW to .mzML, Proteowizard software [18] is used. Proteowizard is capable of converting many other proprietary file formats and guidance is available through their extensive documentation at https://proteowizard.sourceforge.io/doc_users.html accessed on (20 February 2023). Proteowizard comprises two applications: SeeMS and MSConvert.

SeeMS is useful for viewing chromatograms and spectra without access to proprietary software like MassLynx. MSConvert performs conversion of the MS data but depending on the type of MS used, different settings/parameters in MSConvert may be required, detailed in the online step-by-step instructions to complete stage 03 (https://untargeted-metabolomics-workflow.netlify.app/03_conversion-to-open-format/03_msconvert-lcms/ accessed on 27 January 2023).

It is critically important to check the size of .mzML files once converted. They should all be similar. SeeMS can be used to check any that seem unusual and reconvert any with an incongruous file size (problems in conversion can arise, for instance from intermittent internet connection when converting files from a remote drive).

3.2. Preprocessing Data

Untargeted metabolomics datasets can be several GB in size! To get from compressed .mzML files to a tractable peak table that can be interrogated with multivariate statistics, it is necessary to “tidy” the data.

A peak table is a data-frame consisting of aligned spectra with concentration or intensity values against a set of features—mass to charge ratio (m/z) or m/z with retention time (RT). The file size will be dependent on sample number but will be smaller than the .mzML files.

Different downstream tools for multivariate statistics will require the peak table in slightly different formats, so the code included in this guide will help with formatting for some common uses (e.g., MetaboAnalyst one factor and two factor peak tables) as well as helping format treatment information as metadata so that peak tables can be interrogated.

- Depending on the MS approach, different stages are involved but they broadly fall into:
- Baseline correction and/or noise reduction (estimating what part of the detected intensity is the sample and “cleaning” or adjusting the spectra to show only the signal believed to be associated with the sample);
- Normalisation and/or standardisation (these can mean a range of different things to different people but broadly cover accounting for differences in sample volume or concentration or total intensity of the signal);
- Grouping and peak picking (wave-form algorithms are used to determine which parts of the spectra constitute separate peaks utilising their m/z value);
- Alignment or peak matching (assessing across samples to determine whether peaks with slightly different m/z values are the same peak so that samples can be compared more reliably).
- The above criteria are very important when processing data as they can have a big impact on data quality however the parameters may vary with different datasets and different analysis methods. The importance of these factors have been discussed previously by [19].

By the end of this stage, data will be processed into a single table containing all the m/z and intensity values required for down-stream analysis. This stage relies on the use of open-source software (XCMS online [20] for LC-ESI-MS and MassUp [21] for MALDI-ToF-

MS and DI-ESI-MS) to process the data. These provide user interfaces for well-documented R packages (XCMS [22] and MALDIquant [23] respectively) and provide the advantage of coping well with large datasets and, in the case of XCMS online, being run remotely.

For detailed instructions on pre-processing, consult stage 04 of our online guide (https://untargeted-metabolomics-workflow.netlify.app/04_data-preprocessing/ accessed on (27 January 2023)).

R code to extract a peak table from pre-processed data is available in stage 05 of our online guide (https://untargeted-metabolomics-workflow.netlify.app/05_extracting-formatting-peak-table/ accessed on (27 January 2023)).

3.3. Multivariate Analysis

There are often two key questions when analysing a new untargeted metabolomics dataset:

- Are the metabolomic fingerprints distinct classes (treatment groups) different from each other?
- Which features of the metabolomic fingerprint are causing them to be different from each other?

To answer the first question, data ordination is required to provide a global overview of the variability and patterns within the data. Principal Component Analysis (PCA) is a commonly applied ordination tool that reduces the dimensionality of multivariate data to display complex relationships between samples in 2 or 3 dimensions [15]. As it is unsupervised the model is unaware of the classes to which the samples belong, so patterns are unbiased by a priori knowledge of the experimental design. PERMANOVA can be used to provide statistical corroboration of patterns observed in the PCA by statistically evaluating if significant trends exist at the higher levels of the experimental design within multivariate data i.e., if significant treatment and interaction effects are present. Finally, where clear differences between classes in the PCA are apparent, pairwise comparisons between classes (treatment groups) can be investigated via exploring the loadings or using a pairwise analysis such as t-tests or volcano plots. These will provide the user with features of interest that are most important at defining the statistical output [15].

Where patterns are less clear, supervised analysis, such as OPLS-DA (orthogonal projections of latent structures) may be employed to mine for differences between any two classes. The output of supervised analyses will highlight particularly highly abundant features that differ between two randomly assigned classes that may be obscured in global overview if the majority of the metabolome is conserved or unchanging (this can occur in tissues where only small numbers of metabolites respond to a stimulus, but the majority of the metabolome is unaffected). To limit false positives it is important to consider the native separation in the data (i.e., through an unsupervised ordination, like PCA) to provide a robust biological justification for comparing two particular classes. The analyses exemplified here are by no means the only option, and it is highly recommended that tools such as MetaboAnalyst [24] are employed by the researcher to explore all analytical avenues available.

In the online guide, demonstration is given on how to perform these analyses using a free online platform and how to run some alternative code in R. MetaboAnalyst is an online platform on which untargeted metabolomics data can be loaded, normalised, analysed and visualised. However, there is a strong emphasis on detailed statistics that may be more appropriate for targeted analyses, so the user must have a clear understanding of their objectives in choosing amongst the options.

MetaboAnalyst is interoperable with R and the underlying code can be accessed using the button at the top left of the “Results” page. The advantage of running the code is that the user can integrate it with other analyses (and formatting for figures). Examples of figures produced with this approach can be found in Figure 3. In contrast, the advantage of the MetaboAnalyst GUI is that it guides the user through the process and has some useful sense-checks and vignettes available.

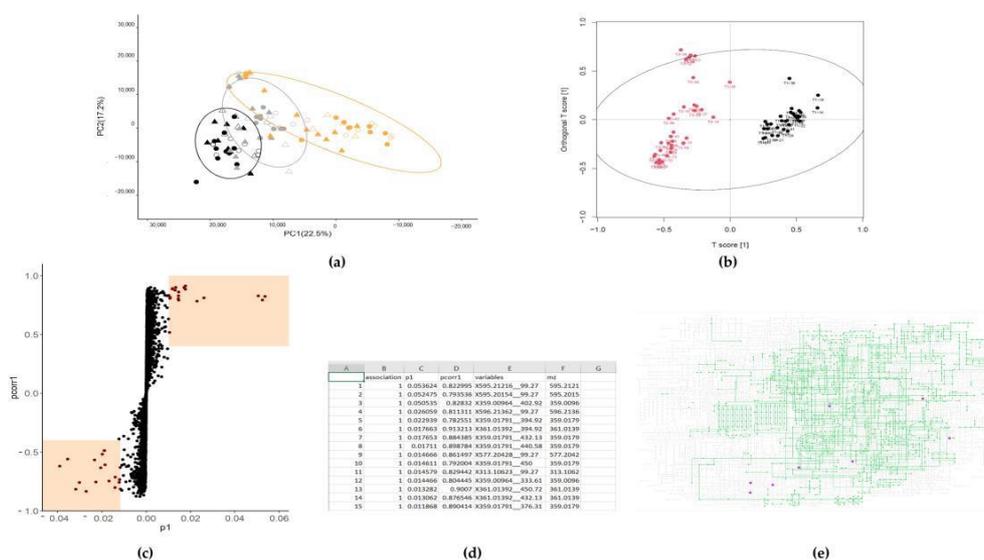


Figure 3. Conceptual diagram of examples of multivariate analysis outputs of untargeted metabolomics analysis, all produced using open-source or freely available software. (a) Principal component analysis (PCA) 2-D scores plot produced with *pcaMethods* and *ggplot2* packages in R; (b) OPLS-DA scores plot produced using the *muma* package in R; (c) scores plot created using *ggplot2* package and data produced by the *muma* package in R; (d) example list of features of interest highlighted by an OPLS-DA using *muma* in R; (e) example of metabolites highlighted within a KEGG pathways global *Escherichia coli* metabolism map.

Details can be found via the excellent tutorials and documentation provided by MetaboAnalyst [25].

It is also possible to analyse the same peak tables using SIMCA (Umetrics) or other proprietary softwares. However, it is much harder (and more costly) to use these remotely, and it is harder to document any analysis for sharing with other researchers. Other software worth considering includes MSDial, MetaboKit and MeV [26–28].

3.4. What Are My Metabolites?

It is very important to consider that this stage of the metabolomic process is not automated and can be incredibly time-consuming and challenging to do, so it is advisable that the preceding analysis has been adequately assessed for its effectiveness before committing time at this stage.

Annotating metabolomic features is challenging—there are some automated annotations included with e.g., XCMS that rely on the CAMERA package [29] amongst others. However, these often struggle with unusual experimental structures and/or large datasets, or “unusual” (i.e., non-human) metabolites. Thus, reducing the number of metabolomic features to those that are causing a significant (in terms of reliability and magnitude) difference between two classes of samples is advisable.

To ascertain the identity of these features, comparing the m/z (or m/z at specific RT) values highlighted by multivariate analysis with databases of reference m/z and with experimental data from the literature (usually available in a publication or in repositories like MetaboLights [30] and Metlin [31]) is key.

Stage 07 of the online guide provides guidance on using a range of databases to help annotate “metabolites of interest” (https://untargeted-metabolomics-workflow.netlify.app/07_putative-metabolite-id/ accessed on (27 January 2023)). These include:

- METLIN to search by m/z ;

- KEGG PATHWAY and KEGG COMPOUND [32] to corroborate likelihood of detecting certain compounds in the study organism/sample and to gain insight on biological function;
- Data repositories such as MetaboLights;
- Details of how to find other relevant databases (MassBank, PubChem, MetaCyc, Metabolomics Workbench [33–36]);
- Reporting Metabolomics Standards Initiative (MSI) identification levels (see also [37]).

3.5. Sharing Metabolomics Data

Metabolomics data from even a small study can be very large. It can also be very complex. But there are ways of sharing it with the wider scientific community (and indeed the public) without too much trouble. It is insufficient to only prepare a data availability statement or simply share graphs or peak tables.

Metabolomics data can be analysed in lots of different ways, so it is important to comply with the FAIR principles [38]:

- Findable
- Accessible
- Interoperable
- Reusable

Institution-based data repositories are an option, but they often require extra levels of support to submit large datasets and there is no guarantee that access to other researchers is feasible.

More useful is a field-specific repository where data will be made available together with other relevant data sets. Furthermore, these repositories provide guidance on appropriate data formatting, allowing it to be compatible with other published data to form part of potential future meta-analyses. Some journals will have specific guidelines on which repository to use [39].

Time should be set aside from the outset of any project for submitting data to a repository. It is not optional!

MetaboLights is a data repository specific to metabolomics studies [30]. Data from NMR, GC-MS, LC-MS, and MALDI amongst others, may be submitted.

The repository is maintained and curated by the European Bioinformatics Institute (EMBL-EBI) meaning that the data it holds is well-formatted and integrated with several other standardised databases and ontologies (ways of describing methods, data and metadata). This “future-proofs” the data stored, making it not only open-access but also more findable and reusable, as well as facilitating integration with other -omics data, if required.

MetaboLights has various stages of submission, validation and then curation by experts to make sure each submission has all the relevant metadata needed to recreate the analysis undertaken. Following curation, there is a review process and finally data can be added to the repository and made available.

Because of the curation process, there can be a significant lag between submission and data being available so early submission is advisable. However, once submitted, there is a reference that can be linked to any publication [30].

Account creation is required, after which, a video tutorial guide on using the submission portal is available. Additional hints and tips on this can be found on the associated website (https://untargeted-metabolomics-workflow.netlify.app/08_data-archiving-citation/02_metabolights/ accessed on (27 January 2023)).

3.6. Citation of the Tools Used in the Workflow

Links to cite the following tools involved in the workflow can be found at https://untargeted-metabolomics-workflow.netlify.app/08_data-archiving-citation/03_citing-tools/ accessed on (21 February 2023). These tools are regularly updated so it is important to cite the version used and/or the date accessed:

- All R packages used;
- R and RStudio versions;
- Proteowizard (SeeMS and MSConvert);
- Metaboanalyst;
- XCMS online and METLIN;
- MassUp;
- MassBank (including access date);
- ECMDDB and any other organism specific metabolite databases used;
- KEGG (including BRITE, COMPOUND and PATHWAY);
- PubChem;
- A data availability statement that links to your archived data (e.g., in MetaboLights).

4. Conclusions

At this point the choice in preparing and analysing metabolomics data is at the discretion of the research group. This guide is a useful starting point that leads the reader through an openly available, best-practice, pipeline. Complex data and analytical processes can be overwhelming, but by engaging in discussion forums, sharing ideas, troubleshooting, and having access to a community of like-minded researchers these processes can become more accessible and facilitate exploration of exciting biological questions.

Author Contributions: Conceptualization, E.J.P. and D.D.C.; software, E.J.P., J.K.P., J.N.P. and R.M.G.; resources, D.D.C.; writing—original draft preparation, E.J.P. and K.C.B.; writing—review and editing, K.C.B., N.A., A.C., D.H., J.A.L., J.N.P., H.J.W. and A.W.; supervision, H.J.W. and D.D.C.; project administration, E.J.P.; funding acquisition, E.J.P. and D.D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by BBSRC, grant numbers BB/M011151/1, and BB/T010789/1. The project was supported by a small grant from the University of Sheffield Library’s Unleash Your Data and Software Competition.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analysed in this study. Research software described in this article is available online https://github.com/LizzyParkerPannell/Untargeted-metabolomics_workflow accessed on (27 January 2023). The associated online guide is available at <https://untargeted-metabolomics-workflow.netlify.app/> accessed on (27 January 2023).

Acknowledgments: With thanks to Erika Hansson, Sophia van Mourik, Emily Magkourilou and Harry Wright for their feedback on the content of the website and to Tim Daniell and Giles Johnson for encouraging the sharing of the workflow. Many thanks to Neil Shephard and Robert Turner from the RSE Team, Department of Computer Science, University of Sheffield for their technical help and guidance in developing the website.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Allwood, J.W.; Williams, A.; Uthe, H.; van Dam, N.M.; Mur, L.A.J.; Grant, M.R.; Pétriacq, P. Unravelling Plant Responses to Stress—The Importance of Targeted and Untargeted Metabolomics. *Metabolites* **2021**, *11*, 558. [[CrossRef](#)] [[PubMed](#)]
2. Want, E.J.; Cravatt, B.F.; Siuzdak, G. The expanding role of mass spectrometry in metabolite profiling and characterization. *ChemBioChem* **2005**, *6*, 1941–1951. [[CrossRef](#)] [[PubMed](#)]
3. Vincent, I.M.; Ehmann, D.E.; Mills, S.D.; Perros, M.; Barrett, M.P. Untargeted metabolomics to ascertain antibiotic modes of action. *Antimicrob. Agents Chemother.* **2016**, *60*, 2281–2291. [[CrossRef](#)]
4. Di Minno, A.; Gelzo, M.; Stornaiuolo, M.; Ruoppolo, M.; Castaldo, G. The evolving landscape of untargeted metabolomics. *Nutr. Metab. Cardiovasc. Dis.* **2021**, *31*, 1645–1652. [[CrossRef](#)]
5. Wei, Y.; Jasbi, P.; Shi, X.; Turner, C.; Hrovat, J.; Liu, L.; Rabena, Y.; Porter, P.; Gu, H. Early Breast Cancer Detection Using Untargeted and Targeted Metabolomics. *J. Proteome Res.* **2021**, *20*, 3133. [[CrossRef](#)] [[PubMed](#)]
6. Schrimpe-Rutledge, A.C.; Codreanu, S.G.; Sherrod, S.D.; McLean, J.A. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1897–1905. [[CrossRef](#)] [[PubMed](#)]

7. Dudzik, D.; Barbas-Bernados, C.; García, A.; Barbas, C. Quality assurance procedures for mass spectrometry untargeted metabolomics. A review. *J. Pharm. Biomed. Anal.* **2018**, *147*, 149–173. [CrossRef]
8. Rainer, J.; Vicini, A.; Salzer, L.; Stanstrup, J.; Badia, J.M.; Neumann, S.; Stravs, M.A.; Verri Hernandez, V.; Gatto, L.; Gibb, S.; et al. A Modular and Expandable Ecosystem for Metabolomics Data Annotation in R. *Metabolites* **2022**, *12*, 173. [CrossRef]
9. Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites* **2018**, *8*, 31. [CrossRef]
10. Misra, B.B. New tools and resources in metabolomics: 2016–2017. *Electrophoresis* **2018**, *39*, 909–923. [CrossRef]
11. Chaleckis, R.; Meister, I.; Zhang, P.; Wheelock, C.E. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Curr. Opin. Biotechnol.* **2019**, *55*, 44–50. [CrossRef]
12. Chang, H.-Y.; Colby, S.M.; Du, X.; Gomez, J.D.; Helf, M.J.; Kechris, K.; Kirkpatrick, C.R.; Li, S.; Patti, G.J.; Renslow, R.S.; et al. A Practical Guide to Metabolomics Software Development. *Anal. Chem.* **2021**, *93*, 1912–1923. [CrossRef] [PubMed]
13. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2010. Available online: <https://www.R-project.org/> (accessed on 23 February 2023).
14. Lu, W.; Su, X.; Klein, M.S.; Lewis, I.A.; Fiehn, O.; Rabinowitz, J.D. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annu. Rev. Biochem.* **2017**, *86*, 277–304. [CrossRef]
15. Pezzatti, J.; Boccard, J.; Codesido, S.; Gagnebin, Y.; Joshi, A.; Picard, D.; González-Ruiz, V.; Rudaz, S. Implementation of liquid chromatography-high resolution mass spectrometry methods for untargeted metabolomic analyses of biological samples: A tutorial. *Anal. Chim. Acta* **2020**, *1105*, 28–44. [CrossRef] [PubMed]
16. Austen, N.; Walker, H.J.; Lake, J.A.; Phoenix, G.K.; Cameron, D.D. The Regulation of Plant Secondary Metabolism in Response to Abiotic Stress: Interactions Between Heat Shock and Elevated CO₂. *Front. Plant Sci.* **2019**, *10*, 1463. [CrossRef] [PubMed]
17. Martens, L.; Chambers, M.; Sturm, M.; Kessner, D.; Levander, F.; Shofstahl, J.; Tang, W.H.; Römpp, A.; Neumann, S.; Pizarro, A.D.; et al. mzML—A Community Standard for Mass Spectrometry Data. *Mol. Cell. Proteom.* **2011**, *10*, R110.000133. [CrossRef] [PubMed]
18. Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **2008**, *24*, 2534–2536. [CrossRef]
19. Forsberg, E.; Huan, T.; Rinehart, D.; Benton, H.P.; Warth, B.; Hilmers, B.; Siuzdak, G. Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS Online. *Nat. Protoc.* **2018**, *13*, 633–651. [CrossRef]
20. Katajamaa, M.; Orešič, M. Data processing for mass spectrometry-based metabolomics. *J. Chromatogr. A* **2007**, *1158*, 318–328. [CrossRef]
21. López-Fernández, H.; Santos, H.M.; Capelo, J.L.; Fdez-Riverola, F.; Glez-Peña, D.; Reboiro-Jato, M. Mass-Up: An all-in-one open software application for MALDI-TOF mass spectrometry knowledge discovery. *BMC Bioinform.* **2015**, *16*, 318. [CrossRef]
22. Smith, C.A.; Want, E.J.; O’Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. *Anal. Chem.* **2006**, *78*, 779–787. [CrossRef] [PubMed]
23. Gibb, S.; Strimmer, K. MALDIquant: A versatile R package for the analysis of mass spectrometry data. *Bioinformatics* **2012**, *28*, 2270–2271. [CrossRef] [PubMed]
24. Xia, J.; Psychogios, N.; Young, N.; Wishart, D.S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucl. Acids Res.* **2009**, *37*, 652–660. [CrossRef]
25. Metaboanalyst Tutorials. Available online: <https://dev.metaboanalyst.ca/docs/Tutorials.xhtml> (accessed on 27 January 2023).
26. Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; van der Gheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **2015**, *12*, 523–526. [CrossRef]
27. Narayanaswamy, P.; Teo, G.; Ow, J.R.; Lau, A.; Kaldis, P.; Tate, S.; Choi, H. MetaboKit: A comprehensive data extraction tool for untargeted metabolomics. *Mol. Omics* **2020**, *16*, 436. [CrossRef]
28. Howe, E.; Holton, K.; Nair, S.; Schlauch, D.; Sinha, R.; Quackenbush, J. MeV: MultiExperiment Viewer. In *Biomedical Informatics for Cancer Research*; Ochs, M., Casagrande, J., Davuluri, R., Eds.; Springer: Boston, MA, USA, 2010; pp. 267–277. [CrossRef]
29. Kuhl, C.; Tautenhahn, R.; Boettcher, C.; Larson, T.R.; Neumann, S. CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* **2012**, *84*, 283–289. [CrossRef]
30. Haug, K.; Cochrane, K.; Nainala, V.C.; Williams, M.; Chang, J.; Jayaseelan, K.V.; O’Donovan, C. MetaboLights: A resource evolving in response to the needs of its scientific community. *Nucleic Acids Res.* **2020**, *48*, D440–D444. [CrossRef]
31. Guijas, C.; Montenegro-Burke, J.R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A.E.; et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* **2018**, *90*, 3156–3164. [CrossRef]
32. Kanehisa, M. KEGG Bioinformatics Resource for Plant Genomics and Metabolomics. In *Plant Bioinformatics; Methods in Molecular Biology*; Edwards, D., Ed.; Humana Press: New York, NY, USA, 2016; Volume 1374. [CrossRef]
33. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; et al. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714. [CrossRef] [PubMed]
34. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem 2023 update. *Nucleic Acids Res.* **2023**, *51*, D1373–D1380. [CrossRef] [PubMed]

35. Caspi, R.; Altman, T.; Billington, R.; Dreher, K.; Foerster, H.; Fulcher, C.A.; Holland, T.A.; Keseler, I.M.; Kothari, A.; Kubo, A.; et al. The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **2014**, *42*, D459–D471. [[CrossRef](#)] [[PubMed](#)]
36. The Metabolomics Workbench. Available online: <https://www.metabolomicsworkbench.org/> (accessed on 27 January 2023).
37. Sumner, L.W.; Lei, Z.; Nikolau, B.J.; Saito, K.; Roessner, U.; Trengove, R. Proposed quantitative and alphanumeric metabolite identification metrics. *Metabolomics* **2014**, *10*, 1047–1049. [[CrossRef](#)]
38. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **2016**, *3*, 160018. [[CrossRef](#)] [[PubMed](#)]
39. Alseekh, S.; Aharoni, A.; Brotman, Y.; Contrepois, K.; D’Auria, J.; Ewald, J.; Ewald, J.C.; Fraser, P.D.; Giavalisco, P.; Hall, R.D.; et al. Mass spectrometry-based metabolomics: A guide for annotation, quantification and best reporting practices. *Nat. Methods* **2021**, *18*, 747–756. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Chapter 3 : Strain specific metabolic responses of diverse *E.coli* lineages to the acquisition of a multi-drug resistant plasmid.

3.1 Abstract

The antibiotic resistance crisis is presenting increasing numbers of infections that are near impossible to treat. Conjugative plasmids are the primary disseminators of resistance genes and have complex interactions with their bacterial hosts which are not fully understood. Metabolomics is an under-utilised tool that can provide valuable molecular insight to phenotypic observation. Metabolomics were conducted on 9 *Escherichia coli* strains from a variety of backgrounds to understand the impact of acquisition of a multi-drug resistant plasmid pLL35. The impact is subtle and strain specific. A few key metabolic functions are affected such as ubiquinone biosynthesis, central energy production and amino acid biosynthesis but the direction and particular pathways highlighted remain strain specific.

3.2 Introduction

Plasmids are the primary facilitators for the worldwide dissemination of antibiotic resistance genes (Carattoli, 2013). *E.coli* cause common infections worldwide and the plasmid-mediated accumulation of multi-drug resistances are continually reducing the number of antibiotics able to combat them (Mathers, Peirano and Pitout, 2015; Stoesser *et al.*, 2016; Dunn, Connor and McNally, 2019). *E.coli* strain group ST131 is an opportunistic pathogen of phylogenetic group B2 of high global concern that causes urinary tract infections (UTIs), bloodstream infections (BSIs), intra-abdominal infections (IAs) and wound infections (Cantón, González-Alba and Galán, 2012; Alhashash *et al.*, 2013; Lee Ventola, 2015; Mathers, Peirano and Pitout, 2015). ST131 is globally the predominant extended spectrum beta lactamase (ESBL) producing *E.coli* and is frequently associated with multidrug resistances, including fluoroquinolone and aminoglycoside resistance (Lahlaoui, Ben Haj Khalifa and Ben Moussa, 2014; Mathers, Peirano and Pitout, 2015; Bevan, Jones and Hawkey, 2017; Peirano and Pitout, 2019). This strain is resistant to most available antibiotics and

forcing hospitals to use last resort antibiotics with increasing frequency such as ertapenem, with decreasing impact (Peirano and Pitout, 2019).

Plasmid mediated horizontal gene transfer can drive the rapid evolution of multidrug resistant bacteria (Hawkey and Jones, 2009; Hall, Brockhurst and Harrison, 2017). Conjugative plasmids can carry large cargoes of accessory genes, including antibiotic resistance genes, which are beneficial to the bacteria in specific environments (Norman, Hansen and Sørensen, 2009). However bacteria-plasmid responses are highly strain dependent. Previous work has shown that both costs and phenotypic effects are dependent on the bacterial strain (Johnson *et al.*, 2015; Lang and Johnson, 2015; Takahashi *et al.*, 2015; Porse *et al.*, 2016a; Alonso-del Valle *et al.*, 2021). Either as a cause or consequence of this the distribution of plasmids can vary significantly between lineages (Benz *et al.*, 2021). Plasmids have natural compatibility host ranges, but even within these species, persistence is dependent on strain specific acquisition costs (Porse *et al.*, 2016a; Prenskey *et al.*, 2021).

The interactions between resistance plasmids and their bacterial hosts are complex. Plasmids induce a diverse range of phenotypic effects in bacteria over and above those encoded by accessory genes. Impacts can include fitness costs, ranging from negligible to severe and significant disruption to cellular processes such as altered expression of hundreds of genes (Coulson *et al.*, 2015; Takahashi *et al.*, 2015; San Millan *et al.*, 2018). Outside of accessory genes plasmids can increase bacterial-encoded antimicrobial resistance functions (Shintani *et al.*, 2010; Takahashi *et al.*, 2015; San Millan *et al.*, 2018; Vasileva *et al.*, 2018), aid in colonisation of bacterial hosts (Gomez-Duarte and Kaper, 1995; Von Bargen *et al.*, 2009; Coulson *et al.*, 2015; Schaufler *et al.*, 2016; Ronin *et al.*, 2017; Ranjan *et al.*, 2018) increase virulence (Song *et al.*, 2013; Patton *et al.*, 2018; San Millan *et al.*, 2018; Vasileva *et al.*, 2018) and facilitate utilisation of novel or alternative energy sources (San Millan *et al.*, 2018; Dunn *et al.*, 2021). These phenotypic traits may promote survival of the bacterium and therefore plasmid vertical transmission (Billane *et al.*, 2022). Other plasmid associated traits meanwhile may promote plasmid fitness in a way that does not necessarily align with bacterial host fitness. For example a reduction of bacterial motility, increase in biofilm formation and inactivation of competitive bacterial T6SS, increasing the likelihood of bacterial cell to cell contact and therefore successful plasmid conjugation (Matsumoto *et al.*, 1998; Valle *et al.*, 2008; Shintani *et al.*, 2010; Parashar

et al., 2013; Lang and Johnson, 2015; Takahashi *et al.*, 2015; Schaufler *et al.*, 2016; Jiang *et al.*, 2017; Vasileva *et al.*, 2018; Huang *et al.*, 2020). Understanding the relationships between plasmid-bacterial interactions across diverse strains can thus help to inform how bacteria respond to plasmid acquisition and importantly, what conditions favour their loss and maintenance.

The success of infection-causing *E. coli* can be partially explained by narrow host range IncF group plasmids, which are large 100-150 kbp plasmids limited to Enterobacteriaceae and have contributed significantly to the resistance levels (Mathers, Peirano and Pitout, 2015). Initial acquisition of IncF plasmids is thought to be costly but some plasmid and strain pairings appear to be more suited than others (Dunn, Connor and McNally, 2019). To promote their own maintenance, IncF plasmids use post-segregational killing and addiction systems but also carry many resistance genes which provide great survival benefits under antibiotic pressure (Mathers, Peirano and Pitout, 2015; Bevan, Jones and Hawkey, 2017).

IncF group plasmids can contain bla_{CTX-M} genes which produce enzymes that inactivate B-lactams (third and fourth generation cephalosporins and monobactams) by hydrolysis, thereby rendering the most frequently prescribed antibiotics in the world impotent (Lahlaoui, Ben Haj Khalifa and Ben Moussa, 2014; Peirano and Pitout, 2019). bla_{CTX-M} genes originate from *Kluyvera* spp. but mobilised and transferred into other *Enterobacteriaceae*, which, since the early 2000s they have proliferated into some of the most treatment resistant infections (Canton and Coque, 2006). Bla_{CTX-M-15}, part of the Bla_{CTX-M-1} subfamily, was first described in 2001 and now the most prevalent variant, alongside Bla_{CTX-M-14} (Bevan, Jones and Hawkey, 2017).

Previous work investigated the impact of MDR plasmid carriage on diverse strains of *Escherichia coli*. The plasmid pLL35, a 106kb incFII ESBL plasmid originally isolated from *Klebsiella pneumoniae*, was introduced to 9 strains spanning a diversity of lineages (Dunn *et al.*, 2021). Dunn *et al.* 2021 showed that while the plasmid was stably maintained in all strains, acquisition induced variable responses at all biological levels studied. Most strains showed very little impact on bacterial growth but for some growth rate was significantly reduced while others actually increased growth rate after acquiring the plasmid. Similarly, the plasmid conferred different levels of resistance to the beta-lactam, Cefotaxime. However when transcriptomics were conducted the

number of genes significantly altered by plasmid acquisition was surprisingly low and did not correlate to the growth impact or resistance (Dunn *et al.*, 2021). There was a consistent subtle differential regulation across the transcriptome, targeting a wide variety of bacterial functions including the cell wall, signal transduction, cell motility, energy production and conversion and carbohydrate transport and metabolism (Dunn *et al.*, 2021). How these changes relate to observed phenotypic responses to plasmid carriage is unclear, meaning that the effects of plasmid carriage on hosts requires further interrogation.

A range of transcriptomic studies showed the most commonly altered genes by plasmid presence were metabolism related, suggesting metabolomic perturbations associated with bacterial acclimatisation to plasmid carriage (Shintani *et al.*, 2010; Lang and Johnson, 2015; Takahashi *et al.*, 2015; San Millan *et al.*, 2018; Vasileva *et al.*, 2018). However, very few metabolomics studies have focused on the impact of plasmid acquisition on a bacterial host (Lang and Johnson, 2015; San Millan *et al.*, 2018). San Millan *et al.* (2018) found in *Pseudomonas aeruginosa*, metabolic evidence of increased RNA synthesis and differential metabolism of glutamine, central energy metabolites such as NADPH and citric acid, and coenzymes such as Pyridoxal-5' phosphate (San Millan *et al.*, 2018). Metabolomics, as the molecular basis of function, can add a level of biological detail inaccessible to genomics or transcriptomics to explain phenotype (Wang *et al.*, 2016; San Millan *et al.*, 2018; Ares-arroyo *et al.*, 2022). The sensitivity of metabolomics can also identify flux in complex networks and add statistical strength to findings with other omics (Ares-arroyo *et al.*, 2022; Radoš *et al.*, 2022).

Contemporary omic research tends to focus on laboratory strains, (Billane *et al.*, 2022) therefore this study emphasised clinically relevant ST131 strains (clades A, B and C) of *Escherichia coli*, as well as the laboratory strain MG1655 and environmental *E.coli* from lineages in which multidrug resistance (MDR) plasmids have never been reported (Dunn *et al.*, 2021). Here we use untargeted metabolomics to obtain a comprehensive, unbiased view of the metabolic profiles of diverse *E.coli* strains after acquisition of a plasmid, pLL35. Originating from *Klebsiella pneumoniae*, this plasmid encodes

multidrug resistances, including beta lactams, aminoglycosides and quinolones (Table 3.1) (Fig 3.1). PLL35 is 106 kb long and belongs to the incFII(K)-9 plasmid group.

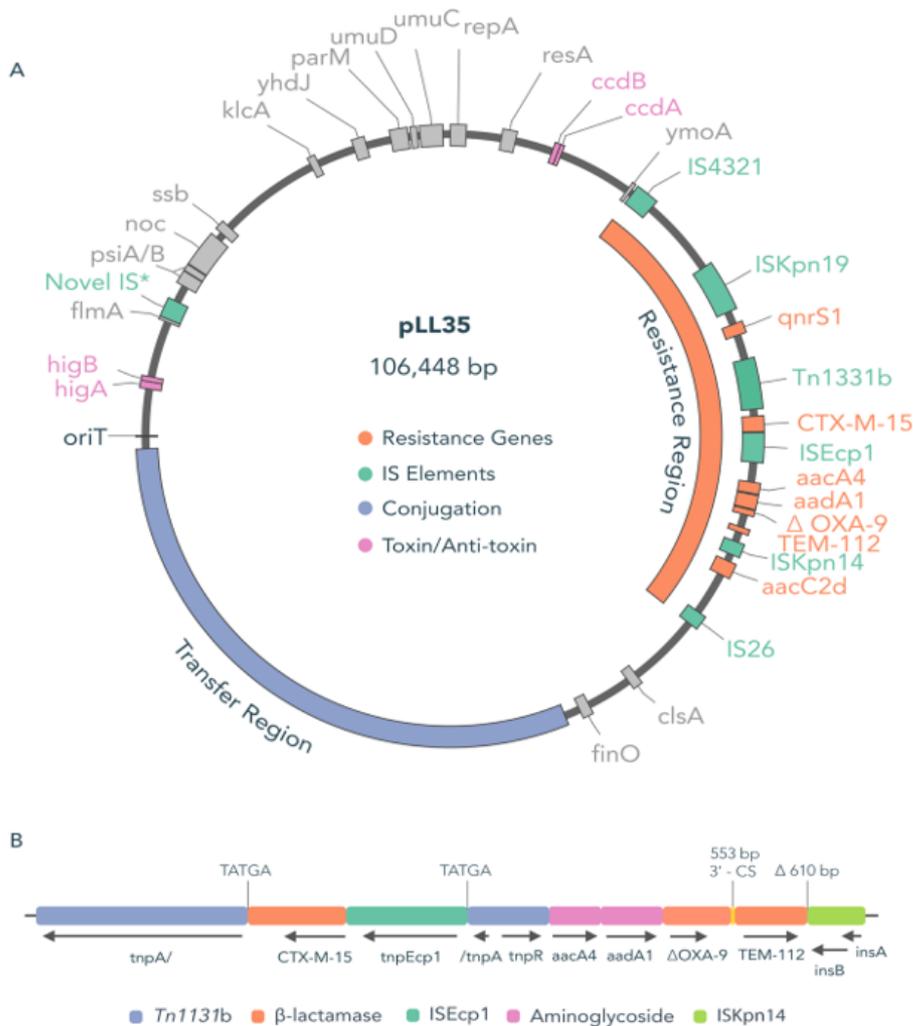


Figure 3.1 Schematic of pLL35, a 106kb IncFII(K)-9 plasmid, originating from *Klebsiella pneumoniae* belonging to ST45. The plasmid encodes full conjugation machinery and contains resistance to beta-lactams (blaCTX-M-15 and blaTEM-112), aminoglycosides (aacA4, aacC2 and aacA1) and quinolones (qnrS1). The OXA-9 gene is truncated by a stop codon. The schematic is the work of Dr. S Dunn (Dunn et al., 2021).

3.3 Methods

Mass Spectrometry

This study used a total of 9 *Escherichia coli* strains (Table 3.1). After undergoing conjugation with the 106kbp pLL35 FII (K)-9 multidrug resistant plasmid (work by Dr. Carrilero) each strain has a plasmid carrying and plasmid free version (Dunn *et al.*, 2021).

Table 3.1 *E.coli* strains used in this experiment and their origins.

Strain	<i>E.coli</i> clonal group	Source
F022	ST131	Clinical bacteremia
F037	ST131	Clinical bacteremia
F047	ST131	Clinical bacteremia
F048	ST131	Clinical bacteremia
F054	ST131	Clinical bacteremia
F0104	ST131	Clinical bacteremia
ELU39	ST1122	River effluent
GU15	ST394	River effluent
MG1655	K-12	Laboratory strain

Full methods are described in chapter 2. The bacteria were grown in nutrient broth at 37°C, 180 rpm to mid-exponential phase, isolated and frozen at -80°C. The samples were prepared for mass spectrometry with chloroform and methanol. Untargeted metabolome analysis was performed using HPLC-QToF MS to identify differentially expressed bacterial metabolites between bacteria with or without the MDR plasmid pLL35. The samples (50µl) were introduced to the Waters G2/G2Si Synapt mass spectrometer. The instrument settings are detailed in the supplementary material (S3). HPLC-QToF MS was performed using the aqueous phase of the samples for DESI MS in positive mode, with a scanning range of 50-1200m/z over 3 minutes.

Data Processing and Analysis

All data processing and analysis followed a standardised methodology based on open-source software, developed and collated into a full guide by the author and peers (<https://untargeted-metabolomics-workflow.netlify.app/> accessed on 27 January 2023) (Parker *et al.*, 2023).

All raw data files were converted to mzML format using the [Proteowizard](#) software MsConvert. [XCMS](#) online was used for peak alignment and retention time correction

(parameters 84500, S3). An average was taken of the technical replicates and output XCMS data tidied in [R](https://untargeted-metabolomics-workflow.netlify.app/) (https://untargeted-metabolomics-workflow.netlify.app/). After peak alignment and retention time correction, over 1000 features were present in every profile. The data was processed to allow comparison of relative metabolite intensities in multigroup analyses across all strains and conditions. Additionally, a series of pairwise analysis by strain were conducted to analyse the relative amount of metabolites present in the 'plasmid carrying' and 'plasmid free' profiles thereby excluding metabolites arising from plasmid transcription and focusing solely on the bacterial metabolome.

Metaboanalyst was used to perform statistical analysis. Data was normalised with pareto scaling (Figure S3.2 and S3.3).

This study employs random forest analysis, a bootstrapping algorithm which combines ensemble learning methods with the decision tree framework to create multiple randomly drawn decision trees from the data and averaging the results. This analysis produces strong predictors of grouping, or treatment, ranked by variable importance. The analysis ran with 7 predictors and 500 trees.

Metabolites are reported if present in all 5 biological replicates and for fold change data, must meet or exceed a threshold of ± 2.0 were reported. Here we define 'significance' as metabolites that have been differentially expressed by a fold change $\geq \pm 2.0$.

Any metabolites highlighted in statistical analysis were putatively identified using [METLIN](#), [KEGG](#) and [ECMD](#). The databases METLIN or ECMD were searched with the m/z values, and must be agreed upon by with the KEGG pathway metabolism map for *E.coli* in order to be reported.

3.4 Results

Strain specific metabolic variability.

Random Forest analysis demonstrated that strain differences dominate the variation between isolates and are greater than the impact of plasmid carriage. Of the 20 most variable masses across the dataset, just 3 were associated with plasmid carriage (Fig 3.2, plots 1, 8 and 9. Table 3.2). Overall global metabolic profiles suggested differences between strains from different ecological backgrounds with environmental and lab strains (ELU39 and GU15) appearing to be distinct from the clinical strains (Fig 3.3) The metabolic signature of environmental strain ELU39 in particular is different to the rest of the strains (Fig 3.3), consistently demonstrated in the discriminating mass bins for the whole dataset (Fig 3.2, plots 4,5,8,9,12,16,17,19). However, variability between strains was driven more by large differences in specific metabolites as overall by strain, the metabolism of plasmid carriers were very similar to plasmid free bacteria (Fig S3.1). For example, in a principal component analysis the majority of the variation along PC1 is driven by clinical strain F022, for which 107 metabolites were greatly upregulated compared to other strains (Fig 3.4), including 4 of the top 20 mass bins as highlighted by a Random Forest analysis (Fig 3.2). F022 carries 8 copies of an IS element which likely account for this very different profile. This interrupted the gene *IrhA*, which is a negative regulator of classes 1, 2 and 3 of the flagellar biosynthesis regulons (Dunn et al 2021).

Plasmid acquisition was characterised by subtle, strain specific variation, affecting a comparatively small proportion of the metabolome; <1% of the metabolome significantly altered in 8 of 9 strains. Environmental strain ELU39 has the largest response to the plasmid with 1.36% of the metabolic profile significantly altered.

The strains represent a continuum of metabolic responses. Some strains showed strong trends towards upregulation (E.g , ELU39 with 17 upregulated and 2 downregulated) while others showed strong downregulation (F047 with 1 upregulated and 11 downregulated) (Fig 3.5). Both of the environmental strains favour upregulation. In other cases strains showed mixed (E.g strain F104 with 3 upregulated and 4 downregulated) or minor responses (E.g F022, F037 and F048 had 3, 2 and 1 significant metabolites respectively) (Fig 3.5).

The specific masses affected by plasmid carriage were largely specific to each strain. A total of 51 unique mass values were significantly differentially up or down regulated by plasmid acquisition, 46 of these were strain specific (Table S3.1). Only 7 unique masses could be identified and among these there is some functional parallelism where the same pathways are implicated in different strains and at times from different metabolites. At a detected mass level 5 masses are significant in up to 5 strains although functional categorization could not always be assigned. For instance, M/z 136.0858 was upregulated in 4 strains F037, F104, ELU39 and GU15, notably over 10-fold in F104 (FC 10.96) and ELU39 (FC 11.74) (Table 3.2). These masses are also not consistent in direction, emphasising strain specificity.

Three main pathways stand out among the metabolites affected by plasmid carriage.

Ubiquinone biosynthesis

4-Hydroxy-3-polyprenyl benzoate (ppm36) was significantly affected in 4 strains (F022, F047, F054 and F104) (Table 3.2, Fig 3.6) and was the first hit in the Random Forest analysis (Fig 3.2 plot 1). As a result of plasmid acquisition this metabolite was upregulated in 3 strains, most notably by nearly 10-fold in F054, (FC 9.473) and downregulated in F047 (FC 0.37) (Table 3.2). 4-Hydroxy-3-polyprenyl benzoate is a key component of the ubiquinone biosynthesis pathway and is a precursory compound to ubiquinol. Furthermore, some of the potential identifications for the M/z 147.06438, which is downregulated in F022, are Benzyl-alcohol and 4-Cresol (ppm 39) (Table 3.2), both part of the toluene degradation pathway which is linked to ubiquinol as an alternative parent compound (Jindrova *et al.*, 2002).

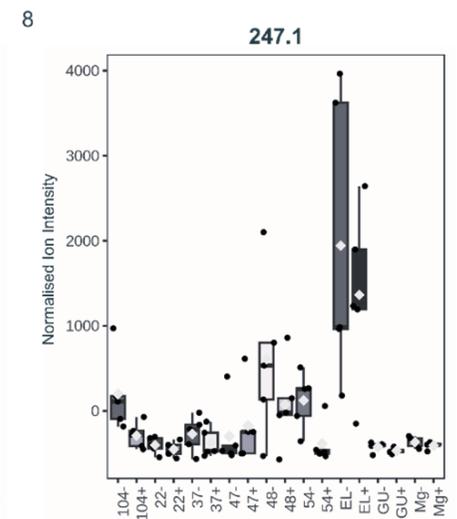
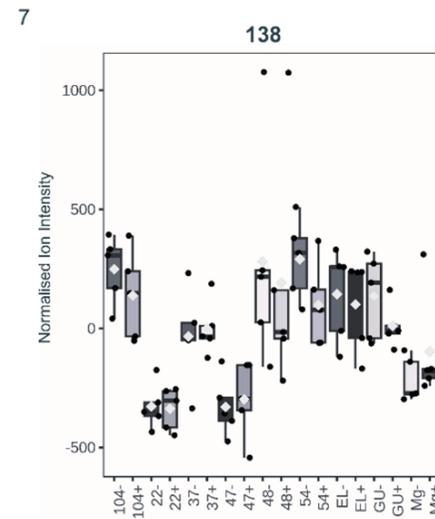
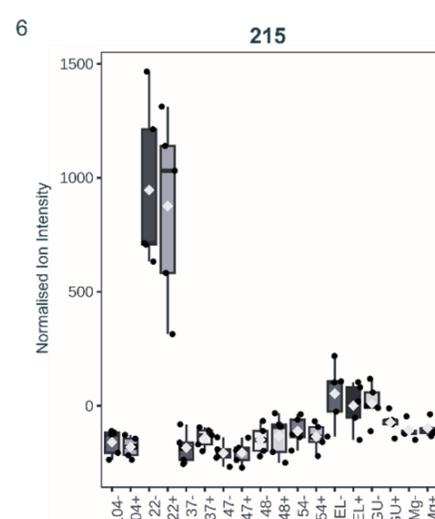
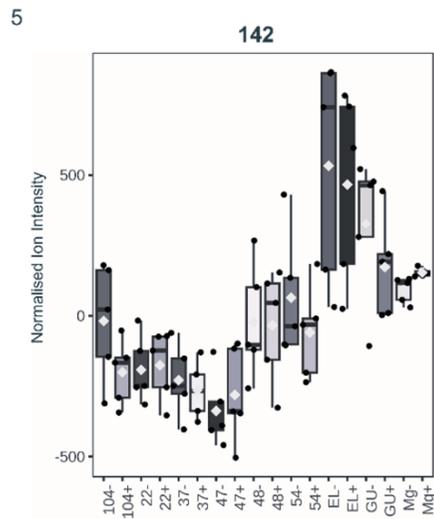
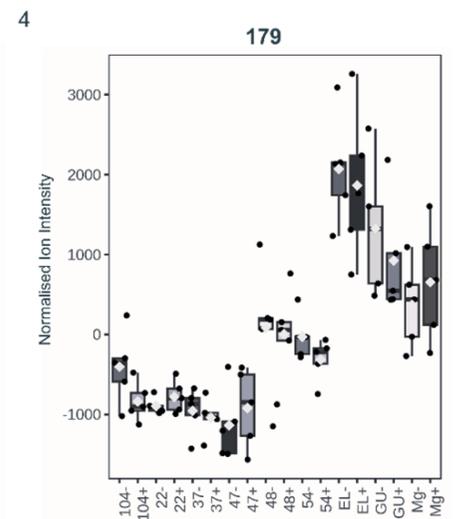
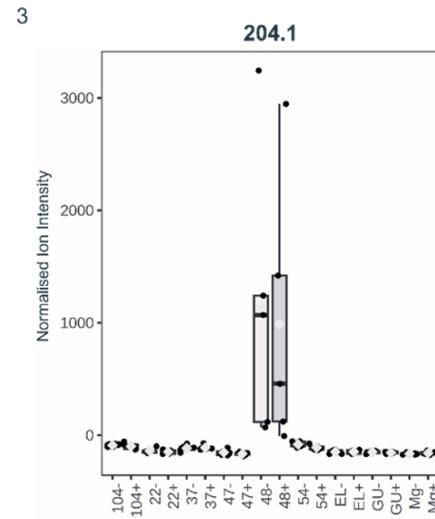
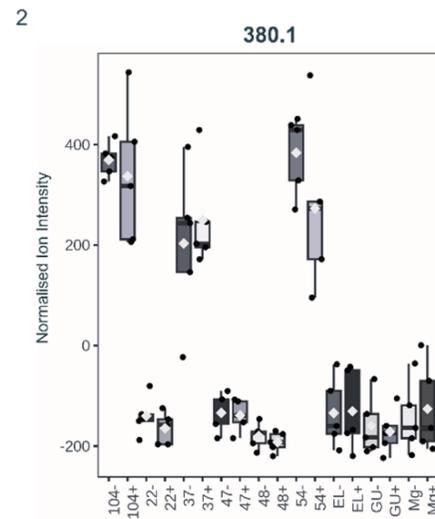
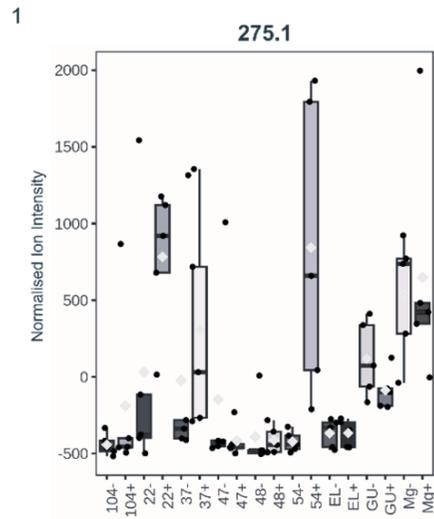
Central energy production

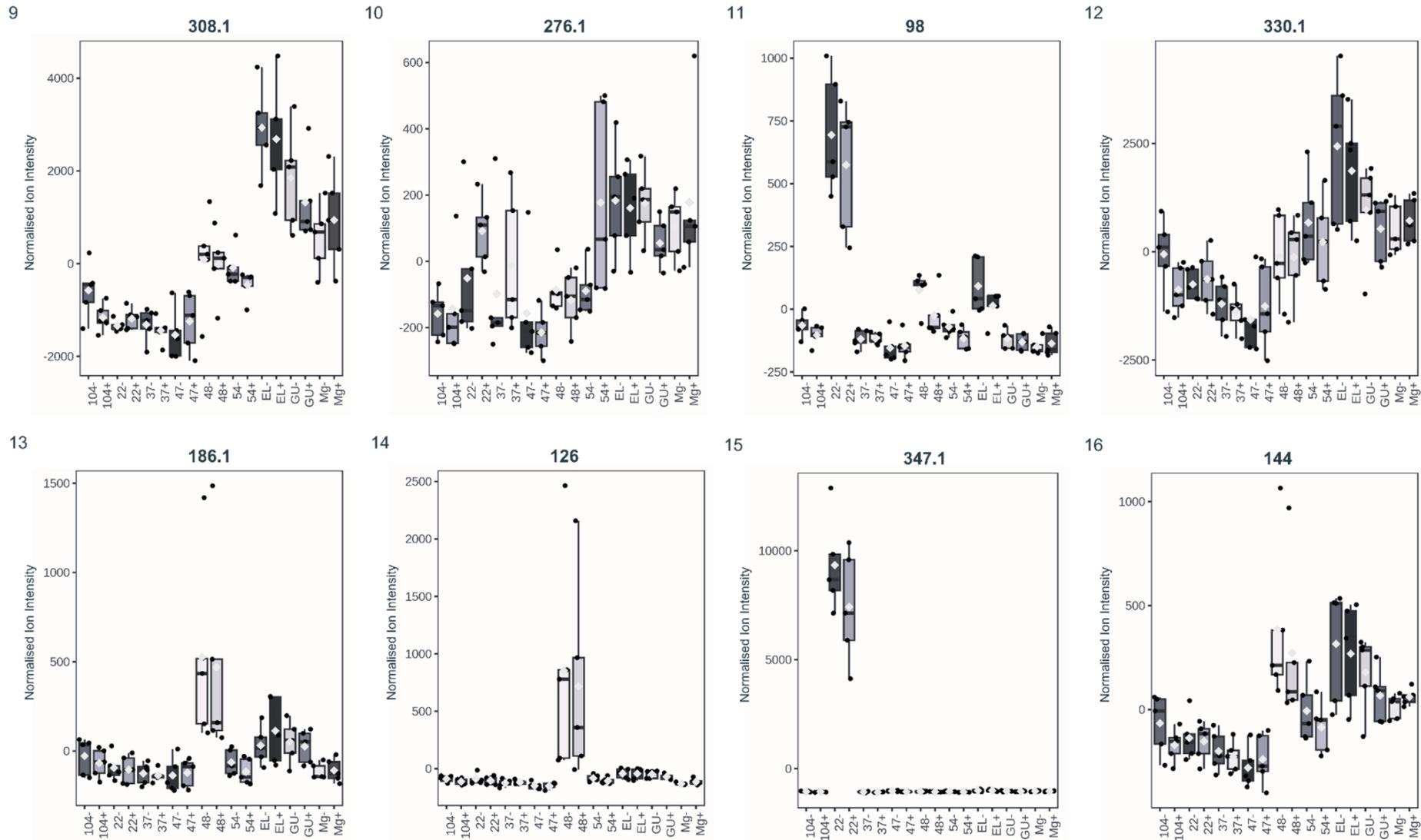
Masses linked to central carbohydrate metabolism were down regulated in 2 strains (F047 and F054) in response to plasmid carriage. The metabolite D-Glyceric acid (ppm 36) was downregulated in F047 (FC 0.45) which is part of the pentose-phosphate pathway (Table 3.2). Methylamine (ppm 5) is upregulated in strain F054 (FC 2.03). This metabolite is part of the methane metabolism pathway but is also linked to D-Glyceric acid as a precursor (Table 3.2) (KEGG).

Amino acid biosynthesis

Amino acid metabolism was associated with several significant metabolites. The metabolite N6-Acetyl-L-Lysine (ppm 25) is part of the lysine degradation pathway and is downregulated in F104 (FC 0.49) but upregulated in MG1655 (FC 2.19) (Table 3.2). There were 2 potential identifications for this mass, the second is decanoate (ppm37), a fatty acid (Ecocyc).

The mass 136.13162 was identified as creatinine (ppm 22) and was significantly altered in 5 strains: F048 (FC 0.11), F054 (FC 15.62), F104 (FC 2.74), GU15 (FC 0.09) and ELU39 (FC 0.29). This metabolite is linked to the arginine and proline metabolism pathway (Table 3.2 Fig 3.6). The response to plasmid carriage in this metabolite was the most extreme observed across the strain collection, being both the most highly upregulated (>15 fold in clinical strain F054) and the most strongly down regulated (>10 fold in environmental strain GU15). The plasmid induces downregulation of creatinine in both environmental strains and F048 but is upregulated in the other two clinical strains F054 and F104.





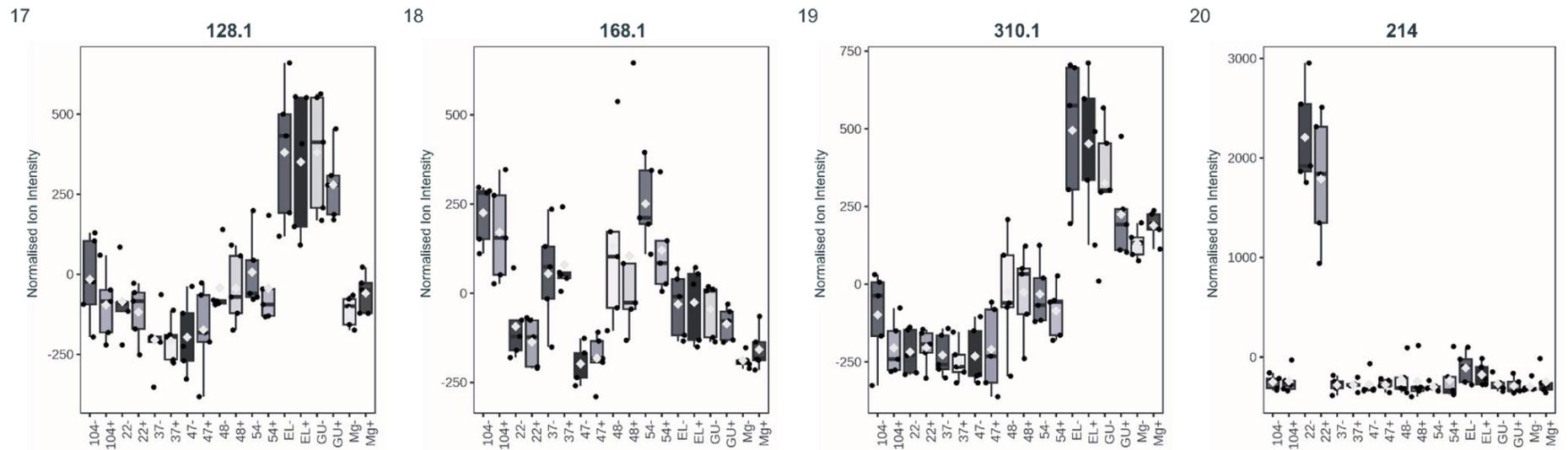


Figure 3.2. The average ion intensities for discriminatory mass bins under 500 Da for all strains with and without the plasmid, from the results of a Random Forest analysis. Strain names are shortened as follows: F104 = 104, F022 = 22, F037 = 37, F047 = 47, F048 = 48, F054 = 54, ELU39 = EL, GU15 = GU and MG1655 = Mg. The symbols – and + indicate plasmid free and plasmid carrying respectively.

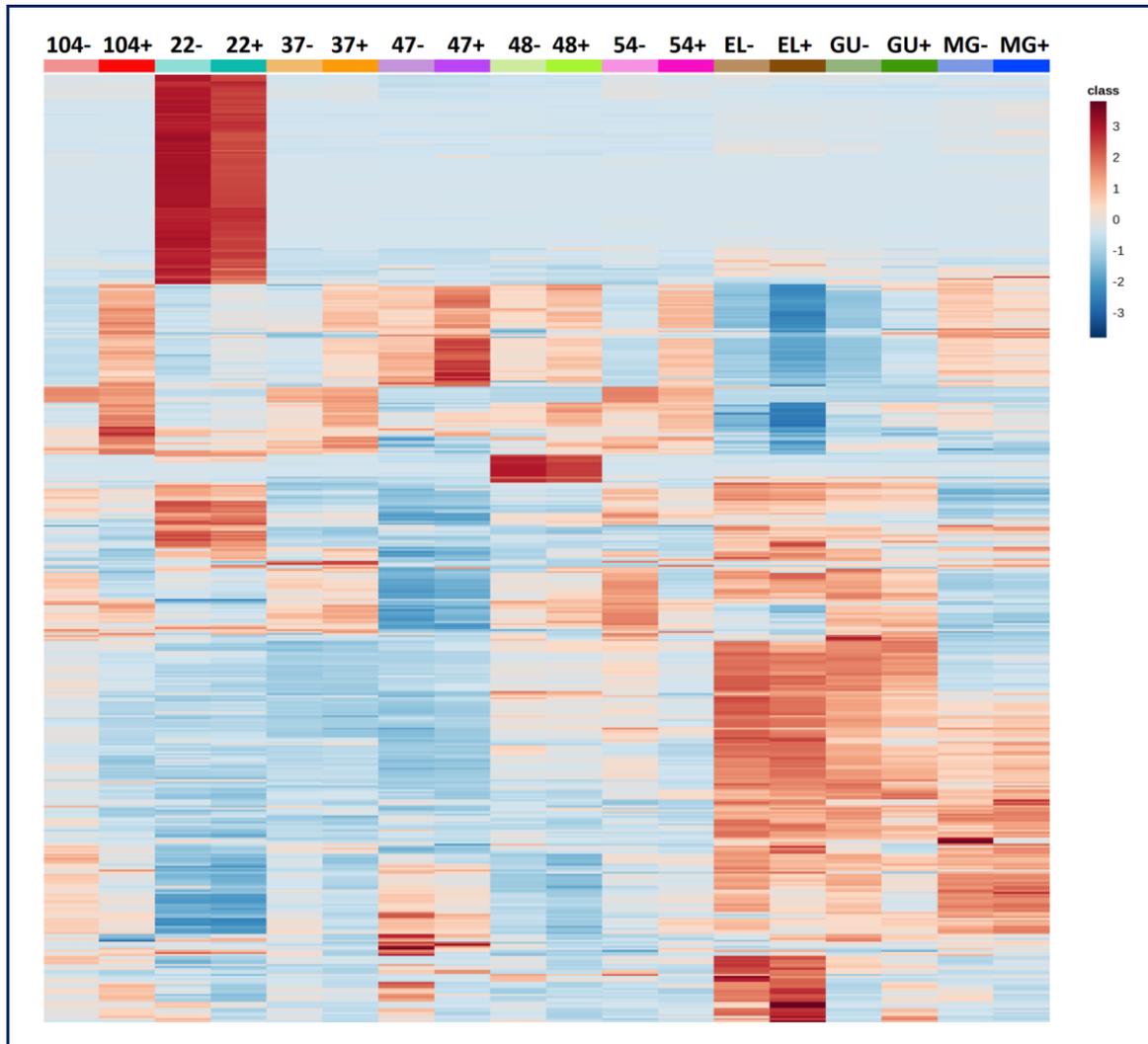


Figure 3.3 A Heatmap showing the full metabolic profiles of 9 strains of *Escherichia coli*. Intensities of detected masses have been scaled -3 to +3. Each strain is in a plasmid free (-) and pLL35 plasmid containing (+) conditions. The biological replicates for each strain have been pooled.

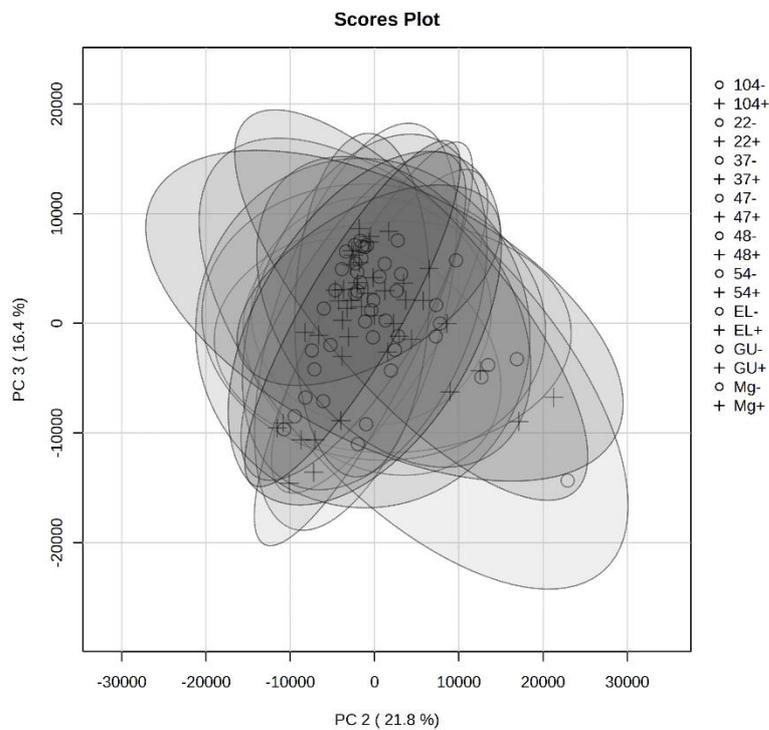
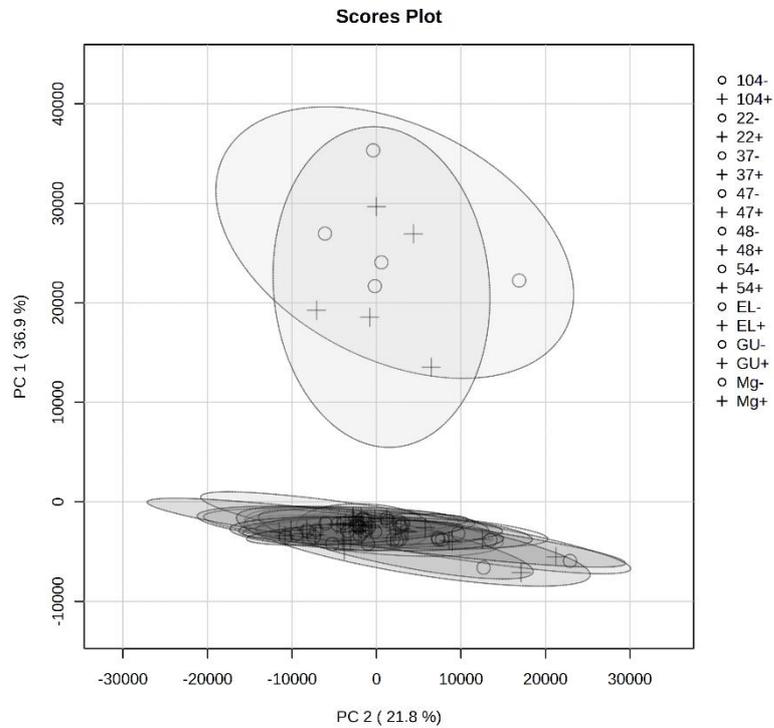
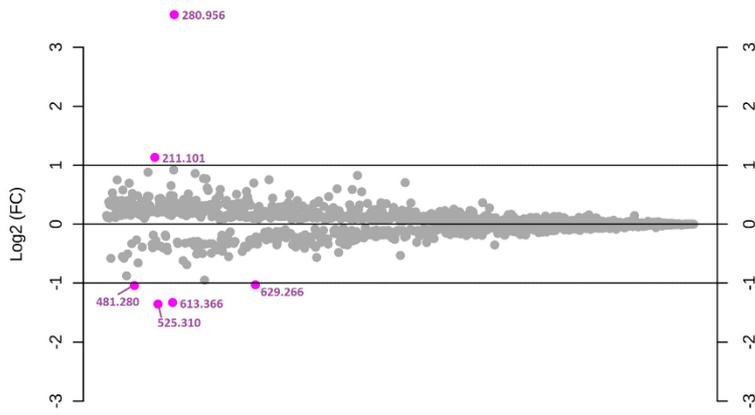


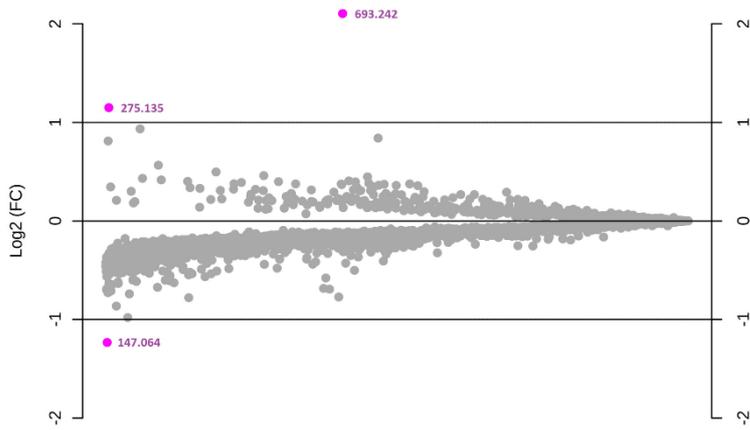
Figure 3.4 – Principle component analysis (PCA) of all metabolomic data. Principle component (PC) 1 (36.9%) and 2 (21.8%) (above) and PC 3 (16.4%) and PC2 (21.8%) (below). A circle represents plasmid free strains and the + symbolises plasmid carrying. The separation in the above plot is caused by the lineage difference of strain F022.

MG1655



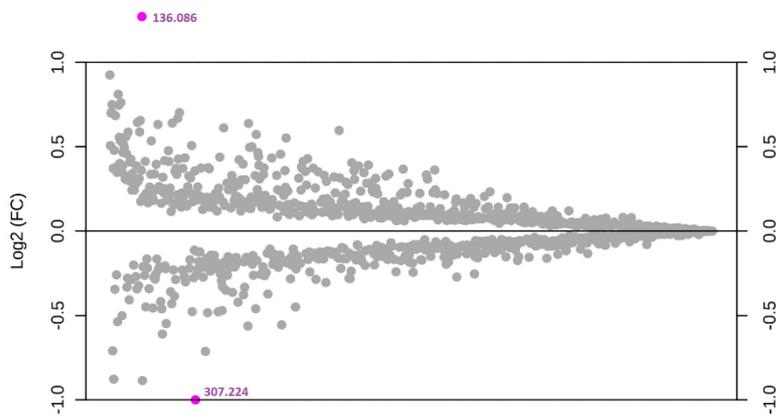
Peaks(mz/rt)

F022



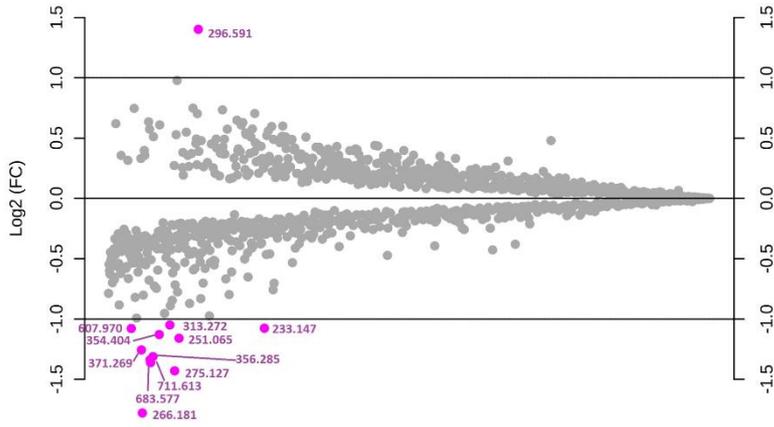
Peaks(mz/rt)

F037



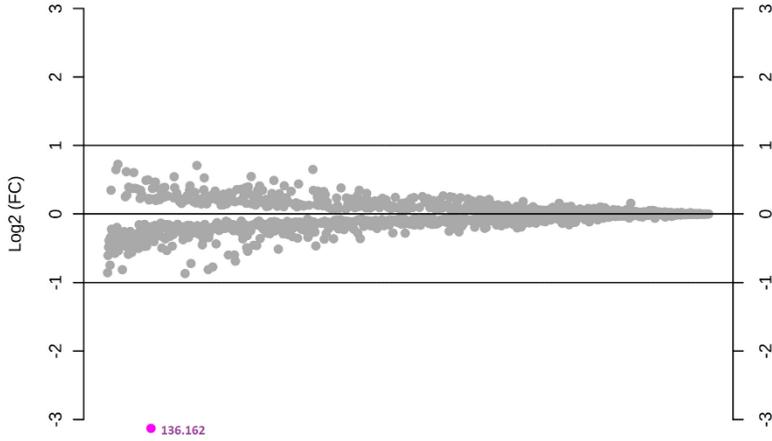
Peaks(mz/rt)

F047



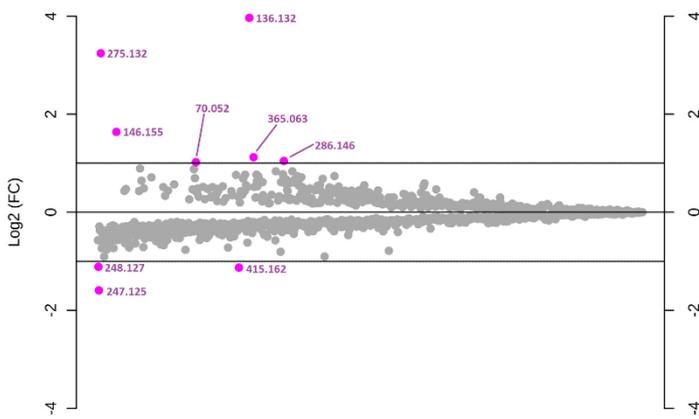
Peaks(mz/rt)

F048



Peaks(mz/rt)

F054



Peaks(mz/rt)

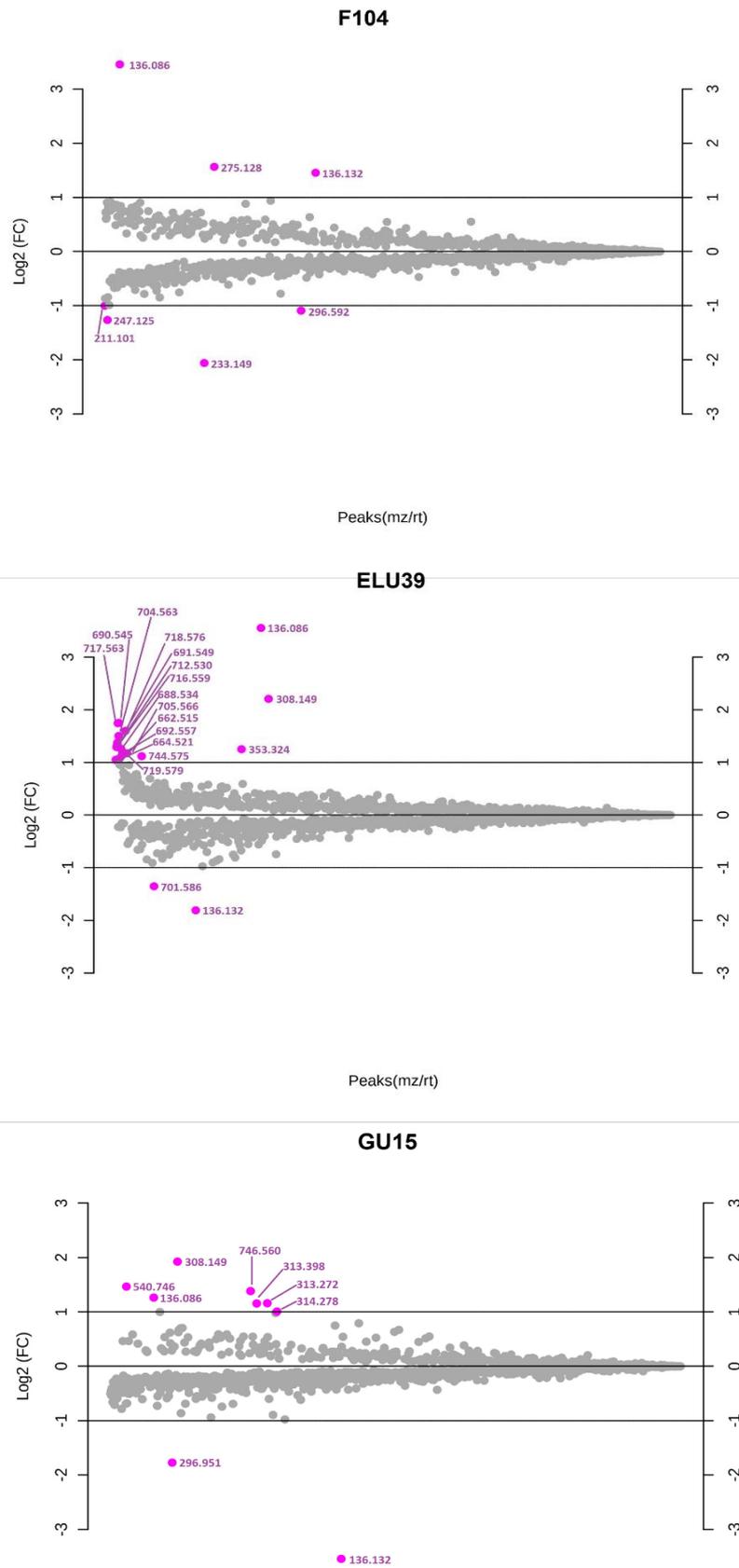


Figure 3.5 Fold change maps highlighting the masses (m/z) that cross the fold change threshold of 2.0 when the plasmid free and plasmid carrying bacterial metabolisms are compared on an individual strain basis.

Table 3.2 – Results from the pairwise ‘plasmid free’ and ‘plasmid containing’ analyses by strain. Mass values with a Fold Change of 2.0 or greater between the 2 treatment groups. Those with putative Identifications have been included here, full table in supplementary (Table S3.1)

Strain	Mass	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	Fold Change	log2(FC)
F054	70.005	Methylamine	C00218	5	M+K	70.0054	31.0422	CH5N	2.0299	1.0214
F048	136.13162	Creatinine	C00791	22	M+Na	136.0481	113.0589	C4H7N3O	0.11428	-3.1294
F054	136.13162	Creatinine	C00791	22	M+Na	136.0481	113.0589	C4H7N3O	15.621	3.9654
F104	136.13162	Creatinine	C00791	22	M+Na	136.0481	113.0589	C4H7N3O	2.7381	1.4532
ELU39	136.13162	Creatinine	C00791	22	M+Na	136.0481	113.0589	C4H7N3O	0.28535	-1.8092
GU15	136.13162	Creatinine	C00791	22	M+Na	136.0481	113.0589	C4H7N3O	0.085993	-3.5396
F022	147.06438	Pyrazinic acid	C19915	10	M+Na	147.0165	124.0273	C5H4N2O2	0.42509	-1.2342
		Benzyl-alcohol	C00556	39	M+K	147.0207	108.0575	C7H8O		
		4-Cresol	C01468	39	M+K	147.0207	108.0575	C7H8O		
F104	211.1012	N6-Acetyl-L-Lysine	C02727	25	M+Na	211.1053	188.1161	C8H16N2O3	0.49985	-1.0004
		Decanoate	C01571	37	M+K	211.1077	171.1385	C10H19O2		
MG1655	211.1012	N6-Acetyl-L-Lysine	C02727	25	M+Na	211.1053	188.1161	C8H16N2O3	2.1908	1.1315
		Decanoate	C01571	37	M+K	211.1077	171.1385	C10H19O2		
F054	248.12694	Tetradecanoate		11	M+Na	248.1747	225.1855	C14H25O2	0.4625	-1.1125
F047	251.06503	D-Glyceric acid	C00258	36	M+H	251.0074	250.0002	C6H10CaO8	0.4477	-1.1594
F047	275.12669	4-Hydroxy-3-polyprenylbenzoate	C05848	36	M+H	275.1642	274.1569	C17H22O3	0.37122	-1.4297
F104	275.12766	4-Hydroxy-3-polyprenylbenzoate	C05848	36	M+H	275.1642	274.1569	C17H22O3	2.9549	1.5631
F054	275.13172	4-Hydroxy-3-polyprenylbenzoate	C05848	36	M+H	275.1642	274.1569	C17H22O3	9.473	3.2438
F022	275.13489	4-Hydroxy-3-polyprenylbenzoate	C05848	36	M+H	275.1642	274.1569	C17H22O3	2.2185	1.1496



Figure 3.6 – KEGG map of *E. coli* metabolites significantly altered by the presence of the pLL35 plasmid. The putative identifications have been summed across all 9 strains and highlighted.

3.5 Discussion

Plasmid carriage induces subtle, strain specific metabolome rewiring in *E.coli*.

Untargeted metabolomic analysis showed a diverse set of *E.coli* strains have highly variable metabolic profiles. The response of these strains to the acquisition of an MDR ESBL plasmid, pLL35, is subtle and strain specific that affected both divergent and shared metabolic pathways. Common pathways impacted by the plasmid related to central energy metabolism, amino acid metabolism and ubiquinone biosynthesis. Among common metabolites, direction of impact was also contingent on the specific strain.

Untargeted metabolomic analysis showed that the response to the acquisition of the plasmid pLL35 by *E.coli* is contingent upon the specific strain, aligning with previous research. In the study conducted by Dunn et al 2021; the plasmid was stably maintained in all strains, but otherwise induced variable responses in resistance to cefotaxime, growth and transcriptomic effects (Dunn *et al.*, 2021). There was a consistent subtle differential regulation across the transcriptome, the most altered of which were associated with; the cell wall, signal transduction, cell motility, energy production and conversion and carbohydrate transport and metabolism (Dunn *et al.*, 2021). While each strain has a unique metabolic profile, there are some common functions. Broadly, the results indicate that the plasmid affected central energy metabolism and alternative energy sources for cellular function and growth.

Key metabolic processes affected by plasmid acquisition across strains were the ubiquinone biosynthesis pathways, central energy metabolism and amino acid metabolism (Table 3.2). In 4 strains; F022, F047, F054 and F104, the ubiquinone biosynthesis pathway is highlighted. Ubiquinone is one of three quinones in *E.coli* and is involved in aerobic respiration (Aussel *et al.*, 2014; Nitzschke and Bettenbrock, 2018). Ubiquinone biosynthesis is a highly conserved pathway because it is essential for aerobic growth, gene regulation and oxidative stress adaptation. The other quinones are for anaerobic respiration, allowing *E.coli* to switch between the two (Nitzschke and Bettenbrock, 2018; Arias-cartin *et al.*, 2023). 3 of the strains; F022, F054 and F104 had this pathway upregulated, suggesting a greater rate of energy

production. In F047 this same pathway was downregulated, but this compliments the transcriptomics which revealed an upregulation of anaerobic genes, suggesting a switch to anaerobic respiration (Dunn *et al.*, 2021).

F047 has the largest metabolic response of the clinical strains and is skewed towards downregulation. A decrease in D-glyceric acid may reflect lower levels of glycolysis, the pentose phosphate pathway or the TCA cycle which could be an indication of stress-induced energy conservation (Wang *et al.*, 2016; Zhao *et al.*, 2019). F047 also had lower levels of cefotaxime resistance, reduced growth and the biggest transcriptional response (Dunn *et al.*, 2021). The transcriptome indicated an upregulation of stress responses, metabolic transport genes, anaerobic respiration genes and the repressor (*marR*) of the *mar* antibiotic resistance and oxidative stress response regulon, although this regulon itself was not significantly impacted (Dunn *et al.*, 2021).

Methane metabolism, fatty acid metabolism and metabolism of alternate sources of energy were also affected pathways (Table 3; Jindrova *et al.*, 2002; Guo AC *et al.*, 2012). Fatty acid metabolism is highlighted in strains F054, F104 and MG1655, and is linked to ubiquinone (Agrawal *et al.*, 2017). Long chain fatty acids are a source of energy for *E.coli* but induce oxidative stress, something that ubiquinone helps to counteract through reduction into ubiquinol because it is an electron carrier in the electron transport chain (Agrawal *et al.*, 2017). Of the strains in which the ubiquinone biosynthesis pathway was highlighted, only F054 and F104 demonstrate significance in fatty acid metabolism.

Nonetheless, parallel responses are seen repeatedly throughout this dataset in the metabolites impacted by the plasmid. Selection pressures tend to affect bacteria by targeting function and not genes, so it follows that there are commonly affected pathways across strains (Wang *et al.*, 2016). While common areas of the metabolism are affected, the direction of change caused by the plasmid is strain specific. In multiple cases the same detected mass, significant in several strains, has displayed both upregulation and downregulation, depending on specific strain. For example, creatinine was identified in 5 strains, upregulated in 2 strains and downregulated in 3 (Table 3.2).

Together the metabolic responses seen in this study may be reflective of stress and energy conservation. Signs of stress in *E.coli* often involve energy metabolism, as energy is a requisite resource for overcoming the underlying stressor (Zhao *et al.*, 2019). Among others, arginine and proline metabolism, methane metabolism, glycolysis and the TCA cycle are all metabolic markers of stress in *E.coli* (Zhao *et al.*, 2019) and together may indicate a strategy of energy conservation or utilisation of alternative energy sources. This ability to acclimate using a metabolic network that can produce energy from lots of different sources and have a responsive regulatory system to enable this is an *E.coli* survival strategy (Ishii *et al.*, 2007).

The metabolomic data revealed trends in response to the plasmid related to ecological background of the strains that were not apparent from the transcriptomics. For instance the metabolome of the environmental strains ELU39 and GU15 responded similarly to each other in masses affected and direction of change. Interestingly ELU39 had a negligible change in growth rate and no significantly differentially expressed genes and yet had the biggest metabolic response of the dataset. GU15 has a greater reduction in growth than ELU39 but a smaller response metabolically, suggesting ELU39 may be better suited to respond to energy demands of the plasmid (Figure 3.5) (Dunn *et al.*, 2021). The different response of the two environmental strains may be partially explained by the likelihood that these strains and the plasmid would encounter each other. Using naturally occurring combinations gleans different results, therefore predictions based on laboratory strains may not be transferable to real-world scenarios (Alonso-del Valle *et al.*, 2021).

It is generally thought that when *E.coli* acquires a plasmid there will be a reduction in fitness thanks to re-directed transcription and translation machinery away from chromosomal operons, and that following this the plasmid will be degraded or adaptive mutations will ameliorate costs (Dunn, Connor and McNally, 2019). However overall, Dunn *et al* reports a very low proportion of the transcriptome differentially expressed upon plasmid acquisition, and a small but varied growth response. GU15 and F047 displayed reduced growth and F022 and F037 increased growth rate after acquiring the plasmid. However, the scale of the transcriptomic response did not correlate to the growth impact (Dunn *et al.*, 2021). In all cases, only a small fraction of the metabolome is changed in response to plasmid acquisition, concurring with the subtle effect seen

in the transcriptomics. Since conjugation of plasmids belonging to *incF* are limited by strain it cannot be ruled out that pLL35 has encountered *E.coli* before, despite originating in *K. pneumoniae*, and has therefore already evolved for minimal disruption (Benz *et al.*, 2021). Alternatively, the clinical strains may already be more generally adapted to plasmid carriage, if they carried other plasmids before conjugation of pLL35, although at the time of conjugation, the strains contained no other MDR plasmids (Dunn *et al.*, 2021). Plasmid presence can increase bacterial permissiveness to further plasmid carriage (San Millan, Heilbron and MacLean, 2014).

A strain specific, low impact of plasmid acquisition is not surprising given a growing body of research demonstrating minimal fitness consequences of MDR plasmids on *E.coli* (Fischer *et al.*, 2019). This is repeated in other studies featuring non-lab strains, highlighting the importance of studying these interactions in ecologically relevant contexts (Alonso-del Valle *et al.*, 2021).

Hernando-amado suggests genetic elements dictate the extent of a resistance plasmid's impact and that effects of plasmid acquisition could be positive as well as deleterious (Hernando-amado *et al.*, 2017). *E.coli* has a robust metabolic network that can re-route in response to perturbations in its environment, consequently metabolic impact is smaller than might be expected, which could also explain why the plasmid does not have large phenotypic effects (Ishii *et al.*, 2007). Furthermore, a multi-omic study noted that changes in mRNA and regulation of energy within *E.coli* were effective enough that metabolite levels did not change a great deal, nor in particular patterns (Ishii *et al.*, 2007).

Identification of metabolites is an area widely recognised as challenging due to highly variable MS techniques, machines and individual sample properties like retention time in LC/MS. Although databases are improving, it is currently lagging behind other omics technologies (Blaženović *et al.*, 2018; Chaleckis *et al.*, 2019). This raises the question of whether *m/z* values which are only a few decimal places different, such as 136.0858 and 136.13162 (Table S3.1) could be the same metabolite, or a very similar one. Although mass spectrometry measurements are very precise, secondary metabolites are highly variable in structure, and overlap of the strains in which these *m/z* are differentially present (± 2.0 FC) compared to plasmid containing groups may not be

a coincidence, especially given the robust replication of the measurements. Although the patterns in this data are backed by robust replication, better identification may strengthen trends seen so far.

Together the results of untargeted metabolomics show acquisition of a large multidrug resistant plasmid in a diverse set of *E.coli* strains has subtle effects that impact central energy metabolism in strain specific ways that may indicate metabolic stress.

Future research should consider variable responses to plasmid acquisition in a community context. Genotype by environment interactions allow the maintenance of even deleterious alleles (Hasik and Siepielski, 2022), which can help to explain why variation in fitness for plasmid and bacteria combinations allows for plasmid persistence. Theoretically, even when a plasmid induces a fitness cost, plasmid by bacteria fitness variation and bacterial genotype by environment interactions allow for the maintenance of plasmids within a population. In eukaryotic hosts, parasite-mediated variation in fitness is a driving factor of selection (Hasik and Siepielski, 2022), the genetic parasitism of plasmids may play a similar role in bacteria. A small plasmid impact on its bacterial host has a place in the spectrum between being too costly which drives the loss of the edge accessory genes provided as alternative resistance mechanisms are favoured (Bottery, Wood and Brockhurst, 2017), and positive impacts which certainly exist (Dunn *et al.*, 2021) but may be less likely.

Chapter 4 : Strain specific/limited parallelism responses to coevolution and antibiotic selection of diverse *E.coli* lineages with a multidrug resistant plasmid.

4.1 Abstract

The antibiotic resistance crisis is costing increasing numbers of lives due to the dissemination of multidrug resistant bacterial strains. This study conducts untargeted metabolomics on 3 *E.coli* strains from the endpoint of an evolutionary experiment having co-evolved with a plasmid, with and without cefotaxime selection. Metabolic profiles varied by strain but showed parallelism across evolutionary treatment within strains. The clinical strain had some of the most variable metabolic profiles. cAMP was a notable exception to the unique metabolic impacts across strains. Coevolution with a plasmid altered amino acid metabolism and energy utilisation, while antibiotic selection displayed stress responses related to the beta-lactam mechanism of action.

4.2 Introduction

Antibiotic resistance is a worldwide crisis set to claim increasing numbers of human lives with untreatable infections (Lee Ventola, 2015). Conjugative plasmids are the primary disseminators of resistance genes thus exacerbating this crisis. Of particular concern are multidrug resistance (MDR) plasmids that can carry multiple resistance genes often against different antibiotic classes and transfer horizontally across bacterial species barriers (Hall, Brockhurst and Harrison, 2017). The emergence of MDR lineages can therefore occur in a single evolutionary event, presenting an extreme clinical challenge to find optimal treatment and to keep antibiotics effective for the long term (MacLean and San Millan, 2019; Peirano and Pitout, 2019; Wu *et al.*, 2019).

Nevertheless, acquiring an MDR plasmid often comes at a fitness cost to the cell, such costs may arise due to a wide range of causes including physiological burden and/or genetic conflicts. The persistence of plasmids in bacterial populations has historically been termed the 'plasmid paradox' because evolutionary theory suggests plasmid loss

should occur over time due to purifying selection acting on the cost of carriage (Harrison and Brockhurst, 2012).

Many studies to date have shown that this paradox is often resolved via the co-evolution of bacteria and plasmids. Specifically genetic amelioration of costs (Brockhurst and Harrison, 2022), where compensatory mutations to ameliorate the cost of plasmid carriage occur in the chromosome, plasmid or both (Hall *et al.*, 2021). Amelioration can take the form of mutations to increase fitness, lose costly functions and to resolve genetic conflict (Porse *et al.*, 2016b; Hernando-amado *et al.*, 2017; Hall *et al.*, 2021). An example of this was revealed in *E.coli* carrying a tetracycline resistance plasmid, where plasmid costs were compensated by mutations to acquire chromosomal resistance and impairing plasmid encoded resistance machinery (Bottery, Wood and Brockhurst, 2019). Compensation can also occur through changes to transcription. For example through increased virulence, colonisation or energy production (Ranjan *et al.*, 2018; San Millan *et al.*, 2018; Vasileva *et al.*, 2018; Billane *et al.*, 2022). Coevolution with plasmids in *P. aeruginosa* induced fitness costs but these were rapidly and repeatedly compensated by mutations to a bacterial regulatory system (GacA/GacS), which then counteracted the genetic conflict, an example of a combination of genetic mutation with downstream transcriptional effects (Harrison *et al.*, 2015; Hall *et al.*, 2021). Compensation can sometimes take place through metabolic mechanisms, as demonstrated in *Haemophilus influenzae*. Coevolution with a ColE1-like plasmid pB1000 demonstrated amelioration could in part be achieved by host alterations to the expression of certain metabolic pathways, such as amino acid metabolism (Ares-arroyo *et al.*, 2022).

As an alternative solution to the paradox, the plasmid may also enable horizontal transfer through differential regulation by increasing biofilm formation and conjugation and decreasing bacterial motility and anti-competitor systems, potentially costly mechanisms to the bacterial host (Parashar *et al.*, 2013; Jiang *et al.*, 2017; Venanzio *et al.*, 2019; Huang *et al.*, 2020; Billane *et al.*, 2022).

Compensatory evolution has now been observed across a wide range of bacteria-plasmid associations, but the physiological basis of amelioration is poorly understood.

Moreover, whether the same pathways of amelioration for a given MDR plasmid are taken in genetically diverse bacterial strains is unclear.

This chapter uses untargeted metabolomics to better understand the evolutionary responses of *E.coli* strains following the acquisition of a multidrug resistant plasmid, pLL35 encoding the extended-spectrum beta lactamase Bla_{CTX-M-15} which provides resistance to cefotaxime. In a previous study, 5 genetically diverse *E.coli* lineages carrying pLL35 were experimentally evolved for approximately 700 generations with or without cefotaxime selection, alongside plasmid free controls (Carrilero, Dunn and Moran, 2023). Here, the metabolomes of the endpoint evolved populations from 3 of the *E.coli* strains: the lab strain MG1655, the clinical strain F022, and the environmental strain ELU39 are analysed. In each of these strains, the initial gain of pLL35 caused a fitness cost associated with strain specific transcriptional changes (Dunn *et al.*, 2021).

After 700 generations of evolution the plasmid was maintained in all populations and MG1655 and F022, but not ELU39, had improved in fitness relative to the plasmid carrying ancestor. Populations evolved with sub-MIC antibiotic selection with cefotaxime had a higher performance across growth kinetics than evolved plasmid free clones (Carrilero, Dunn and Moran, 2023). Resistance levels to cefotaxime did not vary with strain or selection and were largely unchanged but decreased in some plasmid carrying replicates of MG1655 due to upregulation of H-NS and downregulation of the resistance gene Bla_{CTX-M-15} (Carrilero, Dunn and Moran, 2023).

Genome sequencing of the endpoint evolved clones revealed a range of functions targeted by selection in plasmid-carriers not observed in the plasmid-free controls, suggesting that these may be involved in compensatory evolution. Among chromosomal mutations in plasmid carriers evolved with and without antibiotic selection, 22% of single nucleotide variations (SNVs) were parallel. The number of nonsynonymous mutations remained consistent across strains and evolutionary conditions. 14 loci were non-synonymously mutated in one or more evolved clones, involving genes related to cellular metabolism, regulation of mobile genetic elements, and conjugation (Carrilero, Dunn and Moran, 2023).

Many of these functions were involved in metabolism. The functions of the loci mutated in more than one independently evolved clone affected core metabolic functions such as glycerol metabolism, the arginine transporter in F022 and ELU39, fatty acid metabolism in ELU39 and stress response sigma factor in MG1655. MG1655 had several strain specific mutations in aromatic compound metabolism. Other single SNVs mostly related to transcriptional control or metabolism. In MG1655 and ELU39, the transcriptional regulator controlling expression of the NhaA Na⁺/H⁺ antiporter protein was mutated, which controls intracellular pH (Dimroth, 1987; Carrilero, Dunn and Moran, 2023).

Transcriptional analysis of the endpoint evolved MG1655 clones revealed extensive downregulation in plasmid carriers, impacting functions such as cellular structure and motility, DNA damage response, efflux mechanisms, lipopolysaccharide (LPS) production, outer membrane function, and biofilm formation (Carrilero, Dunn and Moran, 2023).

This chapter builds on the work in Chapter 3 which investigated the immediate metabolic response to plasmid acquisition, extending this analysis to understand how evolution acts to resolve plasmid-mediated disruptions of cellular metabolism. This work found that plasmid acquisition had a subtle, strain specific effect. Relatively among the strains, the clinical strain had a low response, and the lab strain had an intermediate response, however both were <1% of the recorded metabolome. The environmental strain ELU39 had the highest metabolic response to plasmid acquisition, which impacted 1.36% of the metabolome. Upon acquisition of the plasmid the transcriptional response was also low, but much more significant in the evolved lines, therefore more pronounced metabolic effects may be expected. Of the identified functions in the previous chapter, there was no commonality between the 3 strains, suggesting *E.coli* from different lifestyles may have different metabolic responses to co-evolution with a plasmid.

Given the predominance of metabolic functions as targets of putative compensatory evolution following gain of pLL35 across all 3 strains (Carrilero, Dunn and Moran, 2023), this chapter conducted untargeted metabolomics on 3 evolved strains from the end point of the evolution experiment; clinical F022 strain, environmental ELU39 strain and lab strain MG1655. More generally, the work in this chapter expands the

application of metabolomics to experimental evolution, where previously it has proven valuable in giving a more comprehensive and nuanced view of evolutionary responses than genomics alone (San Millan *et al.*, 2018; Ares-arroyo *et al.*, 2022).

4.3 Methods

Mass Spectrometry

This study used a total of 3 *Escherichia coli* strains representing 3 environmental backgrounds (Table 4.1).

Table 4.1 Strain types and origins of the *E.coli* strains used in this chapter

Strain	<i>E.coli</i> clonal group	Source
F022	ST131	Clinical bacteremia
ELU39	ST1122	River effluent
MG1655	K-12	Laboratory strain

These strains represent the endpoint of an evolution experiment conducted by Dr. [Laura Carrilero](#) (please see Carrilero, Dunn, Moran 2023 for full methodology). 5 independent colonies from each of the *E.coli* strains were isolated as the ancestral state from which all further cultures were taken. The 106kb FII(K)-9 multidrug resistance plasmid pLL35 was conjugated from *Klebsiella pneumoniae* in static culture into the *E.coli* strains to obtain 5 independent transconjugants per strain. Liquid NB cultures were inoculated with the transconjugants and grown at 37°C for 24 hours to create replication. The MIC varied by strain; F037 had the lowest at ~750mg/L, and several other strain exceeded 2000mg/L (Fig1 of Dunn *et al.*, 2021). The evolution experiment ran 5 independent lines of the strains with 4 biological replicates; plasmid-free clones as the control group, and plasmid-carrying clones in 2 conditions, with or without 4mg/ml cefotaxime treatment. This resulted in 75 lines that were serially transferred daily for a total of 84 days.

The 5 clones of each evolved line of the 3 *E.coli* strains used here were, with 4 biological replicates, grown in nutrient broth at 37°C, 180rpm to mid-exponential phase, isolated and frozen at -80°C. The samples were prepared for mass spectrometry with chloroform and methanol following the method described in chapter 2.

The samples (50µl) were introduced to the Waters G2/G2Si Synapt mass spectrometer. The instrument settings are detailed in the supplementary material (S4). HPLC-QToF MS was performed using the aqueous phase of the samples for DESI MS in positive mode, with a scanning range of 50-1200m/z over 3 minutes.

Data Processing and Analysis

All data processing and analysis followed methodology developed into a user-friendly guide and based on open-source software (Parker et al., 2023) (<https://untargeted-metabolomics-workflow.netlify.app/> accessed on 27 January 2023).

All raw data files were converted to mzML format using the [Proteowizard](#) software MsConvert. [XCMS](#) online was used for peak alignment and retention time correction (parameters detailed in supplementary_location). An average was taken of the technical replicates and output XCMS data tidied in [R](#) (<https://untargeted-metabolomics-workflow.netlify.app/>).

Metaboanalyst was used to perform statistical analysis. Data was normalised with pareto scaling (Figure S4.2).

This study employs random forest analysis, a bootstrapping algorithm which combines ensemble learning methods with the decision tree framework to create multiple randomly drawn decision trees from the data and averaging the results. This analysis produces strong predictors of grouping, or treatment, ranked by variable importance. The analysis ran with 7 predictors and 1000 trees.

Metabolites are reported if present in all 5 biological replicates and for fold change data, must meet or exceed a threshold of +/- 2.0 were reported. Here we define 'significance' as metabolites that have been differentially expressed by a fold change \geq +/- 2.0.

Any metabolites highlighted in statistical analysis were putatively identified using [METLIN](#), [KEGG](#) and [ECMD](#). The databases METLIN or ECMD were searched with the m/z values, and must be agreed upon by with the KEGG pathway metabolism map for *E.coli* in ordered to be reported.

4.4 Results

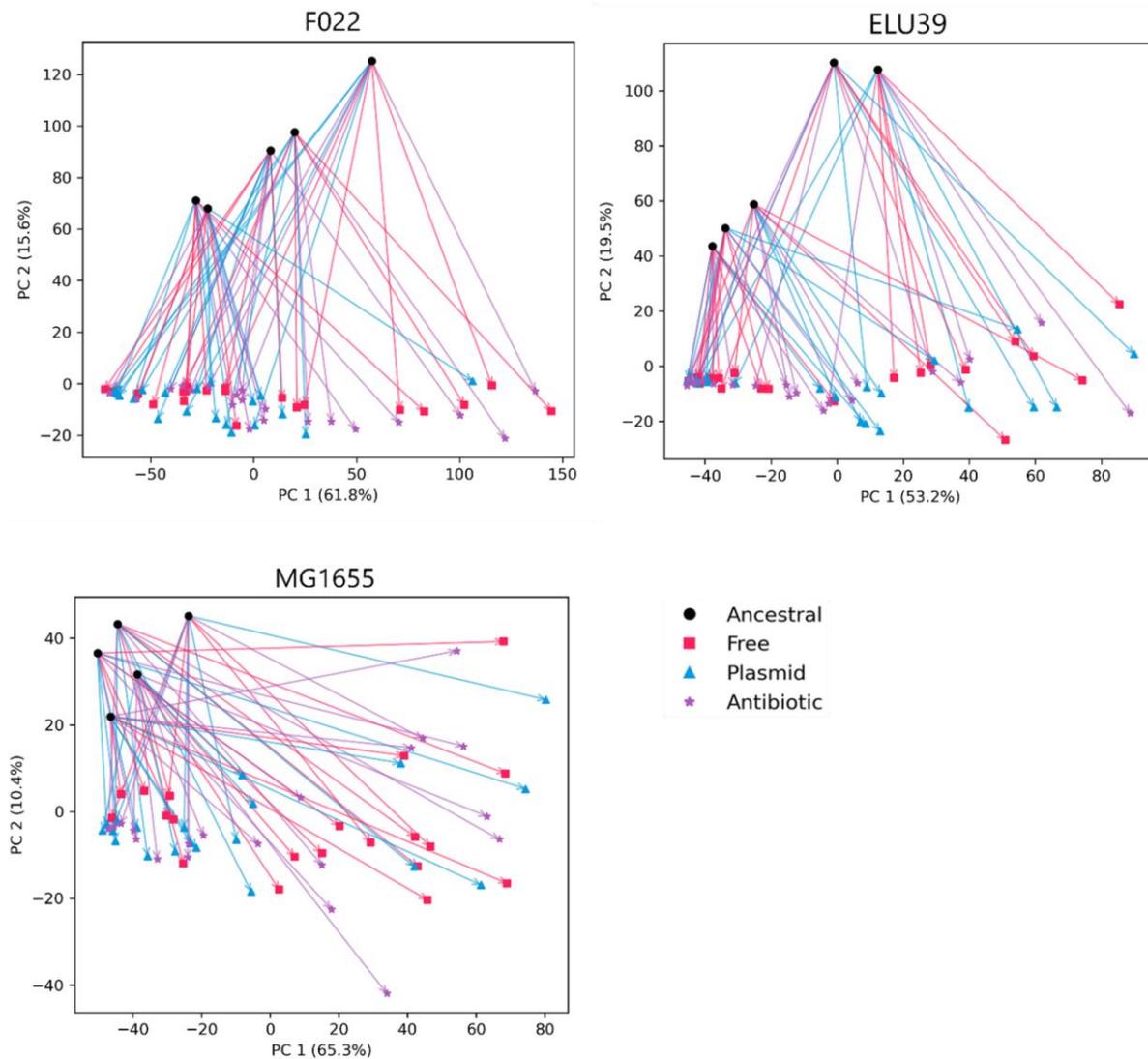


Figure 4.1 trajectory principle component analysis of each strain, grouped by evolution conditions; plasmid free (pink squares) plasmid carrying (blue triangles) and plasmid carrying exposed to cefotaxime (purple stars) and compared against the ancestral state (black circles).

This study performed an untargeted metabolic analysis on the F022, ELU39 and MG1655 strains, 5 replicate lines of each having been evolved for ~700 generation in one of 3 conditions: plasmid free, plasmid carrying and plasmid carrying in the presence of cefotaxime. The metabolome was measured in the same nutrient rich lab medium as the strains had evolved in, without antibiotics therefore alterations seen here are the underlying changes to the metabolic network after antibiotic selection rather than a current response to antibiotic stress.

How has metabolism changed over evolutionary time compared to the ancestor?

The endpoint evolved metabolomes had diverged from their respective ancestors following ~700 generations. Evolutionary paths varied between strains (Fig 4.1, PC2) and among replicates (Fig 4.1, PC1), but the evolved lines from different evolutionary treatment groups overlapped, suggesting that the metabolic signal associated with adaptation to the lab environment overwhelmed plasmid or antibiotic mediated responses.

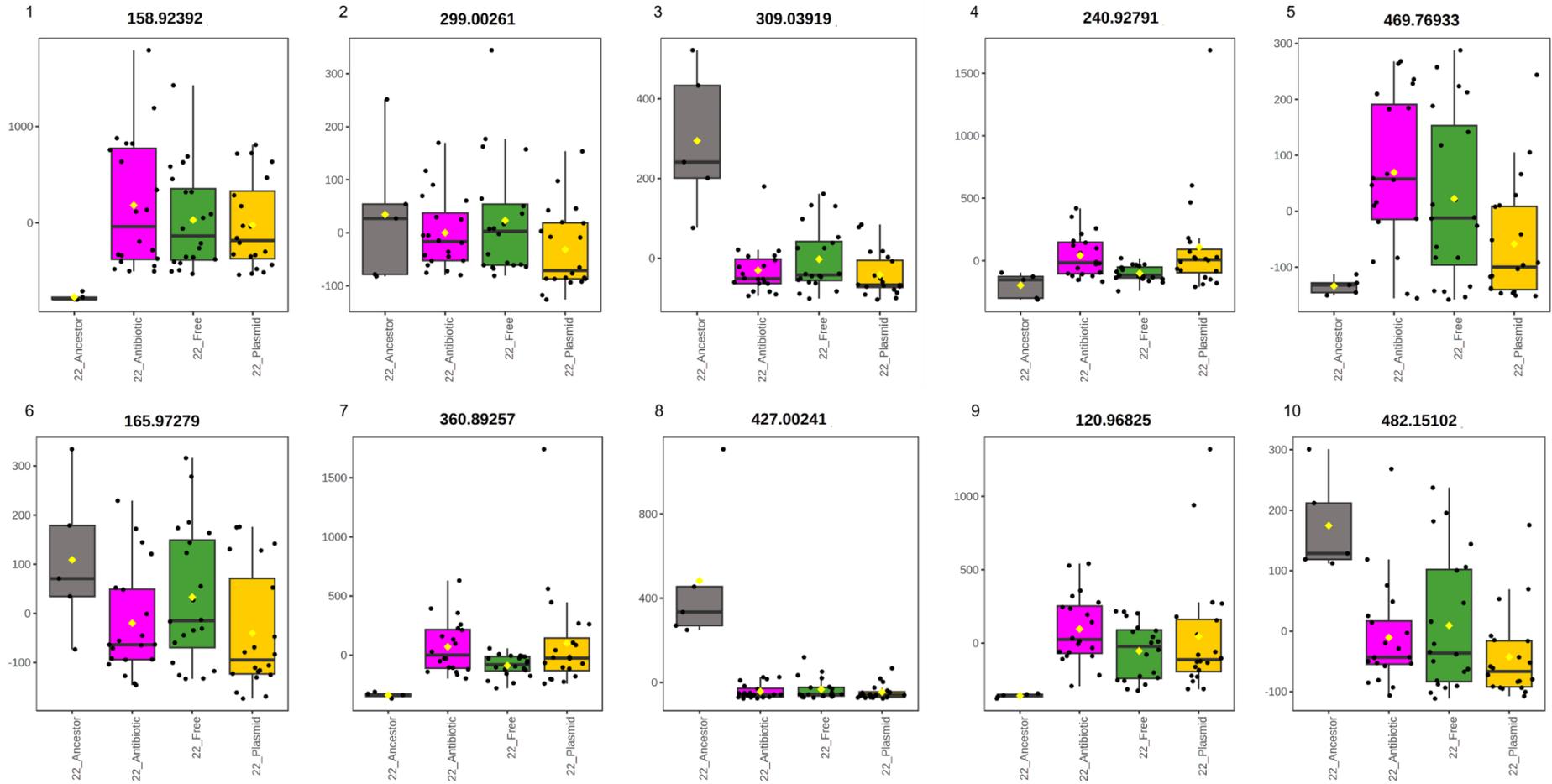
A random forest analysis was conducted to determine the top 20 metabolites driving the variation between the ancestral metabolic state and the evolved lines from each of the 3 evolutionary treatment groups per strain (Fig 4.2, 4.3 & 4.4). When comparing the metabolomic data of the ancestral and evolved states of each strain, affected metabolites showed common functions, namely amino acid biosynthesis, glycolysis and pyrimidine biosynthesis indicating changes to transcription, translation and energy metabolism (Tables 4.2, 4.3, 4.4). All of the strains showed increased levels of amino acid biosynthesis or energy metabolism in the evolved lines. For F022 and MG1655 this agrees with the growth kinetics, conducted as part of the evolution experiment, which were improved after co-evolution with the plasmid (Carrilero, Dunn and Moran, 2023).

Pyrimidine biosynthesis metabolites are a higher intensity in the ancestor for F022 (Fig 4.2 plots 3 and 8) but higher in the evolved lines for ELU39 (Fig 4.3 plot 8). Pyrimidine biosynthesis is essential for synthesis of the bases thymine, cytosine and uracil (Berg *et al.*, 2019). In evolved lines, metabolites involved in this pathway were lower intensity in the clinical strain F022 and higher in the environmental strain. Yet dGDP or ADP, part of deoxyribonucleic acid biosynthesis, is much lower in the evolved environmental strain ELU39. Both of these pathways are essential for transcription and translation, making it difficult to determine a direction for these processes.

Amino acid biosynthesis metabolites are at higher intensities in the endpoint clones for F022 (Fig 4.2, plots 16, 19, 20) and MG1655 (Fig 4.4 plot 11) versus their ancestor. Metabolites involved in energy metabolism, namely glycolysis and the TCA cycle have higher intensities in evolved lines for ELU39 (Fig 4.3, plots 3, 7) and MG1655 (Fig 4.4

plot 15) versus their ancestor. ELU39 also has lower intensities for a metabolite relating to DNA and RNA biosynthesis in the evolved lines (Fig 4.3 plot 11) versus their ancestor. These patterns suggest that after evolution, the bacteria have increased energy and amino acid biosynthesis and altered transcription and translation; however, the differences between ancestral and treatment groups may represent general adaptation of the strains to lab conditions rather than treatment-specific responses.

F022



F022

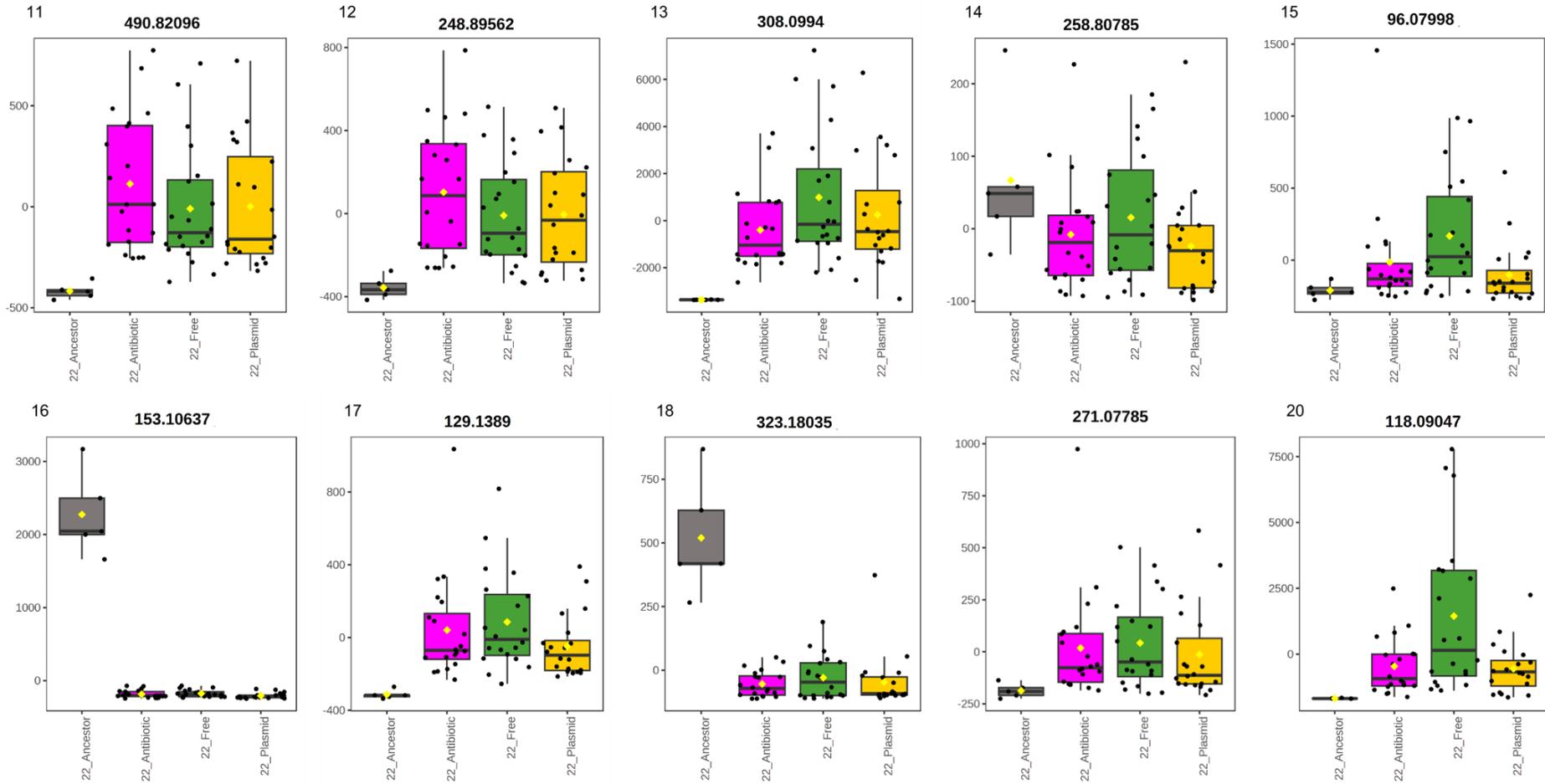


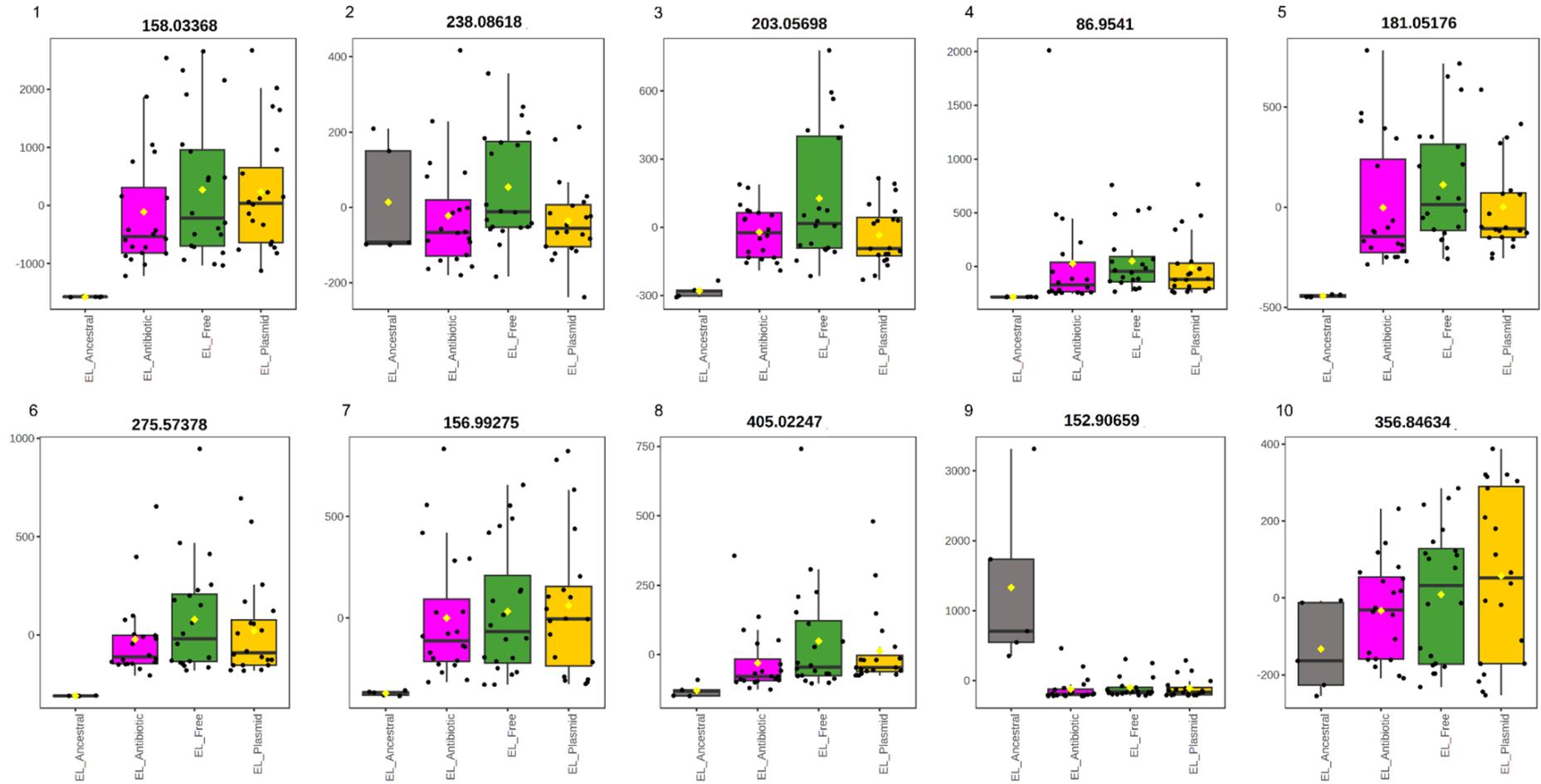
Figure 4.2. Random Forest Analysis showing the top 20 mass values under 500Da which cause the most variation between the ancestral state of *E. coli* strain F022 and the post-evolution clones that underwent 3 treatments of evolution; plasmid free, plasmid carrying and plasmid carrying with exposure to an antibiotic, cefotaxime.

Table 4.2. Random forest analysis rest of top 20 highest ranking masses under 500Da discriminating the ancestral and evolved groups of *E.coli* strain F022. Masses have been putatively identified using ECMDB and KEGG.

F022	Mass	Mean Decrease Accuracy	ID	KEGG	MODULE	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	Function
1	158.9239	0.000664									
2	299.0026	0.000604	Alpha-D-glucose 6-phosphate	C00668	M00001	33	M+K	298.9929	260.0297	C6H13O9P	Glycolysis
			D-Allose 6-phosphate	C02962		33	M+K	298.9929	260.0297	C6H13O9P	
			D-Myo-inositol (1)-monophosphate	C01177		33	M+K	298.9929	260.0297	C6H13O9P	
			-Myo-inositol 4-phosphate	C03546		33	M+K	298.9929	260.0297	C6H13O9P	
			D-Tagatose 6-phosphate	C01097		33	M+K	298.9929	260.0297	C6H13O9P	
			beta-D-Glucose 6-phosphate	C01172		33	M+K	298.9929	260.0297	C6H13O9P	
			Glucose 6-phosphate	C00092		33	M+K	298.9929	260.0297	C6H13O9P	
			Fructose 1-phosphate	C01094		33	M+K	298.9929	260.0297	C6H13O9P	
			Myo-inositol 1-phosphate	C04006		33	M+K	298.9929	260.0297	C6H13O9P	
			Fructose 6-phosphate	C00085	M0001	33	M+K	298.9929	260.0297	C6H13O9P	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes
			D-Mannose 1-phosphate	C00636		33	M+K	298.9929	260.0297	C6H13O9P	
			Glucose 1-phosphate	C00103		33	M+K	298.9929	260.0297	C6H13O9P	
			Galactose 1-phosphate	C00446		33	M+K	298.9929	260.0297	C6H13O9P	
3	309.0392	0.000567	Glycineamideribotide	C03838	M00048	21	M+Na	309.0458	286.0566	C7H15N2O8P	De-novo purine biosynthesis
			dUMP	C00365	M00938	29	M+H	309.0482	308.041	C9H13N2O8P	Pyrimidine deoxyribonucleotide biosynthesis
4	240.9279	0.000546									
5	469.7693	0.000508									
6	165.9728	0.000481									
7	360.8926	0.000448									
8	427.0024	0.000441	Uridine 5'-diphosphate	C00015	M00938 M00053	26	M+Na	426.9914	404.0022	C9H14N2O12 P2	Pyrimidine deoxyribonucleotide

					M00052						and ribonucleotide biosynthesis, deoxyribonucleotide synthesis
9	120.9683	0.000424									
10	482.151	0.000417									
11	490.821	0.000396									
12	248.8956	0.000375									
13	308.0994	0.000368	Glutathione	C00051	M00118	27	M+H	308.0911	307.0838	C10H17N3O6 S	Glutathion biosynthesis
14	258.8079	0.000365									
15	96.07998	0.000353									
16	153.1064	0.000337	Agmatine	C00179	M00133	31	M+Na	153.1111	130.1218	C5H14N4	Polyamine biosynthesis
17	129.1389	0.000332									
18	323.1804	0.000327	Retinal	C00376		10	M+K	323.1772	284.214	C20H28O	
19	271.0779	0.000325	4-(Glutamylamino) butanoate	C15767	M00136	32	M+K	271.0691	232.1059	C9H16N2O5	GABA biosynthesis
			N2-Succinyl-L-ornithine	C03415	M00879	32	M+K	271.0691	232.1059	C9H16N2O5	Arginine succinyltransferase pathway
20	118.0905	0.000317	Betaine	C00719	M00555	31	M+H	118.0868	118.0868	C5H12NO2	Serine and threonine biosynthesis
			L-Valine	C00183	M00019	36	M+H	118.0863	117.079	C5H11NO2	branched chain amino acid metabolism valine- isoleucine metabolism

ELU39



ELU39

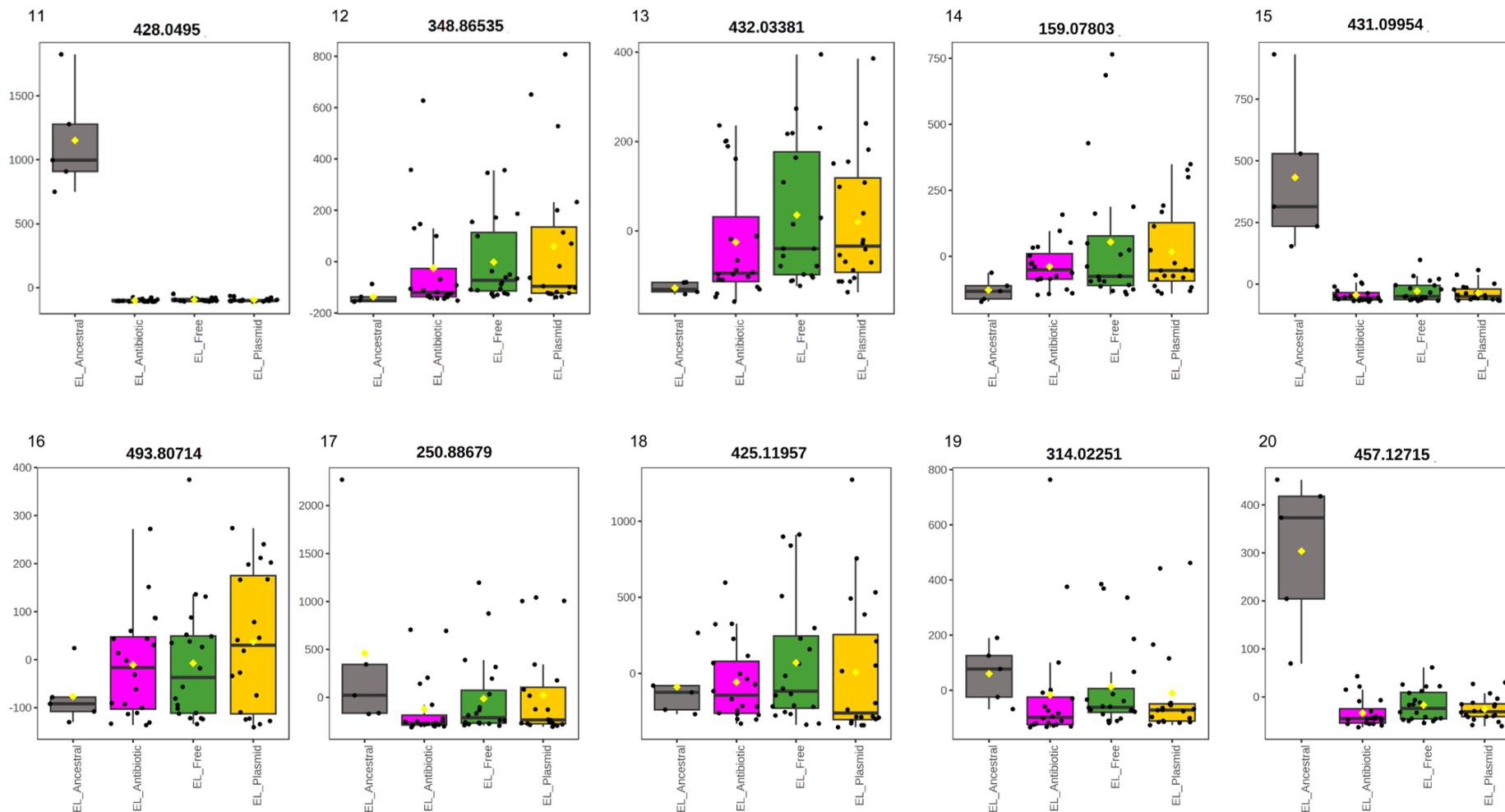


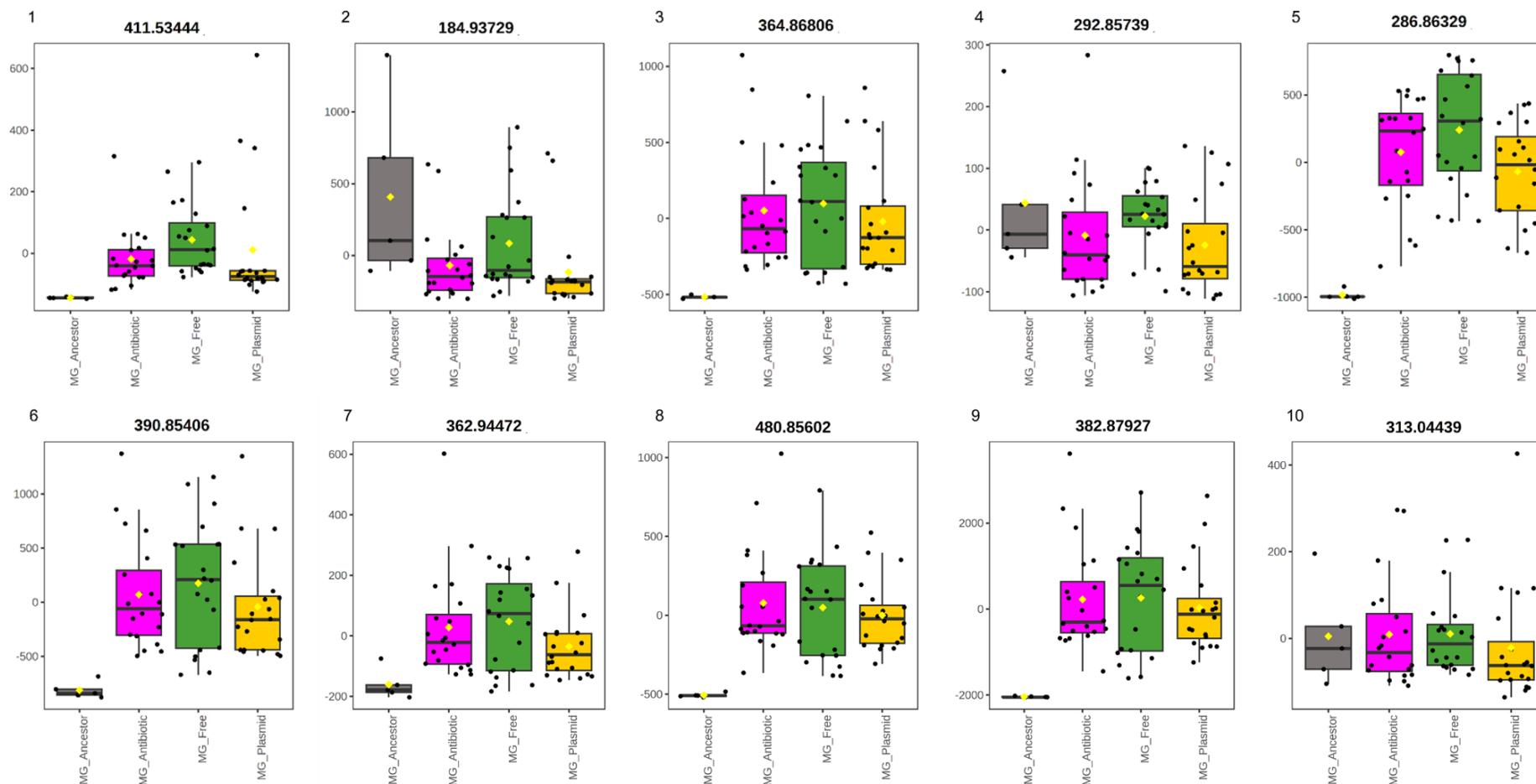
Figure 4.3. Random Forest Analysis showing the top 20 mass values under 500Da which cause the most variation between the ancestral state of *E. coli* strain ELU39 and the post-evolution clones that underwent 3 treatments of evolution; plasmid free, plasmid carrying and plasmid carrying with exposure to an antibiotic, cefotaxime.

Table 4.3. Random forest analysis rest of top 20 highest ranking masses under 500Da discriminating the ancestral and evolved groups of *E.coli* strain ELU39. Masses have been putatively identified using ECMDB and KEGG.

ELU39	Mass	Mean Decrease Accuracy	ID	KEGG	MODULE	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	Function
1	158.0337	0.000636									
2	238.0862	0.000612	Dyspropterin	C03684		31	M+H	238.0935	237.0862	C9H11N5O3	Tetrahydrobiopterin biosynthesis
3	203.057	0.000565	L-Rhamnonate	C01934		22	M+Na	203.0526	180.0634	C6H12O6	
			D-Allose	C01487		22	M+Na	203.0526	180.0634	C6H12O6	
			Inositol	C00137		22	M+Na	203.0526	180.0634	C6H12O6	
			alpha-D-Glucose	C00267	M00001 M000549 M00909	22	M+Na	203.0526	180.0634	C6H12O6	
			D-Fuctose	C00095		22	M+Na	203.0526	180.0634	C6H12O6	
			beta-D-glucose	C00221		22	M+Na	203.0526	180.0634	C6H12O6	
			D-Mannose	C00159		22	M+Na	203.0526	180.0634	C6H12O6	
			Alpha-D-Galactose	C00984	M00554 M00632	22	M+Na	203.0526	180.0634	C6H12O6	
			D-Galactose	C00124	M00632	22	M+Na	203.0526	180.0634	C6H12O6	
			D-Glucose	C00031		22	M+Na	203.0526	180.0634	C6H12O6	
4	86.9541	0.000521									
5	181.0518	0.000517	5-Methylthioribose	C03089		6	M+H	181.0529	180.0456	C6H12O4S	
			Trans-2,3-Dihydroxycinnamate	C12623	M00545	12	M+H	181.0495	180.0423	C9H8O4	Trans-cinnamate degradation
			4-Hydroxyphenylpyruvic acid	C01179	M00025	12	M+H	181.0495	180.0423	C9H8O4	tyrosine biosynthesis
6	275.5738	0.000507									
7	156.9928	0.000475	Succinic acid	C00042	M00009 M00011 M00027 M00956	19	M+K	156.9898	118.0266	C4H6O4	Citrate cycle, GABA shunt lysine degradation
8	405.0225	0.00047	Phosphoribosyl formamidocarboxamide	C04734	M00048	4	M+K	405.02	366.0577	C10H15N4O9P	De novo purine biosynthesis
			Uridine 5'-diphosphate	C00015	M00938 M00053 M00052	32	M+H	405.0095	404.0022	C9H14N2O12P2	Pyrimidine deoxyribonucleotide and ribonucleotide biosynthesis, deoxyribonucleotide synthesis
9	152.9066	0.000467	Thiosulfate	C00320		7	M+K	152.9077	113.9445	H2O3S2	
10	356.8463	0.000434									

11	428.0495	0.000432	dGDP	C00361	M00053	30	M+H	428.0367	427.0294	C10H15N5O10P2	Deoxyribonucleotide biosynthesis
			ADP	C00008	M00053	30	M+H	428.0367	427.0294	C10H15N5O10P2	Deoxyribonucleotide biosynthesis
12	348.8654	0.000415									
13	432.0338	0.00041									
14	159.078	0.000394									
15	431.0995	0.000392									
16	493.8071	0.00039									
17	250.8868	0.000382									
18	425.1196	0.000378									
19	314.0225	0.000374	2-Methyl-4-amino-5-hydroxymethylpyrimidine diphosphate	C047	M00127	10	M+NH4	314.0192	295.9854	C6H8N3O7P2	Thiamine
20	457.1272	0.000373	FMNH(2)	C01847		33	M+H	457.1119	456.1046	C17H21N4O9P	Riboflavin biosynthesis
			Flavin Mononucleotide	C00061	M00125	33	M+H	457.1119	456.1046	C17H21N4O9P	Riboflavin biosynthesis

MG1655



MG1655

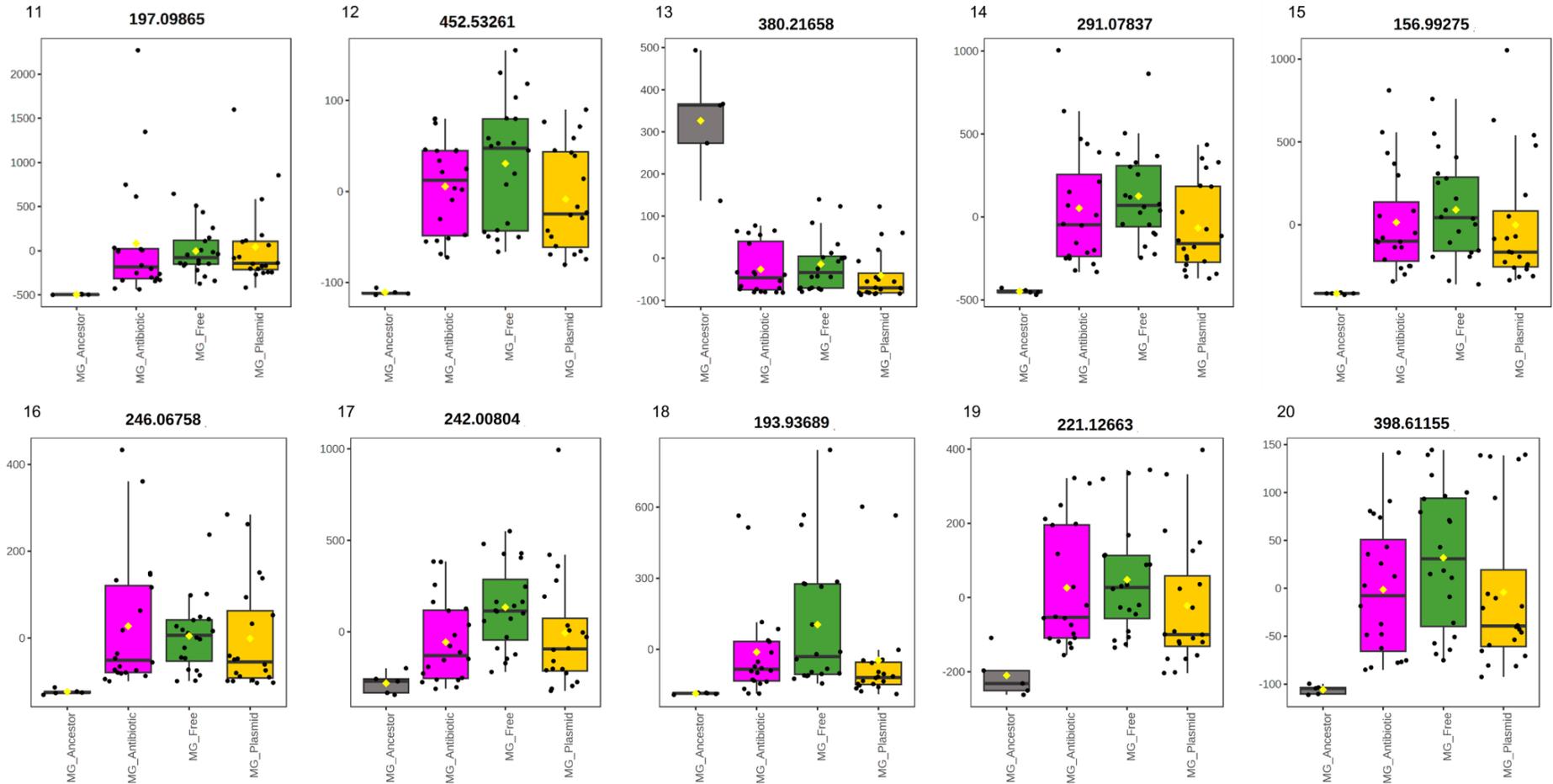


Figure 4.4. Random Forest Analysis showing the top 20 mass values under 500Da which cause the most variation between the ancestral state of *E. coli* strain MG1655 and the post-evolution clones that underwent 3 treatments of evolution; plasmid free, plasmid carrying and plasmid carrying with exposure to an antibiotic, cefotaxime.

Table 4.4. Random forest analysis rest of top 20 highest ranking masses under 500Da discriminating the ancestral and evolved groups of *E.coli* strain MG1655. Masses have been putatively identified using ECMDB and KEGG.

MG1655	Mass	Mean Decrease Accuracy	ID	KEGG	MODULE	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	Function
1	411.5344	0.001098									
2	184.9373	0.000838									
3	364.8681	0.000778									
4	292.8574	0.00073									
5	286.8633	0.000628									
6	390.8541	0.000585									
7	362.9447	0.000578									
8	480.856	0.000576									
9	382.8793	0.000561									
10	313.0444	0.000549									
11	197.0987	0.000548	L-Arginine	C00062	M00844	11	M+Na	197.1009	174.1117	C6H14N4O2	Arginine biosynthesis
12	452.5326	0.000546									
13	380.2166	0.000507									
14	291.0784	0.000484	Inosine	C00294	M00958	29	M+Na	291.07	268.0808	C10H12N4O5	Adenine ribonucleotide degradation
			2(alpha-D-Mannosyl)-D-glycerate	C11544		33	M+Na	291.0687	268.0794	C9H16O9	
15	156.9928	0.000474	Succinic acid	C00042	M00009 M00011 M00027 M00956	19	M+K	156.9898	118.0266	C4H6O4	Citrate cycle, GABA shunt lysine degradation
16	246.0676	0.000467									
17	242.008	0.000462									
18	193.9369	0.000461									
19	221.1266	0.000461									
20	398.6116	0.000459									

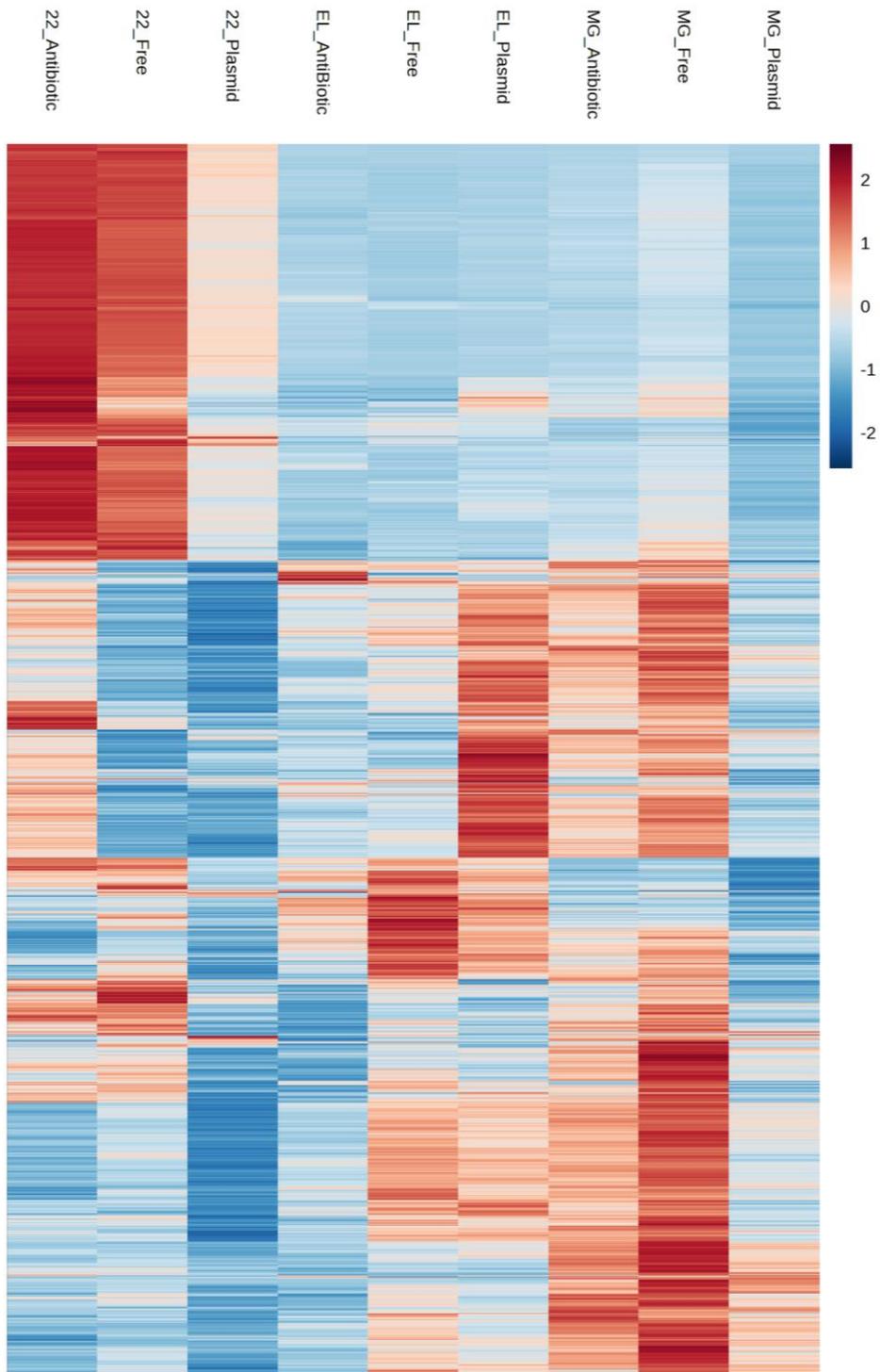


Figure 4.5 Heatmap showing the full metabolic profiles of 3 strains of *Escherichia coli*, scaled by intensity. Each strain is labelled by the conditions during evolution; plasmid free 'Free' and plasmid containing 'Plasmid' and plasmid containing under antibiotic stress 'Antibiotic'. The biological replicates for each strain have been pooled. Strain names 22 = F022, EL = ELU39 and MG = MG1655.

Metabolic responses vary by conditions they were evolved in.

To understand differences in the metabolome determined by evolutionary conditions, a series of pairwise analyses was next performed contrasting evolved lines between treatments per strain to understand differences in the metabolome determined by evolutionary conditions. Together, these analyses help separate functional changes in the metabolism as a result of plasmid carriage or of positive selection on the resistance genes carried by the plasmid.

The impact of plasmid co-evolution both in the presence and absence of antibiotic selection revealed distinct strain specific patterns in the scale and functions affected (Fig4.5). When evolved plasmid carrying treatments with or without cefotaxime were compared to the evolved plasmid-free control treatment, the cefotaxime treatment group had fewer significantly differently expressed metabolites than did the plasmid carrying treatment without antibiotic exposure (Fig 4.6 D-I). The clinical strain F022 had the largest response of the strains to coevolution with the pLL35 plasmid (2.6% of the metabolome) (Fig 4.6 plot A).

Comparing the plasmid treatment groups reveals the impact of selection by cefotaxime on plasmid carrying bacteria. The scale of impact varied and saw both the highest (3.64% in F022) and lowest (0.16% in MG1655) impacts on the metabolome. This may be indicative of alterations to the metabolic network beyond cost of plasmid carriage or stress of antibiotic exposure. Metabolic burden of plasmid carriage is measured and defined by the adverse effects on bacterial growth (Silva, Queiroz and Domingues, 2012) yet after co-evolution, F022 and MG1655 show improved growth kinetics compared to the ancestor (Carrilero, Dunn and Moran, 2023), suggesting that the metabolic modifications may actually indicate a net positive impact of plasmid carriage in these strains.

Masses reported in the fold change analyses between evolution treatments are mostly different from those reported by random forest analysis. The singular exception to this is the mass 96.07998, present in both the F022 random forest results (Fig 4.2 plots 1, 15) and the F022 plasmid carriage fold change results (Table S4.1), however this value was not identified.

Cyclic AMP is the singular commonly affected metabolite between strains.

Within strains, there were commonalities in the metabolic pathways impacted across treatments occasionally with directional differences. There was no overlap of pathways altered by evolution treatment between strains, except for one notable exception; the global regulator cyclic adenosine monophosphate (cAMP) which appeared in all 3 strains.

The universal secondary messenger cAMP is highlighted in all 3 treatments, and when looking solely at the effect of antibiotic selection, by comparing the treatments 'plasmid containing' and 'plasmid containing under antibiotic stress', is present and upregulated in all 3 strains simultaneously. This compound was downregulated (FC 0.43) in F022 when the plasmid co-evolved metabolome was compared with the plasmid free control. There was an alternative identification (Tables 4.5, 4.6, 4.7) which is indistinguishable by ppm but was nonetheless identified as a derivative of cAMP. cAMP was upregulated (FC 4.77) in ELU39 when the cefotaxime treatment was compared to the plasmid free control. Selection induced a common response in all 3 strains of upregulation of cAMP by 2.22-3.40 fold. Together this suggests the plasmid-free cAMP level is intermediate, and co-evolution with a plasmid induces downregulation of this, while selection with cefotaxime induces upregulation.

Strain specific metabolic responses to co-evolution with a plasmid.

The difference in bacterial metabolome caused by long term plasmid carriage can be seen by comparing the evolved datasets for plasmid free strains and plasmid carrying strains. Each strain had a unique set of metabolites affected by plasmid selection.

The clinical strain F022 has the most differentially expressed metabolites when coevolved with a plasmid compared to plasmid free controls (2.6% of the metabolome) and primarily displays downregulation (Fig 4.6 plot A) which could imply increased consumption of these metabolites as a result of plasmid carriage. Downregulation of amino acid metabolism represents the majority of identified plasmid carriage impact compared to plasmid free lines in strains F022 (Table 4.5). Amino acids have been shown to be the primary source of carbon needed for biomass synthesis (Maser *et al.*, 2020). Pathways implicated include biosynthesis of; serine, threonine, valine-isoleucine, leucine and lysine. Of the amino acids, serine is highly utilised for the TCA cycle and acetate biosynthesis (Maser *et al.*, 2020). Leucine is one of the most variable

amino acids, alongside arginine, because they have control functions for the biosynthesis of almost all other amino acids between them (Radoš *et al.*, 2022).

Coevolution with a plasmid induced downregulation of a metabolite identifiable as betaine or L-valine (Table 4.5). Betaine or valine appears downregulated 3 times in F022 in different ionised states, meaning the differential expression of this pathway is likely stronger than apparent from the individual metabolites alone. In every instance, ppm determined that betaine is the more likely identification. Betaine is part of the serine and threonine biosynthesis pathways (*KEGG*) and is an osmoprotectant whose presence increases cell water content (Cayley, Lewis and Record, 1992; Metris *et al.*, 2014). A downregulation of Betaine would result in a reduction of intracellular water (Cayley, Lewis and Record, 1992).

F022 also displays downregulation of the shikimate pathway, which produces intermediaries for the biosynthesis of the aromatic amino acids phenylalanine, tyrosine and tryptophan (Berg *et al.*, 2019).

Ubiquinol, formed by reduction of ubiquinone as part of the electron transport chain, was downregulated (FC 0.38). This implies that more ATP is being generated aerobically in strain F022 evolved with the plasmid compared to the plasmid free controls.

By contrast the environmental strain ELU39 had 15 upregulated metabolites (0.85% of the metabolome) in evolved lines with the plasmid versus the plasmid free treatment. Few metabolites of these could be identified (Table S4.1). The metabolite gamma-Glutamyl-gamma-butyraldehyde was identified, which is part of the Gamma-amino Butyric acid (GABA) pathway (FC 4.7)(Table 4.5). GABA is a non-protein amino acid involved in the tricarboxylic acid cycle (TCA) and aids in bacterial resistance to acidic conditions (Sarasa *et al.*, 2020).

MG1655 had 7 altered metabolites in total (0.35% of the metabolome) and favoured downregulation. Just 1 of these could be identified, and this was involved in non-central carbohydrate metabolism, specifically the UDP-N-acetyl-D-glucosamine biosynthesis pathway and was downregulated in plasmid carrying lines relative to plasmid free (Table 4.5). This pathway is involved in cell wall peptidoglycan synthesis,

liposaccharide biosynthesis and biosynthesis and addition of sugars to proteins (KEGG).

How does antibiotic selection change the metabolism of MDR plasmid carrying bacteria?

Comparing the plasmid free control treatment with the plasmid carrier cefotaxime treatment shows the effect of coevolution and antibiotic selection. Of the three analyses, this pairing has the smallest count of significantly altered metabolites. In response to antibiotic selection, MG1655, significant fold change differences in the metabolome accounted for downregulation of 4 masses, or 0.2% of the metabolome. ELU39 had 8 upregulated and 3 downregulated metabolites (0.63% of the metabolome) and F022 had 3 upregulated and 5 downregulated metabolites (0.37% of the metabolome) (Fig 4.6 plots D-F).

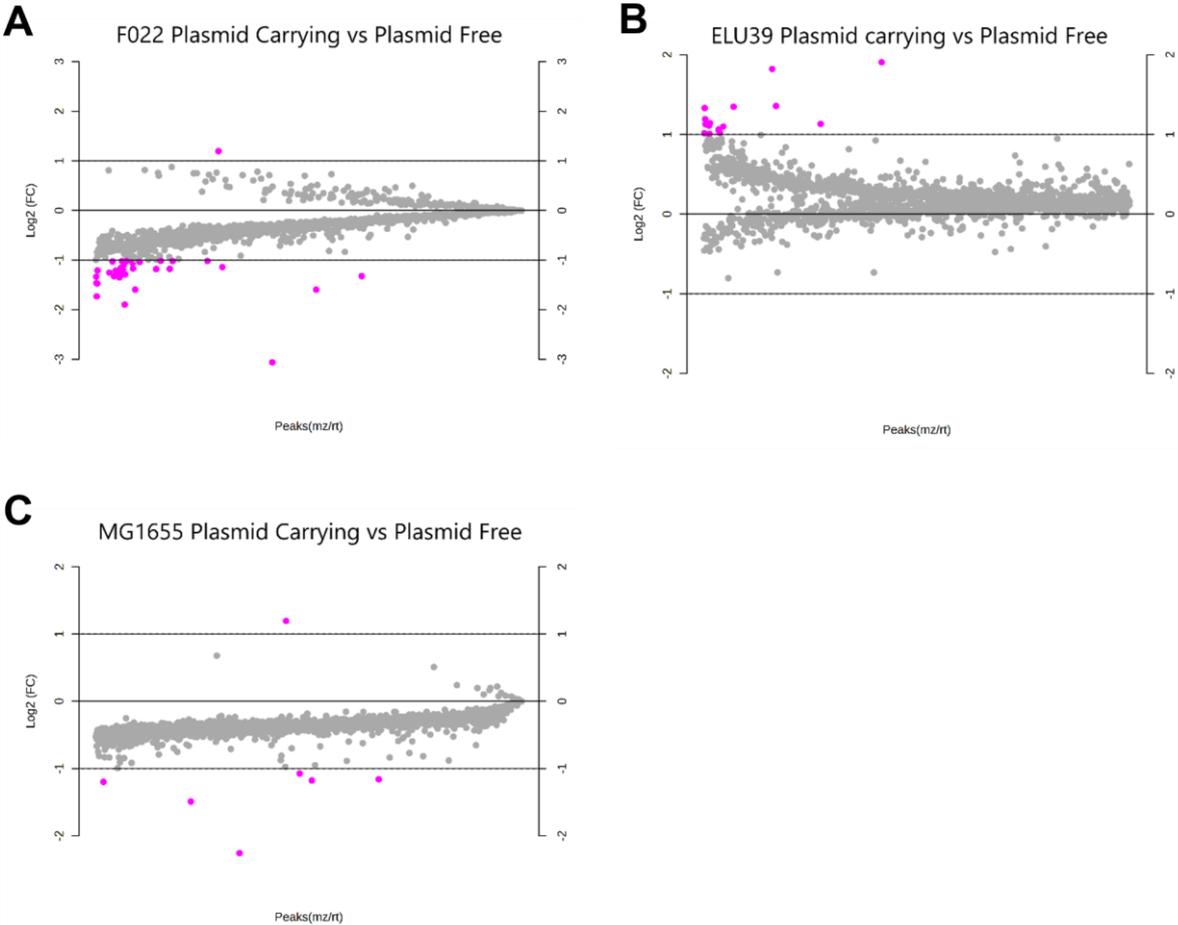
Many of the metabolites identified in the plasmid free versus plasmid carrying comparison are also highlighted in the plasmid free versus antibiotic plasmid carrying treatment comparison, indicating these differences in the bacterial metabolome are likely driven by carriage of the plasmid (Tables 4.5, 4.6).

Unique features in the plasmid-free versus antibiotic plasmid carrying comparison imply a combined effect of plasmid carriage and antibiotic selection. All identified features are downregulated and include amino acid biosynthesis; phenylalanine in ELU39, tyrosine in MG1655. A component of folate biosynthesis was also identified in MG1655. Phosphoribosylamine, part of de-novo purine biosynthesis, was downregulated in F022.

Direct comparison of plasmid carriers with versus without antibiotics show the effect of antibiotic selection on the bacterial metabolome. Antibiotic exposure altered 3.64% of the metabolome in a predominantly upregulated direction in the clinical strain F022 (78 metabolites upregulated and 3 downregulated), affecting amino acid and energy metabolism (Fig 4.6 plots G-I and Table 4.7). Selection caused upregulation of all metabolites related to amino acid metabolism, specifically the biosynthesis pathways of lysine, arginine, leucine and valine-isoleucine.

Glycerol, which feeds into glycolysis, was upregulated (FC 2.05) and propylene glycol, a by-product of glycolysis, was downregulated. This could imply a reduction in glycolysis but none of the identified metabolites were part of glycolysis directly.

In ELU39 1.3% of the metabolome is altered by selection, (6 metabolites upregulated and 11 downregulated) and 0.16% in MG1655 (2 metabolites upregulated and 1 downregulated). None of these metabolites were able to be identified, excepting cAMP.



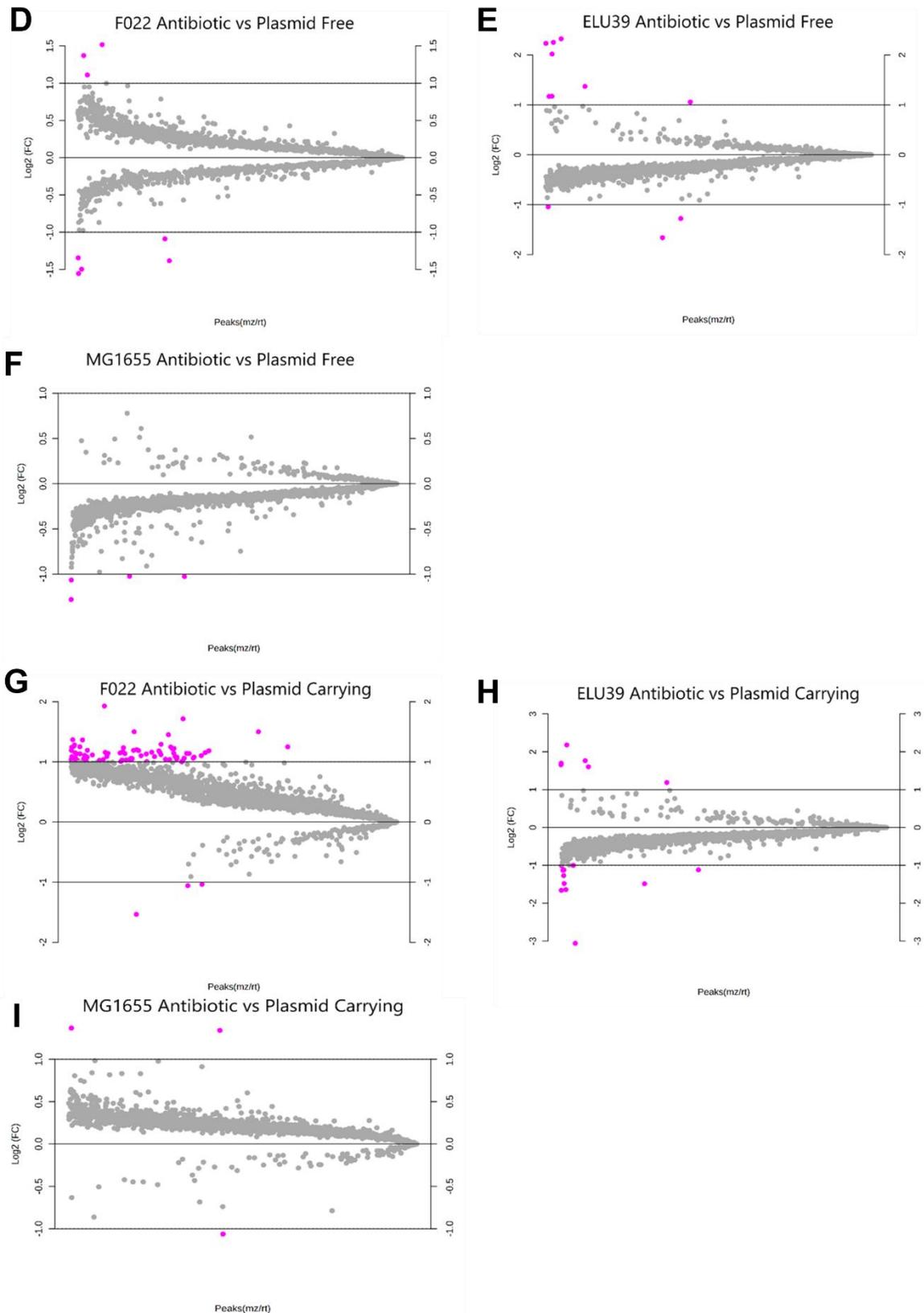


Figure 4.6 Fold change analysis highlighting in pink the masses that cross the fold change threshold of ± 2.0 . Y axis is the Log₂ of the fold change and the X axis are peak masses (mz/rt) corrected for retention time. A-C Evolved plasmid carrying *E. coli* metabolic profiles compared to evolved plasmid free *E. coli* metabolic profiles. D-F *E. coli* metabolic profiles evolved in antibiotic stress compared to the evolved plasmid free *E. coli* metabolic profiles. G-I *E. coli* metabolic profiles evolved in antibiotic stress compared to the evolved plasmid carrying *E. coli* metabolic profiles.

Table 4.5 Fold change summary of the masses (mz/rt) that cross the fold change threshold of +/- 2.0 comparing the plasmid carrying profiles against the plasmid free profiles by strain. Putative identifications made using KEGG and ECMDB. Full table, including masses not identified in the supplementary (S4.1).

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	log2(FC)	Function
F022	118.09047	Betaine	C00719	31	M+H	118.0868	118.0868	C5H12NO2	0.36273	-1.463	Serine and threonine biosynthesis
		L-Valine	C00183	36	M+H	118.0863	117.079	C5H12NO2			Branched chain amino acid metabolism valine-isoleucine metabolism
F022	140.07101	Betaine	C00719	16	M+Na		118.0868	C5H12NO2	0.36083	-1.4706	Serine and threonine biosynthesis
		L-Valine	C00183	20	M+Na		117.079	C5H12NO2			Branched chain amino acid metabolism
F022	153.04427	Xanthine	C00385	23	M+H				0.46548	-1.1032	Adenine and guanine metabolism
		2-Keto-3-methyl-valerate	C00671	48	M+Na	153.0517	129.0552	C6H9O3			Branched chain amino acid metabolism valine-isoleucine metabolism
F022	153.07936	Ribitol	C00474	24	M+H	153.0757	152.0685	C5H12O5	0.47956	-1.0602	
		L-Arabitol	C00532	24	M+H	153.0757	152.0685	C5H12O5			
F022	156.04794	Betaine	C00719	34	M+K	156.0427	118.0868	C5H12NO2	0.30117	-1.7313	Serine and threonine biosynthesis
		L-Valine	C00183	37	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism
F022	213.02325	3-Carboxy-3-hydroxy-isocaproate	C02504	34	M+K	213.016	174.0528	C7H10O5	0.48593	-1.0412	Branched chain amino acid metabolism leucine biosynthesis
		2-Isopropyl-3-oxosuccinate	C04236	34	M+K	213.016	174.0528	C7H10O5			Branched chain amino acid metabolism leucine biosynthesis
		Shikimic acid	C00493	34	M+K	213.016	174.0528	C7H10O5			Shikimate pathway
ELU39	217.11175	gamma-Glutamyl-gamma-butyraldehyde	C15700	30	M+H	216.111	217.1183	C9H16N2O4	2.544	1.3471	GABA biosynthesis
ELU39	226.9613	2,5-Dichloro-4-oxohex-2-enedioate	C12835	46	M+H	226.9509	225.9436	C6H4Cl2O5	2.2061	1.1415	
F022	270.15899	Ubiquinol-1	C00390	41	M+NH4	270.17	252.1362	C14H20O4	0.33108	-1.5948	Ubiquinone biosynthesis
MG1655	277.08735	Nicotinamide riboside	C03150	26	M+Na	277.08	255.0981	C11H15N2O5	0.44779	-1.1591	
		D-Galactosamine 6-phosphate	C06377	28	M+NH4	277.0795	259.0457	C6H14NO8P			
		alpha-D-Glucosamine 1-phosphate	C06156	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis

		Glucosamine-1P	C04501	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
F022	329.07394	N-Succinyl-L,L-2,6-diaminopimelate	C04421	2	M+K	329.0746	290.1114	C11H18N2O7	0.43099	-1.2143	Lysine metabolism
		Argininosuccinic acid	C03406	36	M+K	329.0858	290.1226	C10H18N4O6			Arginine biosynthesis
		2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	C16519	39	M+H	329.0867	328.0794	C14H16O9			Menaquinone biosynthesis
F022	347.08038	Adenosine 2',3'-cyclic phosphate	C02353	17	M+NH4	347.0863	329.0525	C10H12N5O6P	0.42755	-1.2258	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	17	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation

Table 4.6 Fold change summary of the masses (mz/rt) that cross the fold change threshold of +/- 2.0 comparing the antibiotic selection profiles against the plasmid free profiles by strain. Putative identifications made using KEGG and ECMDDB. Full table, including masses not identified in the supplementary (S4.2).

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	Log2(FC)	Function
F022	118.09047	Betaine	C00719	31	M+H	118.0868	118.0868	C5H12NO2	0.39376	-1.3446	Serine and threonine biosynthesis
		L-Valine	C00183	36	M+H	118.0863	117.079	C5H11NO2			Branched cahin amino acid metabolism
F022	156.04794	Betaine	C00719	34	M+K	156.0427	118.0868	C5H12NO2	0.34037	-1.5548	Serine and threonine biosynthesis
		L-Valine	C00183	37	M+K	156.0421	117.079	C5H11NO2			Branched cahin amino acid metabolism
ELU39	166.0853	D-Phenylalanine	C02265	6	M+H	166.0863	165.079	C9H11NO2	0.48577	-1.0417	Phenylalanine biosynthesis
		L-Phenylalanine	C00079	6	M+H	166.0863	165.079	C9H11NO2			Phenylalanine biosynthesis
MG1655	174.94267								0.4903	-1.0283	
MG1655	198.04356	7-Cyano-7-carbaguanine	C15996	25	M+Na	198.0386	175.0494	C7H5N5O	0.41098	-1.2829	Folate biosynthesis
ELU39	217.10864	gamma-Glutamyl-gamma-butyraldehyde	C15700	44	M+H	217.1183	216.111	C9H16N2O4	4.7016	2.2332	GABA biosynthesis
MG1655	220.0286	L-Tyrosine	C00082	38	M+K	220.0371	181.0739	C9H11NO3	0.47746	-1.0666	Tyrosine biosynthesis, thiamine biosynthesis
F022	247.06603	-Phosphoribosylamine	C03090	12	M+NH4	247.069	229.0351	C5H12NO7P	0.35436	-1.4967	Denovo purine biosynthesis
F022	270.15827	Ubiquinol-1	C00390	43	M+NH4	270.17	252.1362	C14H20O4	0.38333	-1.3834	Ubiquinone biosynthesis
ELU39	347.08318	Adenosine 2',3'-cyclic phosphate	C02353	9	M+NH4	347.0863	329.0525	C10H12N5O6P	4.7682	2.2534	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	9	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation

Table 4.7. Fold change summary of the masses (mz/rt) that cross the fold change threshold of +/- 2.0 comparing the antibiotic selection profiles against the plasmid carrying profiles by strain. Putative identifications made using KEGG and ECMDDB. Full table, including masses not identified in the supplementary (S4.3).

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	Log2(FC)	Function
F022	99.04636	Propylene glycol	C00583	48	M+Na	99.0416	76.0524	C3H8O2	0.3453	-1.5341	by-product of glycolysis
F022	115.0412	Glycerol	C00116	40	M+Na	115.0366	92.0473	C3H8O3	2.0506	1.036	feeds into glycolysis
F022	135.069	R)-2,3-Dihydroxy-isovalerate	C04272	28	M+H	135.0652	134.0579	C5H10O4	2.1982	1.1363	Valine/isoleucine branched chain amino acid metabolism
		Deoxyribose	C01801	28	M+H	135.0652	134.0579	C5H10O4			
F022	153.0794	Ribitol	C00474	24	M+H	153.0757	152.0685	C5H12O5	2.0533	1.038	
		L-Arabitol	C00532	24	M+H	153.0757	152.0685	C5H12O5			
MG1655	174.9427								0.47841	-1.0637	
F022	195.0083	Glycerol 3-phosphate	C00093	28	M+Na	195.0029	172.0137	C3H9O6P	2.1478	1.1028	
F022	213.0233	3-Carboxy-3-hydroxy-isocaproate	C02504	34	M+K	213.016	174.0528	C7H10O5	2.2291	1.1565	branched chain amino acid metabolism leucine biosynthesis
		2-Isopropyl-3-oxosuccinate	C04236	34	M+K	213.016	174.0528	C7H10O5			branched chain amino acid metabolism leucine biosynthesis
		Shikimic acid	C00493	34	M+K	213.016	174.0528	C7H10O5			shikimate pathway
F022	329.0711	N-Succinyl-L,L-2,6-diaminopimelate	C04421	11	M+K	329.0746	290.1114	C11H18N2O7	2.2688	1.1819	Lysine metabolism
		Argininosuccinic acid	C03406	45	M+K	329.0858	290.1226	C10H18N4O6			arginine biosynthesis
		2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	C16519	47	M+H	329.0867	328.0794	C14H16O9			menaquinone biosynthesis
F022	347.0599								2.8285	1.5001	
ELU39	347.0818	Adenosine 2',3'-cyclic phosphate	C02353	13	M+NH4	347.0863	329.0525	C10H12N5O6P	3.4007	1.7658	precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	13	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
MG1655	347.0851	Adenosine 2',3'-cyclic phosphate	C02353	4	M+NH4	347.0863	329.0525	C10H12N5O6P	2.5321	1.3403	precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	4	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F022	347.0879	Adenosine 2',3'-cyclic phosphate	C02353	4	M+NH4	347.0863	329.0525	C10H12N5O6P	2.2155	1.1476	precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	4	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F022	347.5848								2.3	1.2016	

F022	358.0724	S-Formylglutathione	C01031	12	M+Na	358.0679	335.0787	C11H17N3O7 S	2.0873	1.0617	Methane metabolism
------	----------	---------------------	--------	----	------	----------	----------	-----------------	--------	--------	--------------------

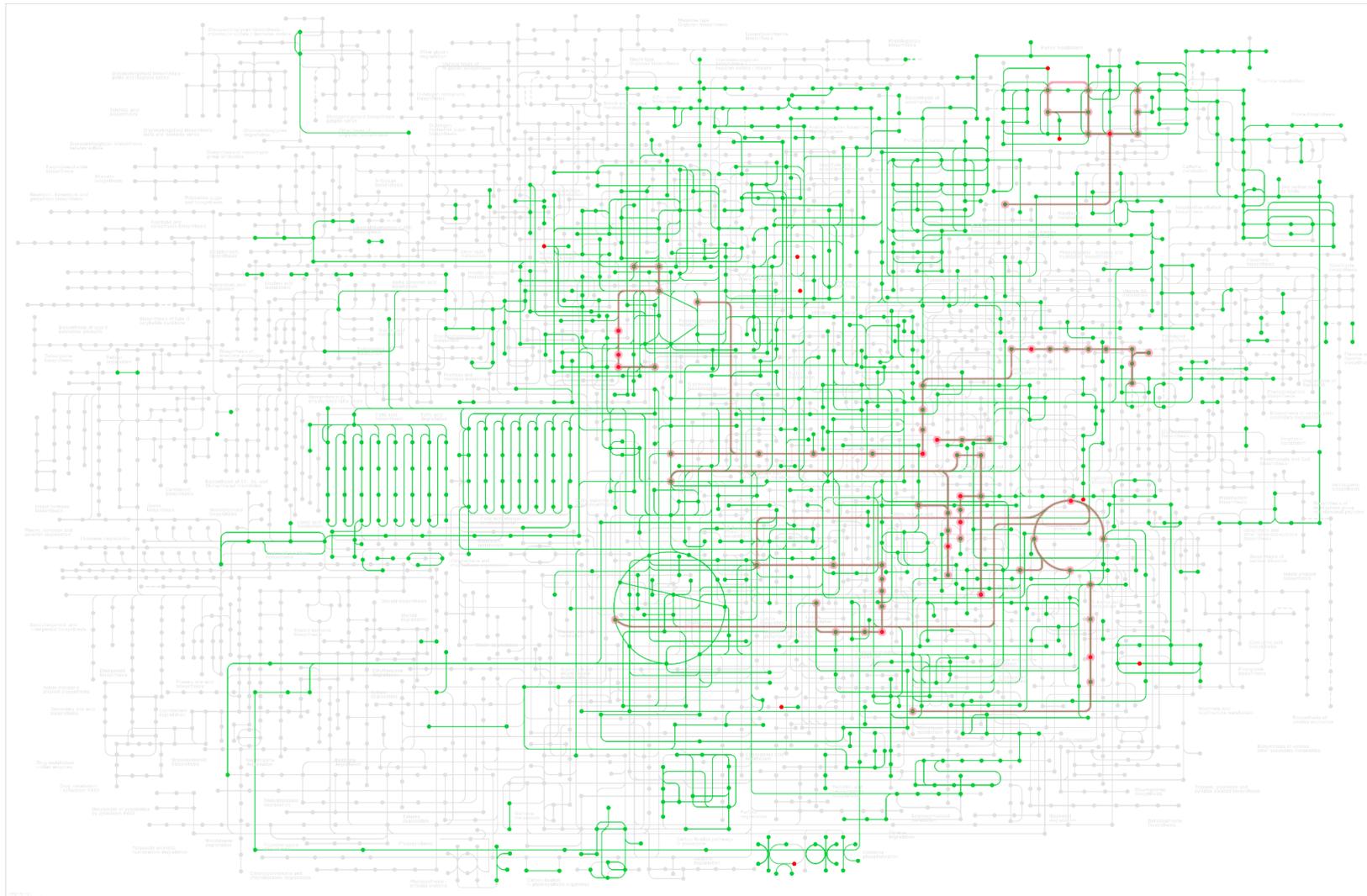


Figure 4.7. All possible putative identifications of metabolites and pathways affected by plasmid carriage compared to plasmid free lineages in all 3 strains, F022, ELU39 and MG1655, as described in Table 1 (pink). Visualised on the KEGG metabolic pathways map for *E. coli* K-12 MG1655 (green).

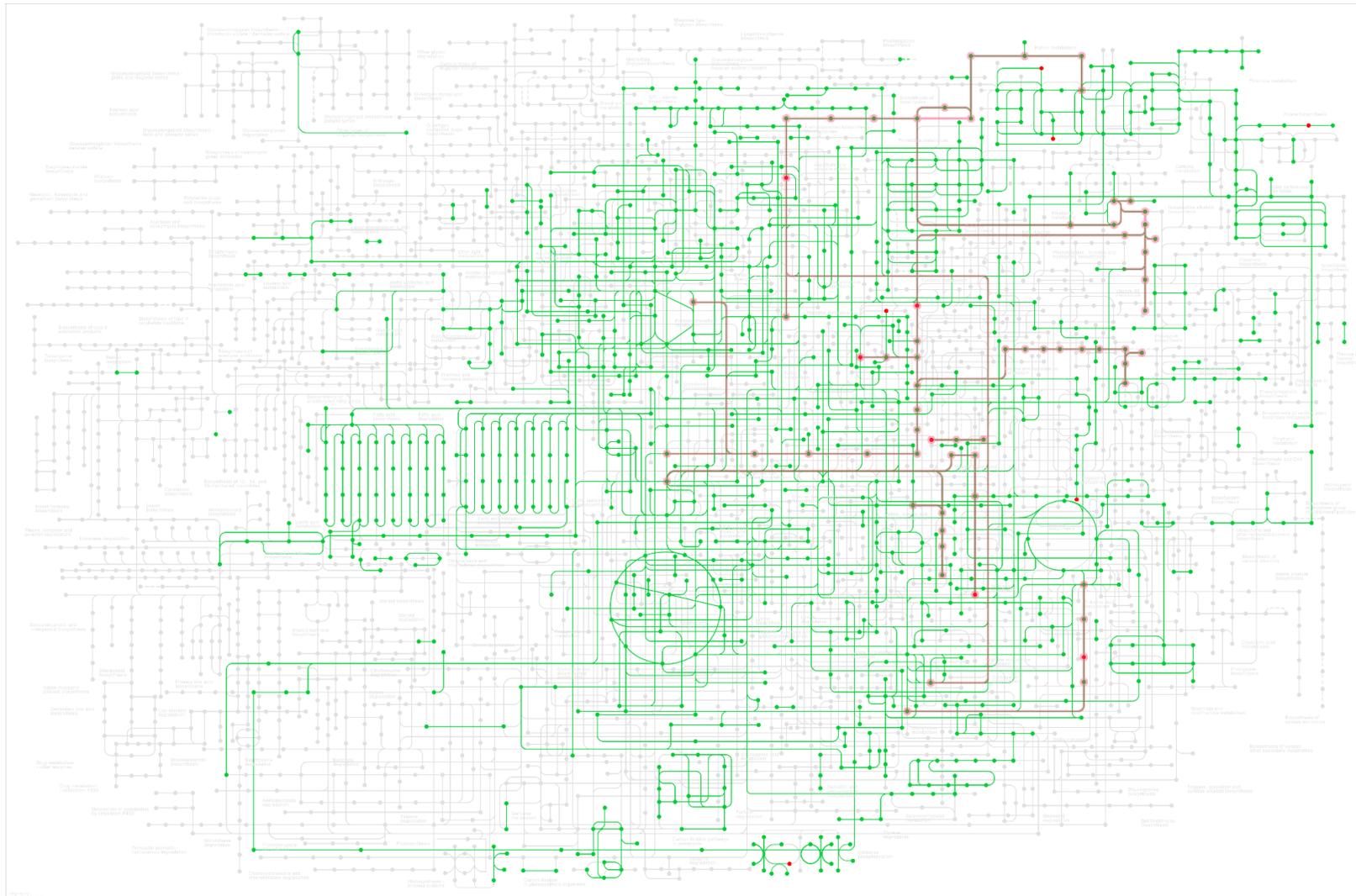


Figure 4.8. All possible putative identifications of metabolites and pathways affected by antibiotic selection of a plasmid compared to plasmid free lineages in all 3 strains, F022, ELU39 and MG1655, as described in Table 2 (pink). Visualised on the KEGG metabolic pathways map for *E.coli* K-12 MG1655 (green).

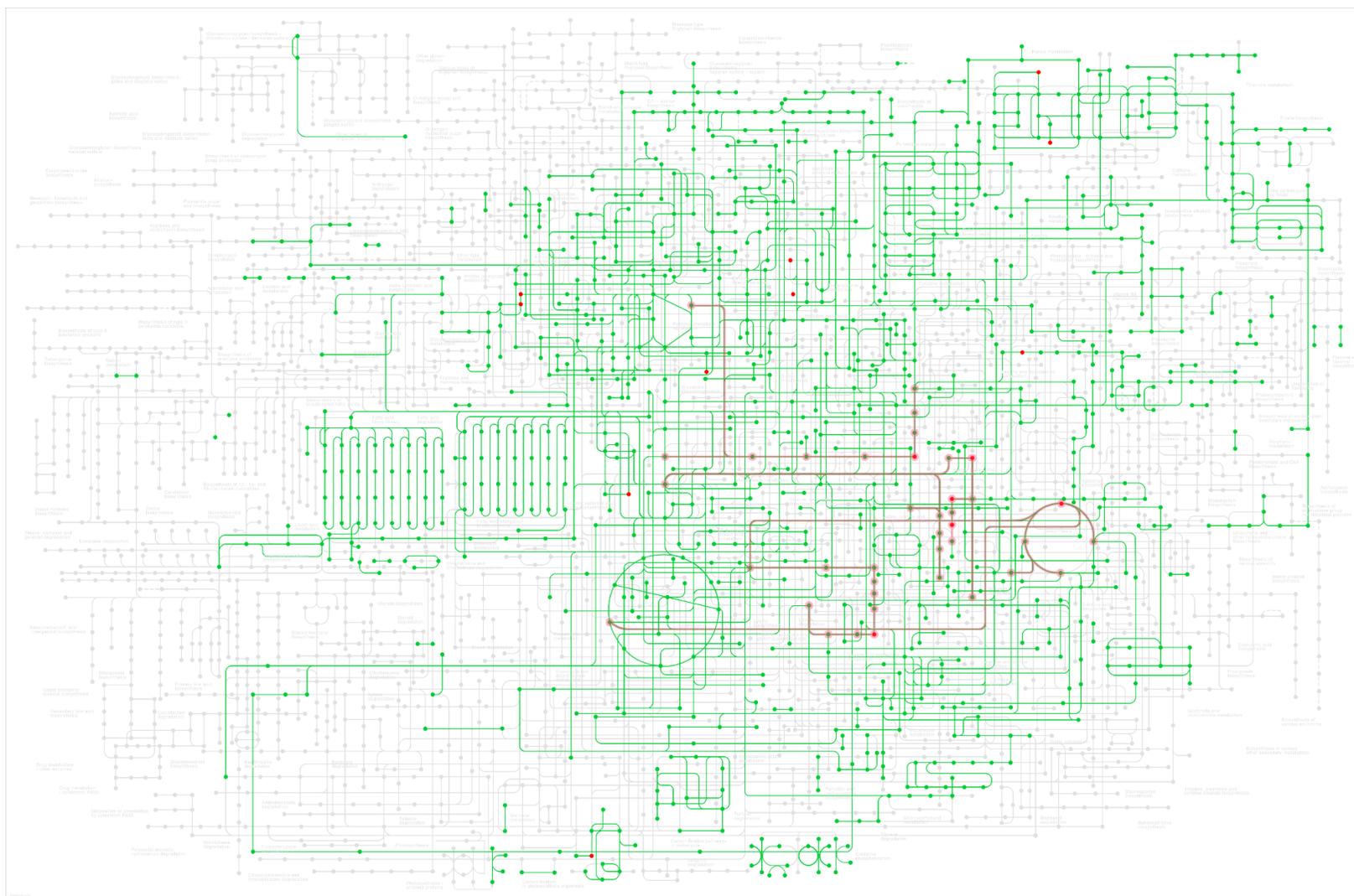


Figure 4.9. Putative identifications of significant metabolites and pathways comparing plasmid carriers with and without selection in strains F022, ELU39 and MG1655, as described in table 3 (pink). Visualised on KEGG metabolic pathways map for *E. coli* K-12 MG1655 (green).

4.5 Discussion

Metabolomics were used to understand the metabolic responses of 3 diverse *E.coli* strains to an MDR plasmid and the antibiotic cefotaxime following 700 generations of evolution. The *E.coli* strains are from a clinical (F022), environmental (ELU39) or laboratory (MG1655) background, and were each evolved in a plasmid free state, plasmid carrying, and plasmid carrying with antibiotic selection. Comparison of evolved lines to their ancestor revealed large-scale changes in metabolism that were consistent among treatments within strains, indicating pervasive metabolic changes associated with adaptation to the lab environment, including amino acid metabolism and glycolysis. Targeted pairwise contrasts were then used to determine the metabolic responses to the selection treatments per strain. Each strain had specific aspects of the metabolism altered by evolutionary conditions, (Fig 4.7, 4.8 and 4.9) and the only commonly changed metabolite was global regulator cAMP. Within strains, common altered pathways were often seen in plasmid carriers evolved both with and without antibiotic, but sometimes changing in opposing directions (Tables 4.5, 4.6, 4.7). The clinical strain F022 showed the largest metabolic impact as a result of coevolution with a plasmid, with and without antibiotic selection.

Functions changed by plasmid carriage or antibiotic selection

The metabolic responses to evolutionary conditions were strain specific. Although common broad functions were affected, such as amino acid metabolism, the specific pathways affected were unique to each strain. For example, the environmental strain ELU39 saw changes in phenylalanine metabolism, whereas the lab strain MG1655 showed altered levels of tyrosine and the clinical strain F022 showed altered levels of lysine and leucine metabolism (Tables 4.5, 4.6, 4.7).

Within strains a degree of functional consistency was observed between the evolution treatment groups but the analysis was able to distinguish metabolic effects derived from plasmid carriage and those derived from additional antibiotic selection. Many of the functions affected by antibiotic exposure were associated with metabolic stress, such as arginine, proline, serine, threonine and methane metabolism, glycolysis, and the TCA cycle (Zhao *et al.*, 2019).

The functions affected by plasmids acquisition (Chapter 3) are all present in the post evolution data. Plasmid acquisition caused a very low metabolic response in F022, lower than either ELU39 or MG1655 but F022 had the most metabolic changes after co-evolution with the plasmid. The clinical strain F022 showed the most change in the metabolome after ~700 generations of plasmid carriage and was the strain with the most differences caused by antibiotic exposure. This strain also demonstrated some of the highest levels of mutations and transcriptional alterations (Carrilero, Dunn and Moran, 2023).

Cyclic AMP is the singular commonly affected metabolite between strains.

In previous work, genomics showed that over a fifth of mutations acquired by evolved clones were common between the strains (Carrilero, Dunn and Moran, 2023), there was only one commonly affected metabolite in all 3 strains, cyclic AMP, which was altered in response to coadaptation with a plasmid under antibiotic exposure. Cyclic AMP responds to environmental stimuli and in pathogenic bacteria and has several important regulatory functions, including virulence, type II secretion, carbon metabolism and biofilm formation (McDonough and Rodriguez, 2012). cAMP is synthesised from ATP at low glucose concentration (Shimizu, 2013).

cAMP is downregulated in the plasmid carrying treatment without cefotaxime and upregulated in all 3 strains in the cefotaxime treatment. As a regulator, accumulation of cAMP may indicate an increase in downstream processes regulated by cAMP. These include biofilms which can reduce sensitivity to antibiotics (Crabbé *et al.*, 2019). Together, this suggests the accumulation of cAMP may induce biofilm formation as an additional defense response by the bacteria against cefotaxime, despite being evolved in liquid shaken conditions that would normally be expected to disfavour biofilms (Carrilero, Dunn and Moran, 2023). Additional experimental work would be required to definitively prove the link between the observed changes in cAMP levels and biofilm production.

Evidence of beta-lactam specific metabolic responses

Several altered metabolic functions were consistent with the cell stresses induced by beta-lactam antibiotics. By impeding cell wall synthesis and initiating a futile cycle of

peptidoglycan degradation, beta-lactams induce an energy demand and systemic cellular toxicity associated with ROS (Cho, Uehara and Bernhardt, 2014; Dwyer *et al.*, 2014; Adolfsen and Brynildsen, 2015). Despite containing Bla_{CTX-M-15} which can degrade the antibiotic, the altered metabolites in the evolutionary treatment with cefotaxime highlight alterations to energy and redox regulation through cAMP, glycolysis and alternative energy sources and the electron acceptor ubiquinol.

Additionally, cefotaxime selection led to a potential decrease in glycolysis because the by-product propylene glycol is downregulated but glycerol, which supplies glycolysis, is upregulated in F022. This accumulation of the supply and reduction in by-product suggests the central reaction is not as active. As none of the metabolites identified are directly part of glycolysis it remains difficult to determine the energy dynamics (Table 4.7). Furthermore, evolved MG1655 and F022 plasmid carriers had notably higher growth rates compared to their ancestors indicating compensatory evolution for plasmid fitness costs (Carrilero, Dunn and Moran, 2023). This may have been associated with improved energy efficiency, as this strain also contained a mutation in glycerol metabolism genes in plasmid carriers (Carrilero, Dunn and Moran, 2023) but metabolic efficiency and growth rate are not proportionally linked (Metris *et al.*, 2014). Alternatively, the increase in metabolites part of amino acid metabolism and alternative carbon sources may instead point to alternative ways of supplying the TCA cycle. As the most efficient method of energy production in bacterial cells, a demand on energy imposed by stressors would increase catabolism of metabolites to supply the TCA (Arense *et al.*, 2010; Rui *et al.*, 2010; Passalacqua, Charbonneau and O’Riordan, 2016).

Ubiquinol is downregulated by plasmid carriage and antibiotic selection. A reduced abundance of a metabolite can indicate its consumption (Laviña *et al.*, 2020), therefore in plasmid carrying F022 ubiquinol may be actively consumed as part of the electron transport chain in aerobic respiration (Aussel *et al.*, 2014; Agrawal *et al.*, 2017; Nitzschke and Bettenbrock, 2018). Ubiquinol also has a role in counteracting redox stress because as a reducing agent ubiquinol is reduced to form ubiquinone. Therefore a decrease in ubiquinol in the clinical strain F022 may reflect consumption of ubiquinol to facilitate ATP production and aid in combating cefotaxime-induced redox stress. The ubiquinone biosynthesis pathway was also implicated as being affected by plasmid

acquisition in the ancestral clones of F022 (Chapter 3), suggesting this quinone is important in both short term and long term responses to the cellular impacts of plasmid carriage.

Amino acid metabolism

Comparison of the mutations acquired by the evolved lines reported by Carrilero et al. with the metabolomics data reported here reveals some overlapping functions (Carrilero, Dunn and Moran, 2023). Plasmid carrying MG1655 acquired mutations in aromatic amino acid metabolism. Tyrosine is one of the aromatic amino acids and is downregulated by plasmid carriage with cefotaxime selection.

Upon acquisition of the plasmid, the lysine degradation pathway is upregulated in MG1655. Lysine metabolism is upregulated in F022 when the plasmid carrying treatment groups with and without cefotaxime selection were compared. This same metabolite was identified in the plasmid free and plasmid carrying comparison but was down regulated. This implies that plasmid carriage reduced lysine metabolism in F022, but it was driven up again by positive selection to a level not significantly different from the plasmid free group. However, this metabolite has an alternative identification involved in arginine biosynthesis, and the arginine transporter gene *art P* was mutated in plasmid carrying F022 and ELU39 (Carrilero, Dunn and Moran, 2023).

The mutated functions do not, however, match perfectly with changes in metabolites. For example, ELU39, F022 and MG1655 evolved plasmid carriers all gained mutations in operons involved in anaerobic respiration and fatty acid metabolism (Carrilero, Dunn and Moran, 2023) but these functions were not altered in the metabolome. This is likely to reflect the fact that these experiments were performed in aerobic conditions and thus these pathways may not have been expressed. Some caution is required however, because a large proportion of the significant masses were unable to be identified, so future improvements in identification systems may uncover these functions.

In certain instances, pathways highlighted in this evolutionary analysis were differentially expressed upon plasmid acquisition in other strains analysed in Chapter

3. For example, formylglutathione was upregulated following evolution under antibiotic selection in F022 here, and methylamine was upregulated following plasmid acquisition in F054 in Chapter 3. Both metabolites belong to the methane metabolism pathway, and both strains are of a clinical background.

Conclusions

The metabolome of *E. coli* at the endpoint of an evolution experiment demonstrated distinct strain specific profiles. While overarching functions were shared across strains, the particular pathways impacted were unique to individual strains. The singular exception to this was the global regulator cAMP which was altered in all strains as a consequence of evolution under antibiotic selection with plasmid pLL35. The large metabolic changes seen in the clinical strain F022 coupled with improved growth kinetics relative to the ancestor following evolution, demonstrates compensatory evolution for plasmid carriage and resisting antibiotics occurs at all biological levels. Furthermore, the evidence of metabolic responses to beta-lactam-specific stresses within the antibiotic selection treatment group highlights the importance of metabolic responses for cells to mitigate effects of exposure to antibiotics even when they carry the relevant resistance genes.

Further questions

Future studies should consider how the nature of plasmid carrying bacteria changes when confronted with antibiotics, with an emphasis on replicating or using common plasmid and bacterial combinations and their natural environmental conditions, because metabolic signatures of antibiotic stress are seen even in bacteria carrying the relevant resistance genes. Lab conditions are not representative of real-world scenarios, and with such specific responses this is important to replicate. For instance, there is a difference in metabolic response dependent on planktonic or biofilm states which has implications for infection conditions (Zhao *et al.*, 2019) More specifically and relevant here is the fact the gut is anaerobic and that related functions appear to be important but not seen here due to aerobic experiments; expanding the conditions to make them more host relevant is key priority. (Dunn *et al.*, 2021; Carrilero, Dunn and Moran, 2023).

Chapter 5 : **Metabolic responses of 4 *E.coli* strains carrying a multidrug-resistant plasmid to antibiotic exposure.**

5.1 Abstract

Multidrug-resistant (MDR) bacterial strains pose a significant threat to public health, leading to increased mortality rates and prolonged hospitalisation. Conjugative plasmids facilitate the rapid dissemination of resistance genes, contributing to the accumulation of multidrug resistances in bacterial pathogens. Understanding the metabolic consequences of antibiotic stress is valuable for identifying potential drug targets and treatment strategies. Untargeted metabolomics were performed to understand the metabolic impacts of antibiotic stress on 4 *Escherichia coli* strains from diverse ecological backgrounds. The strains were carrying the MDR plasmid pLL35, originating from *Klebsiella pneumoniae* which conferred resistance to 3 classes of antibiotics. The strains were stressed with a sub-MIC concentration of kanamycin, cefotaxime and ciprofloxacin. The metabolic responses to antibiotic displayed parallel functions across strains, but the extent and direction of change in particular pathways and metabolites were strain specific. Broadly, functions affected included alterations in amino acid and nucleotide metabolism, energy production pathways and cofactor biosynthesis. Notably, redox stress mitigation mechanisms were observed across all strains in the ciprofloxacin treatment, and cell wall and biofilm biosynthesis was indicated in all treatments, suggesting bacterial resistance mechanisms were active alongside plasmid-encoded resistance. Strain specificity has implications for future research on innovative combination therapies.

5.2 Introduction

Multidrug-resistant (MDR) bacterial strains pose a severe threat to public health, contributing to increased mortality rates and prolonged hospitalisation for both initial admission infections and iatrogenic infections (*Antimicrobial Resistance*, 2021). MDR plasmids enable the rapid spread of multiple resistance genes by inter and intraspecies conjugation (Carattoli, 2013; Hall, Brockhurst and Harrison, 2017; San Millan, 2018). While a given strain can accumulate plasmids conferring different resistances over time, multidrug resistant plasmids can confer resistance to multiple classes of antibiotic in a singular transfer event, posing a particular clinical threat to combination antibiotic therapies (Carattoli, 2013). For example, *E.coli*, which commonly causes a wide range of serious human and animal infections worldwide, frequently carry MDR plasmids and consequently are becoming increasingly challenging to treat due to the diminishing effectiveness of available antibiotics (Mathers, Peirano and Pitout, 2015; Stoesser *et al.*, 2016; Dunn, Connor and McNally, 2019). Multidrug resistant plasmids have been integral to the evolution of the predominant MDR *E.coli* lineage ST131, the globally disseminated strain that can produce extended-spectrum beta-lactamases (ESBLs) and are resistant to frontline antibiotics such as carbapenems and cephalosporins (Stoesser *et al.*, 2016; Dunn, Connor and McNally, 2019). Such MDR plasmids are often stably maintained by toxin-antitoxin systems and compensatory amelioration of costs mediated by both the bacterial host and the plasmids (Carattoli, 2013; Porse *et al.*, 2016).

The rate of novel antimicrobial development is not fast enough to keep up with the evolution and dissemination of resistance (Krell and Matilla, 2022). Understanding the relationship between bacteria and their MDR plasmids and thus devising interventions to control plasmid spread is therefore a vital part of tackling the global AMR crisis. Analysing metabolic profiles under antibiotic stress offers valuable insights for identifying potential metabolic targets in the development of treatment strategies. For example, revealing metabolites that are suppressed by the bacteria in response to antibiotics allows development of combination therapies. The proposed treatment would deliver metabolites that are normally suppressed, enhancing susceptibility to the antibiotics, alongside the antibiotic itself (Peng *et al.*, 2015). Redirecting metabolism to

prevent or diminish resistance phenotypes emerges as a promising future strategy to re-sensitise bacteria to antibiotics (Meylan, Andrews and Collins, 2018; Crabbé *et al.*, 2019). However, a significant challenge lies in the current lack of understanding regarding the metabolic workings of clinical strains that could serve as viable treatment targets (Kok *et al.*, 2022). This is especially relevant for cases where the resistance is plasmid encoded because plasmids induce a range of disruptions to bacterial metabolism phenotypes that could act as potential therapeutic targets (Coulson *et al.*, 2015; Takahashi *et al.*, 2015; San Millan *et al.*, 2018; Billane *et al.*, 2022).

While metabolic fingerprints reliably relate to antibiotic mechanisms of action (Hoerr *et al.*, 2016), the interaction between environment, and plasmid associated changes to the metabolome has not been investigated. There is a need for biological systems-wide studies as increasingly, bacterial responses to antibiotic stress are acknowledged to be a combination of resistance genes and cell-wide network dependent responses involving global transcriptional regulators (Deter, Hossain and Butzin, 2021). Bacteria employ responses on a network scale to antibiotic stress in addition to resistance provided by accessory genes on plasmids (Deter, Hossain and Butzin, 2021; Kok *et al.*, 2022). The transcriptional response to ampicillin in *E.coli* affected hundreds of genes - notably the most frequently affected functional groups were amino acid transport and metabolism, energy production and conversion and protein synthesis, modification, and degradation (Deter, Hossain and Butzin, 2021).

To further the understanding of intracellular dynamics of *E.coli* harbouring an MDR plasmid, this chapter uses untargeted metabolomics to obtain a comprehensive view of the metabolic changes in multiple *E.coli* strains carrying MDR plasmid pLL35 under antibiotic stress. The metabolome, as the molecular underpinning of phenotype, reveals early responses to antibiotic stress and the genetic and molecular consequences of adaptations necessary for sustaining resistance mechanisms (Kok *et al.*, 2022). Therefore, this is a vital tool to aid the discovery of emerging drug targets and treatment strategies associated with metabolism (Kok *et al.*, 2022). Previous work presented in this thesis has shown that the acquisition of the MDR plasmid pLL35 induced subtle, strain specific metabolic impacts. Acquisition affected energy metabolism, fatty acid metabolism, amino acid metabolism and oxidative phosphorylation (Chapter 3). Experimental evolution of the bacteria plasmid relationship in the presence of an antibiotic - where the plasmid is essential for bacterial

survival - resulted in amelioration of plasmid costs through genetic mutations and alterations in the transcriptome and metabolome (Carrilero, Dunn and Moran, 2023; and Chapter 4). This suggests that plasmid associated disruption to the bacteria continues in the presence of antibiotics, despite providing a net fitness benefit.

The interaction between antibiotic exposure and carriage of an MDR plasmid pLL35 on the bacterial metabolome has not been rigorously tested. Moreover, although data in Chapter 4 suggests metabolic impacts of cefotaxime exposure, pLL35 confers resistance to 2 additional classes of antibiotic; aminoglycoside and fluoroquinolone. The mechanisms of resistance to each class of antibiotic differ markedly. The bactericidal mechanism of aminoglycosides involves inhibition of protein synthesis by binding to parts of the ribosomal complex (Kotra, Haddad and Mobashery, 2000). Mechanisms of resistance include phosphorylation, nucleotidylation, efflux, altering of the target and acetylation (Dunn, Connor and McNally, 2019). The *aacA4*, *aadA1* and *aacC2d* genes on the plasmid (Table 5.1), modify the kanamycin molecule by acetylation or adenylation and reduce its affinity for the ribosomal complex (Kotra, Haddad and Mobashery, 2000). Beta-lactams disrupt the biosynthesis of cell walls (Padda and Nagalli, 2022) and the resistance mechanisms act in the form of hydrolysis, efflux and alteration of the target (Dunn, Connor and McNally, 2019). CTX-M-15 and TEM-112 code for beta-lactamase enzymes that hydrolyse beta-lactam antibiotics. CTX-M enzymes can lyse cephalosporins like cefotaxime and have disseminated worldwide, causing near untreatable infections (Poole, 2004; Cantón, González-Alba and Galán, 2012; Bevan, Jones and Hawkey, 2017). In particular *E.coli* ST131 carrying *bla*_{CTX-M} genes is considered a high risk clone (Peirano and Pitout, 2019). Fluoroquinolones disrupt DNA replication, repair and supercoiling by targeting DNA gyrase and topoisomerases. The resistance mechanisms to combat quinolones include efflux, alteration of the target and acetylation (Dunn, Connor and McNally, 2019). Ciprofloxacin targets DNA topoisomerase II (Ojkic *et al.*, 2020), and the *qnrS1* gene encoded by the plasmid produces the QnrS1 protein which interferes with the antibiotic by binding to DNA topoisomerase and therefore altering the target (Strahilevitz *et al.*, 2009).

In this study, the impact of different antibiotic exposures on 4 plasmid carrying strains was investigated. For my metabolomic experiments, Kanamycin, Cefotaxime and Ciprofloxacin were chosen as representatives of the classes of AMR gene carried by

pLL35. *E. coli* strains MG1655 (lab strain), ELU39 (environmentally derived strain), F054 and F022 (clinically derived strains) carrying pLL35 were grown aerobically in nutrient rich media, with sub-minimum inhibitory concentrations of the 3 antibiotics, alongside antibiotic free controls. The impact of antibiotic exposure on the metabolic profile of these strains collectively (see disclaimer) was analysed, showing that bacteria exhibit stress responses and resistance mechanisms outside of plasmid encoded resistance genes.

Table 5.1 Table summarising the resistance genes on the plasmid pLL35. See Fig 3.1 for full schematic.

Gene	Resistance
<i>aacA4</i>	Aminoglycosides
<i>aacC2d</i>	Aminoglycosides
<i>aadA1</i>	Aminoglycosides
<i>bla_{CTX-M-15}</i>	Beta-lactams
<i>bla_{TEM-112}</i>	Beta-lactams
<i>qnrS1</i>	Quinolones

5.3 Methods

E. coli strains MG1655, ELU39, F022 and F054 carrying plasmid pLL35 were grown in nutrient broth at 37°C, 180 rpm, for 24 hours. 1% of the population was then transferred and grown for 3 hours to mid-exponential phase, followed by exposure to sub-minimum inhibitory concentrations of the antibiotics. The cefotaxime MIC >1024 µg/mL, kanamycin 102.4 µg/mL and ciprofloxacin 102.4 µg/mL. After testing shock duration and checking survival rate, it was decided the concentrations of antibiotic would be: Kanamycin (5µg/ml), Cefotaxime (5µg/ml) or Ciprofloxacin (2µg/ml) for 1 hour, along with an antibiotic free control. The samples were then isolated and frozen at -80°C. Prepared for mass spectrometry with chloroform and methanol, following the method described in chapter 2.

The samples (50µl) were introduced to the Waters G2/G2Si Synapt mass spectrometer. The instrument settings are detailed in the supplementary material (S5) HPLC-QToF MS was performed using the aqueous phase of the samples for DESI MS in positive mode, with a scanning range of 50-1200m/z over 3 minutes.

Data Processing and Analysis

All data processing and analysis followed methodology developed into a user-friendly guide and based on open source software (Parker *et al.*, 2023) (<https://untargeted-metabolomics-workflow.netlify.app/> accessed on 27 January 2023).

All raw data files were converted to mzML format using the [Proteowizard](#) software MsConvert. [XCMS](#) online was used for peak alignment and retention time correction (parameters 84500). An average was taken of the technical replicates and output XCMS data tidied in [R](#) (<https://untargeted-metabolomics-workflow.netlify.app/>).

Metaboanalyst was used to perform statistical analysis of each antibiotic treatment against the antibiotic free control. Data was normalised with pareto scaling (Figure S5.1).

A random forest analysis is used to identify significantly affected metabolites using a bootstrapping algorithm which combines ensemble learning methods with the decision tree framework to create multiple randomly drawn decision trees from the data and averaging the results. This analysis produces strong predictors of grouping, or

treatment, ranked by variable importance. The analysis ran with 7 predictors and 1000 trees.

Metabolites are reported if present in all 5 biological replicates and for fold change data, must meet or exceed a threshold of ± 2.0 to be reported. Here we define 'significance' as metabolites that have been differentially expressed by a fold change $\geq \pm 2.0$.

Any metabolites highlighted in statistical analysis were putatively identified using [METLIN](#), [KEGG](#) and [ECMD](#). The databases METLIN or ECMD were searched with the m/z values, and must be agreed upon by with the KEGG pathway metabolism map for *E.coli* in order to be reported.

Using this method, the total across strains of the significant metabolites in the identifiable range by treatment are as follows: Ciprofloxacin, 79.55 %. Cefotaxime, 93.97 %. Kanamycin, 78.36 %.

Of the metabolites in the identifiable range, the percentage of those with a putative identification by treatment are as follows: Ciprofloxacin, 33.33 %. Cefotaxime, 30.07 %. Kanamycin, 19.01 %.

It should be noted that masses representing these metabolic pathways sometimes have multiple putative identifications which have identical or close ppm, making the compound identity indistinguishable between the options, meaning the exact identity of some of the metabolites is uncertain. In every instance, the identification with the lowest ppm possible is used.

5.4 Results

In this study, 4 *E.coli* strains into which an MDR plasmid had recently (within 30 generations) been conjugated, were exposed to 4 treatment conditions; a sub-MIC of one of 3 antibiotics and an antibiotic-free control. The treatments were ciprofloxacin, cefotaxime and kanamycin, representing 3 antibiotic classes. The metabolome was arrested during exposure to the treatment, and untargeted metabolomic analysis performed to obtain an unbiased assay of the bacterial metabolome under active antibiotic stress. All treatment groups were different to the antibiotic-free control (PCAs supp S5.5, S5.6, S5.7, S5.8) and had significant differences to the control metabolome, when significance is defined as a fold change $\geq \pm 2.0$.

Antibiotic induced disruption varies by strain.

Overall, when all strain responses were totalled for each antibiotic treatment, ciprofloxacin caused the most metabolic disruption and significantly altered 624 metabolites. This was much higher than Kanamycin, which caused differential expression of 357 metabolites, and nearly twice the 315 metabolites altered by cefotaxime treatment. This correlated with the resistance encoded on pLL35, which encoded 1 gene for quinolones, compared to 2 for beta-lactams and 3 for aminoglycosides (Table 5.1).

The metabolic responses of each *E.coli* strain to antibiotic exposure varied by antibiotic class and were strain specific (Fig 5.1, 5.2, 5.3, 5.4). For 3 out of 4 strains ciprofloxacin induced the largest metabolic responses, altering 10.71-12.42% of the recorded metabolome (Table 5.2). The exception was strain F022, where just 4.92% of the recorded metabolome was significantly altered in response to ciprofloxacin (Fig 5.1 Table 5.2). This strain was most impacted by cefotaxime, which induced a significant response in 5.26% of the recorded metabolome and was the highest response to cefotaxime across any of the strains (Fig 5.1 FC Table 5.2). The response to kanamycin was the lowest of the antibiotics in 3 of 4 strains, altering 2.7 – 3.95% of the recorded metabolome (Table 5.2). The exception to this is the second clinical strain, F054 which had 10.71% of the recorded metabolome altered by kanamycin, the

highest response to this treatment across any of the strains (Fig 5.2 Table 5.2). These data demonstrate differing levels of disruption in the metabolic network in a strain dependent manner, despite carrying the same resistance plasmid.

Functional parallelism

Overall, these results highlight commonly affected functional groups between treatment groups and between strains in each treatment group. Within antibiotic treatments, the proportion of identified metabolites that were common in 2 strains or more was 37.93% in the ciprofloxacin treatment, 27.12% in the cefotaxime treatment and 10.42% in the kanamycin treatment. However, the specific pathways, direction of change and extent of change is variable by strain and treatment. Among all treatment groups, the main metabolic function categories were amino acid metabolism, nucleotides metabolism, energy metabolism and cofactor metabolism (Fig 5.5, 5.6, 5.7) but the distribution of affected metabolites across these categories was highly strain specific.

Amino Acid Metabolism

Amino acid metabolism commonly features in bacterial stress responses (Zhao *et al.*, 2019) and in these data amino acid biosynthesis and catabolism was featured in all strains and treatments (Fig 5.5, 5.6, 5.7). Most of the amino acid metabolism pathways identified in the dataset were featured in every antibiotic treatment. Common to all treatments was metabolism of: valine, cysteine, arginine, tryptophan, phenylalanine, lysine, ornithine, proline, isoleucine, methionine. Common to ciprofloxacin and cefotaxime treatments only was metabolism of serine, threonine and histidine. Common to ciprofloxacin and kanamycin treatments only was metabolism of leucine and tyrosine and unique to kanamycin was metabolism of alanine.

There are some examples of a uniform parallel response, for example L-tryptophan is upregulated to similar levels (FC range 2.3 – 3.7) in all 4 strains in response to cefotaxime (Table S5.2). However, as in the case of the arginine biosynthesis pathway, the same metabolite can be present in multiple strains and treatment groups but is altered in a different way in different strains.

Arginine metabolism was highlighted in all antibiotic treatments and in particular the metabolites pyrroline hydrocarboxylic acid and citrulline were seen repeatedly. Although sometimes these masses had other potential identifications, the identifications discussed here had the lowest ppm and are therefore the most likely. Targeted metabolomic analysis would be needed to determine these identifications with certainty. Citrulline, a key component of the arginine biosynthesis pathway, was identified in multiple strains and treatments. In strains ELU39, F022 and MG1655 citrulline was downregulated in response to ciprofloxacin (Table S5.3). In ELU39 alone, citrulline was upregulated in response to kanamycin (Table S5.1). In the environmental strain ELU39 Pyrroline hydrocarboxylic acid was upregulated in response to kanamycin but downregulated in response to ciprofloxacin (Tables S5.1, S5.3). This same metabolite was downregulated in the F054 response to kanamycin, an example of an opposite response to the same stressor by a clinical strain compared to the environmental strain (Table S5.1). In the identified metabolome this metabolite did not appear in F022, demonstrating the highly strain specific ways the metabolism can change even when the strains have historically been exposed to a similar environment.

All strains responding to ciprofloxacin had a downregulated (FC range 0.09 – 0.41) metabolite, 4(Glutamylamino)butanoate, which is involved in GABA biosynthesis, suggesting a demand on GABA (Table S5.3). However, this was indistinguishable from N2-succinyl-L-ornithine, which is involved in the arginine succinyltransferase pathway. In contrast, ELU39 upregulated the same metabolite (FC 2.16) in response to kanamycin (Table S5.1). GABA has a role as an acidity protectant, which arginine succinyltransferase is essential for aerobic arginine catabolism in *E.coli* and is one way of catabolizing ornithine. The downregulated metabolites in arginine and proline metabolism in response to ciprofloxacin indicate cellular stress (Zhao *et al.*, 2019).

Energy metabolism

All treatments indicated an increased demand on energy needed to overcome antibiotic stress. Cell wall modifications and increased expression of efflux pumps, among other resistance mechanisms, create an energy demand (Wagner *et al.*, 2007; Parsons and Rock, 2013; Da Silva and Domingues, 2017; Pacheco *et al.*, 2017; Zhao

et al., 2019). Central energy metabolism is heavily featured in the response to cefotaxime (Fig 5.6). Downregulation of components of the TCA cycle such as malic acid (Table S5.2) may indicate energy consumption (Laviña *et al.*, 2020). Amino acid degradation may also form part of the response to a demand in energy as carbon skeletons from degradation of amino acids can feed into the biosynthesis of essential coenzymes and intermediates for the TCA cycle.

Further evidence for an increase in energy consumption is that deoxyadenosine monophosphate, a by-product of the ATP-ADP reaction is a commonly upregulated metabolite. Deoxyadenosine monophosphate (dAMP) was significantly altered in all 4 strains in the ciprofloxacin treatment and all but ELU39 in the cefotaxime and kanamycin treatments. MG1655 had dAMP upregulated the most out of the strains, particularly in the ciprofloxacin treatment (FC 83.56) and the cefotaxime treatment (FC 373.07) (Table S5.2, S5.3). F054 again displays directional divergence from the trend, and had dAMP downregulated in the kanamycin (FC 0.44) and cefotaxime (FC 0.027) treatments (Table S5.1 S5.2).

Differentially regulated components of aerobic respiration are indicative of an energetic demand caused by induction of resistance mechanisms. Key components of oxidative phosphorylation were affected in all the treatments. Ubiquinone which is upregulated 40-fold in MG1655 in response to cefotaxime (FC 40.24), over 50-fold in response to kanamycin (FC 56.01) (Table S5.2, S5.3). In the MG1655 response to ciprofloxacin, ubiquinone is upregulated (FC 126.29), while simultaneously ubiquinol-1 is downregulated (FC 0.086). The downregulation of ubiquinol implies oxidative phosphorylation is taking place, because to produce energy ubiquinol is reduced to form ubiquinone, which is upregulated indicating an ongoing reaction at a higher rate than the control.

Pantothenic acid is a precursor of Coenzyme A, and a downregulation of pantothenic acid implies a demand on coenzyme A (López-Sámano *et al.*, 2020). Pantothenic acid is downregulated in all strains in response to ciprofloxacin (FC range 0.14 – 0.49) (Table S5.3). Strain F054 also downregulated pantothenic acid (FC 0.49) and an additional metabolite involved in coenzyme A biosynthesis, pantetheine 4'-phosphate

in response to cefotaxime, which together suggest a demand on coenzyme A biosynthesis (Table S5.2). In contrast, ELU39 upregulated pantothenic acid (FC 2.16) in response to kanamycin (Table S5.1).

Pyridoxal-P, PLP is essential for phosphorylytic cleavage of glycogen and amino acid degradation (Berg *et al.*, 2019). In the cefotaxime treatment, a metabolite involved in its biosynthesis, O-phospho-4-hydroxy-L-threonine is downregulated to very similar levels in ELU39 and F054 (FC 0.493 and 0.497 respectively), suggesting a demand on PLP in these strains (Table S5.2). In the ciprofloxacin treatment, 2-Amino-3-oxo-4-phosphonooxybutyrate is upregulated in MG1655 and F022 (FC 2.04 and 2.59 respectively) and twice in ELU39 (FC 2.49 and 3.49) (Table S5.3). However in this untargeted analysis, this mass is indistinguishable from 2-Aspartyl-4-phosphate, which is involved in the lysine, threonine and methionine biosynthesis pathway.

Biotin is an essential cofactor of carboxylases, decarboxylases and transcarboxylases, enzymes involved in fatty acid synthesis, gluconeogenesis and amino acid metabolism (Sirithanakorn and Cronan, 2021). Biotin was a unique metabolite to the ciprofloxacin response, and was upregulated in the strains ELU39, F022 and MG1655 (FC range 2.44 – 4.84) (Table S5.3).

Nucleotide Metabolism

There are parallel responses in nucleotide and energy metabolism reflected in significant changes in purine and pyrimidine metabolism in every antibiotic treatment. Purine metabolism was highlighted 5, 7 and 10 times in the kanamycin, cefotaxime and ciprofloxacin treatments respectively. Pyrimidine metabolism was highlighted 5, 4 and 13 times in the kanamycin, cefotaxime and ciprofloxacin treatments respectively. Pyrimidine and purine are fundamental parts of DNA and RNA synthesis as essential components of nucleic acids and are essential to ATP structure and signal transduction (Berg *et al.*, 2019). Furthermore, in the response to ciprofloxacin, cytidine and its precursor cytosine are identified multiple times due to different ionisation adducts such as hydrogen or sodium, or slightly different recorded masses (<0.01 m/z). Cytosine is upregulated in all 4 of the strains (FC 2.62 – 9.65) and cytidine is upregulated in F054, F022 and MG1655 (FC 2.12 – 7.89) (Table S5.3). An exception is one of the 4 instances of cytidine in F022 is instead downregulated (FC 0.43) (Table S5.3).

Accumulation of a metabolite can be an indication of redundancy (Laviña *et al.*, 2020). Therefore, accumulation of cytidine suggests a decreased use of this metabolite for nucleotide biosynthesis. This suggests interrupted transcription or replication which would follow given ciprofloxacin binds to DNA topoisomerase (Ojkic *et al.*, 2020).

Bacterial resistance mechanisms occur alongside plasmid-encoded resistance.

Peptidoglycan is a key component of the bacterial cell wall and its amplification is one of the mechanisms of resistance to aminoglycosides (Plumbridge, 2015; Garneau-Tsodikovaa and Labby, 2016). N-acetyl-D-muramoate, a precursor for peptidoglycan biosynthesis, was upregulated over 500-fold (FC 565.44) in response to Kanamycin in clinical strain F022 (Table S5.1). This metabolite was unique to strain F022. This contrasts with the plasmid-encoded resistance, which focuses on modification of the kanamycin molecule. N-acetyl-D-muramoate was also upregulated in all 4 strains in the ciprofloxacin treatment, but the fold change varied drastically. F022, ELU39 and MG1655 all had several hundred-fold upregulation of this metabolite: FC 423.59 in F022, FC 628.07 in ELU39 and FC 841.6 in MG1655, the biggest alteration of any single metabolite seen throughout this thesis (Table S5.3). Contrastingly, strain F054 had a fold change of 3.43 for N-acetyl-D- muramoate (Table S5.3).

In the cefotaxime treatment, all 4 strains displayed a similar level (FC 3.12 – 5.42) of upregulation for a mass whose identity could not be distinguished by ppm between 3'-AMP, AMP or dGMP (Table S5.2). 3'AMP is a derivative of cAMP, which has a role in biofilm regulation. Biofilm formation is a common response to antibiotic exposure that increases bacterial tolerance to antimicrobials (Crabbé *et al.*, 2019). This metabolite was also upregulated in strains F022 and ELU39 (FC 3.02 and 2.08 respectively) in the kanamycin treatment (Table S5.1). cAMP or a derivative also appears downregulated (FC 0.49) in strain F022 and a derivative of cAMP is upregulated (FC 2.0) in strain F054 in the ciprofloxacin treatment (Table S5.3). The appearance of this metabolite in all 3 treatments for strain F022 suggests altering this pathway is important in the response to antibiotic stress for this strain, although the direction of change varies with the antibiotic.

Table 5.2 A table detailing the percentage of the recorded metabolome by strain that was significantly (± 2.0 Fold Change) impacted by the antibiotic treatment of ciprofloxacin, cefotaxime and kanamycin when compared to the antibiotic free control. The numbers of the significantly altered metabolites that are up or downregulated compared to the control are also recorded.

Treatment	Strain	% Metabolome	Upregulated	Downregulated
Ciprofloxacin	F054	12.42 %	31	124
Ciprofloxacin	MG1655	12.22 %	45	126
Kanamycin	F054	10.71 %	12	144
Ciprofloxacin	ELU39	10.20 %	76	106
Cefotaxime	F022	5.26 %	116	6
Cefotaxime	ELU39	5.21%	59	43
Ciprofloxacin	F022	4.92 %	60	56
Kanamycin	MG1655	3.95 %	67	13
Cefotaxime	MG1655	3.83 %	18	30
Cefotaxime	F054	3.43 %	8	35
Kanamycin	F022	3.32 %	66	8
Kanamycin	ELU39	2.7 %	45	2

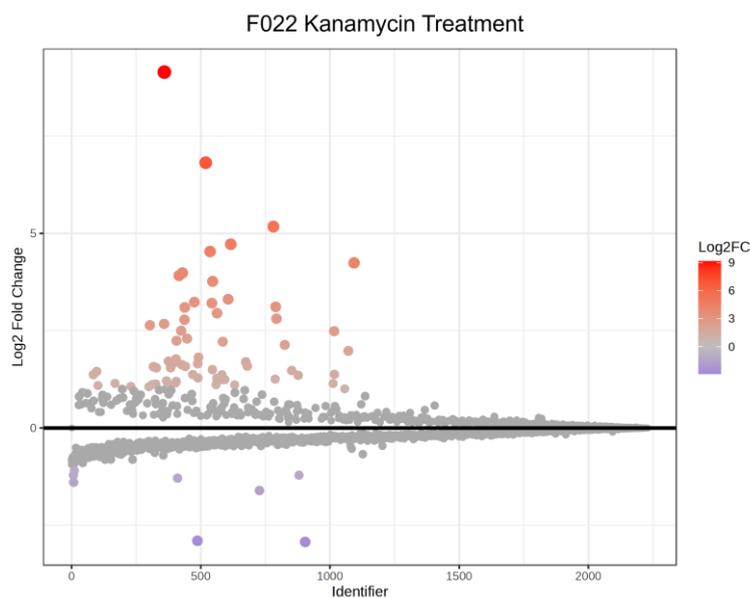
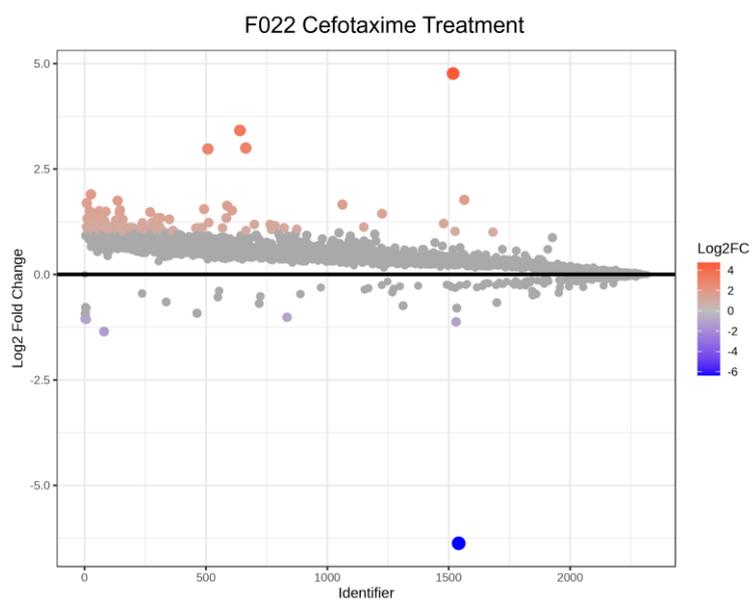
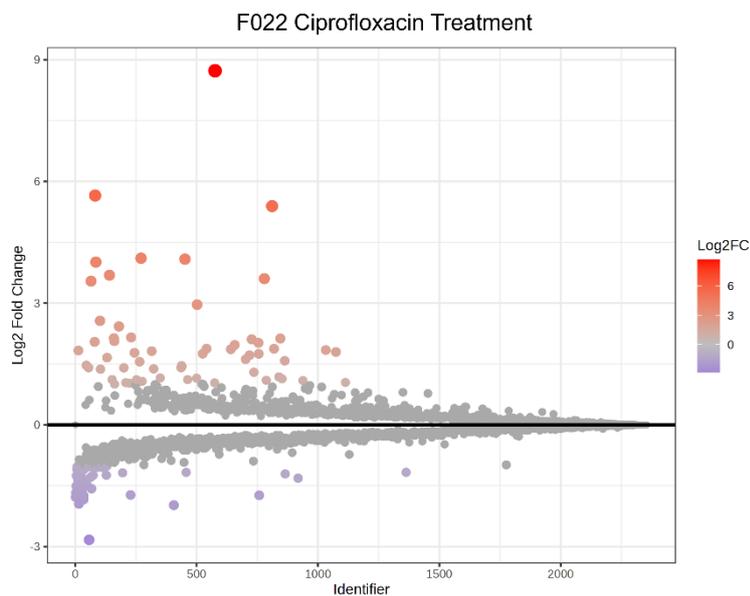


Figure 5.1. Fold change maps for strain F022 of the metabolomes exposed to the antibiotic treatment ciprofloxacin, cefotaxime and kanamycin each compared to the control group for that strain, highlighting significant metabolites. The threshold for significance is defined as the fold change ± 2.0 . See table 5.2 for further values.

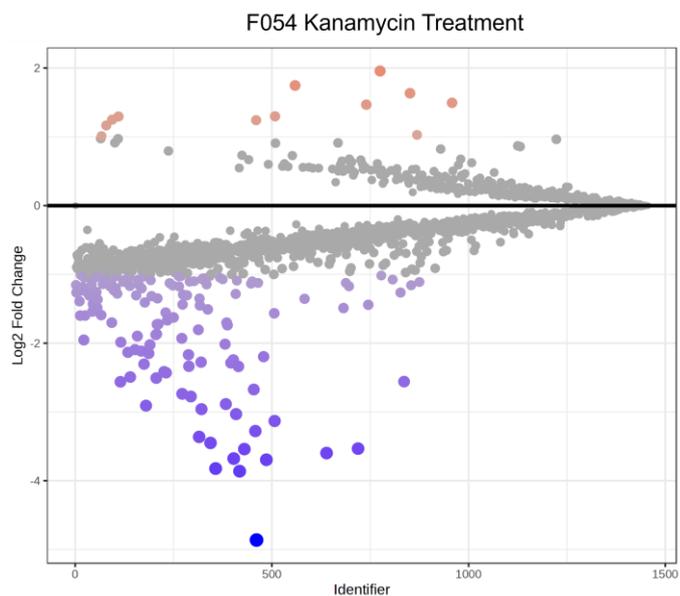
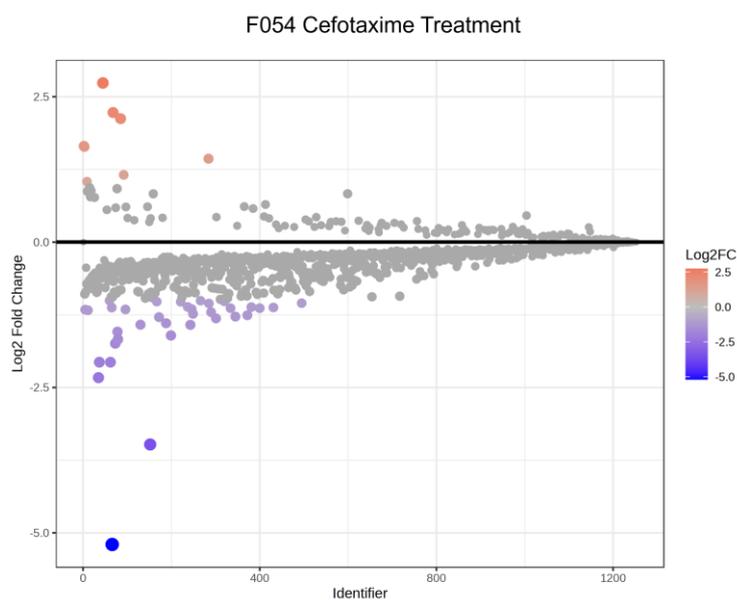
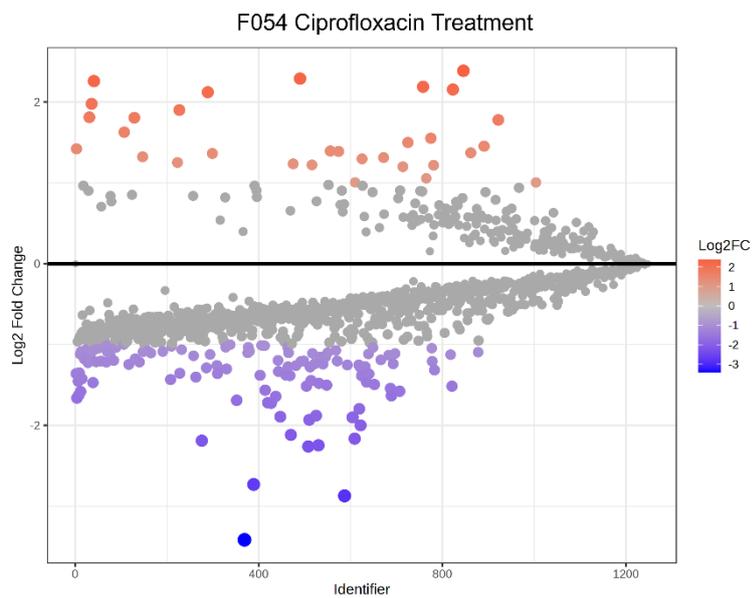
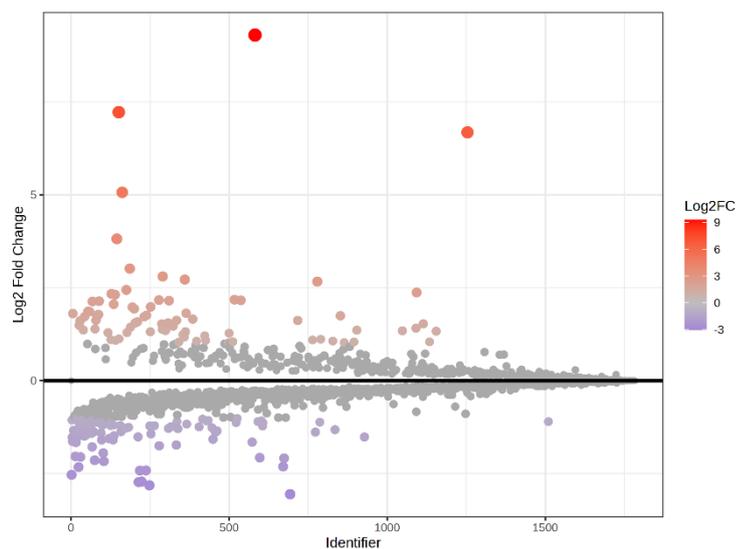
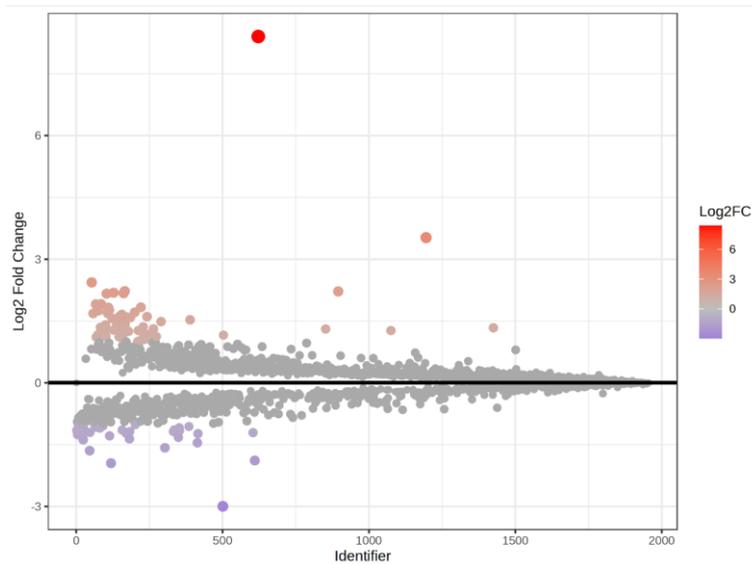


Figure 5.2. Fold change maps for strain F054 of the metabolomes exposed to the antibiotic treatment ciprofloxacin, cefotaxime and kanamycin each compared to the control group for that strain, highlighting significant metabolites. The threshold for significance is defined as the fold change ± 2.0 . See table 5.2 for further values.

ELU39 Ciprofloxacin Treatment



ELU39 Cefotaxime Treatment



ELU39 Kanamycin Treatment

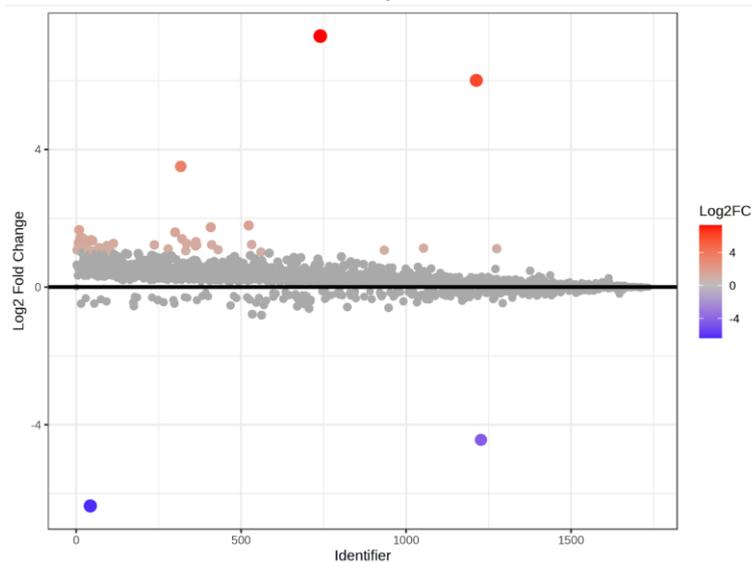


Figure 5.3. Fold change maps for strain ELU39 of the metabolomes exposed to the antibiotic treatment ciprofloxacin, cefotaxime and kanamycin each compared to the control group for that strain, highlighting significant metabolites. The threshold for significance is defined as the fold change +/- 2.0. See table 5.2 for further values.

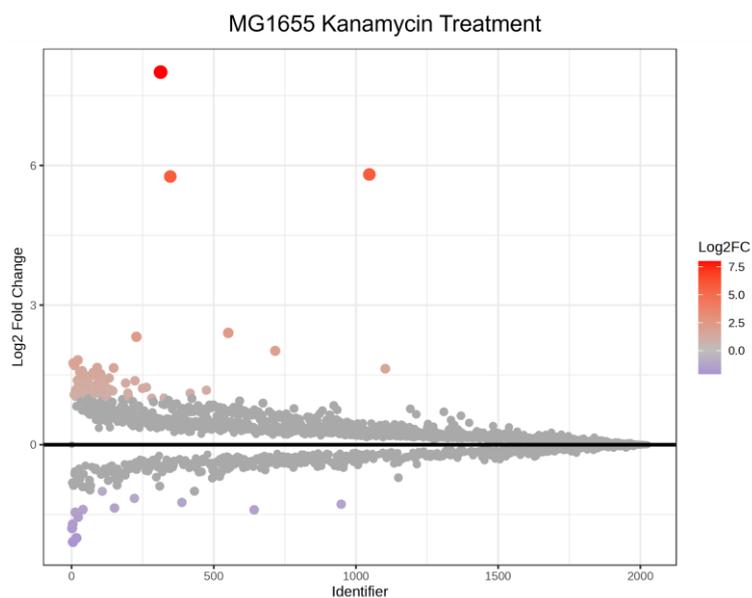
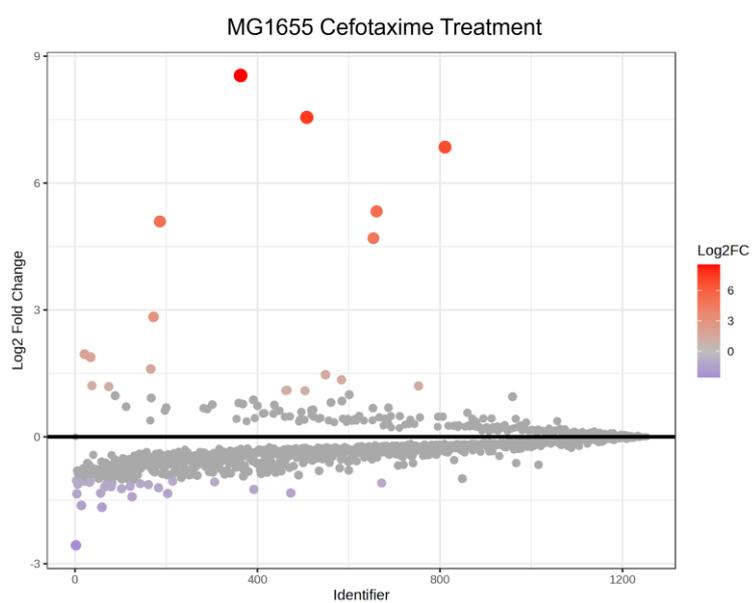
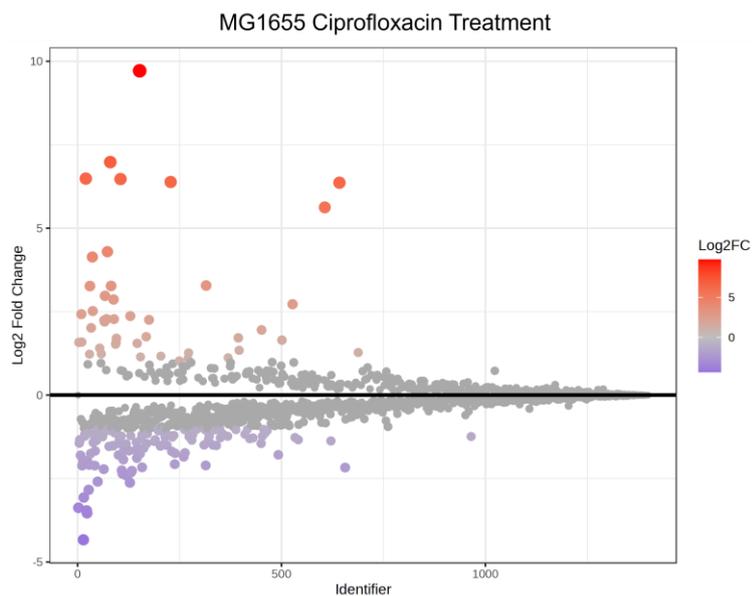


Figure 5.4. Fold change maps for strain MG1655 of the metabolomes exposed to the antibiotic treatment ciprofloxacin, cefotaxime and kanamycin each compared to the control group for that strain, highlighting significant metabolites. The threshold for significance is defined as the fold change ± 2.0 . See table 5.2 for further values.

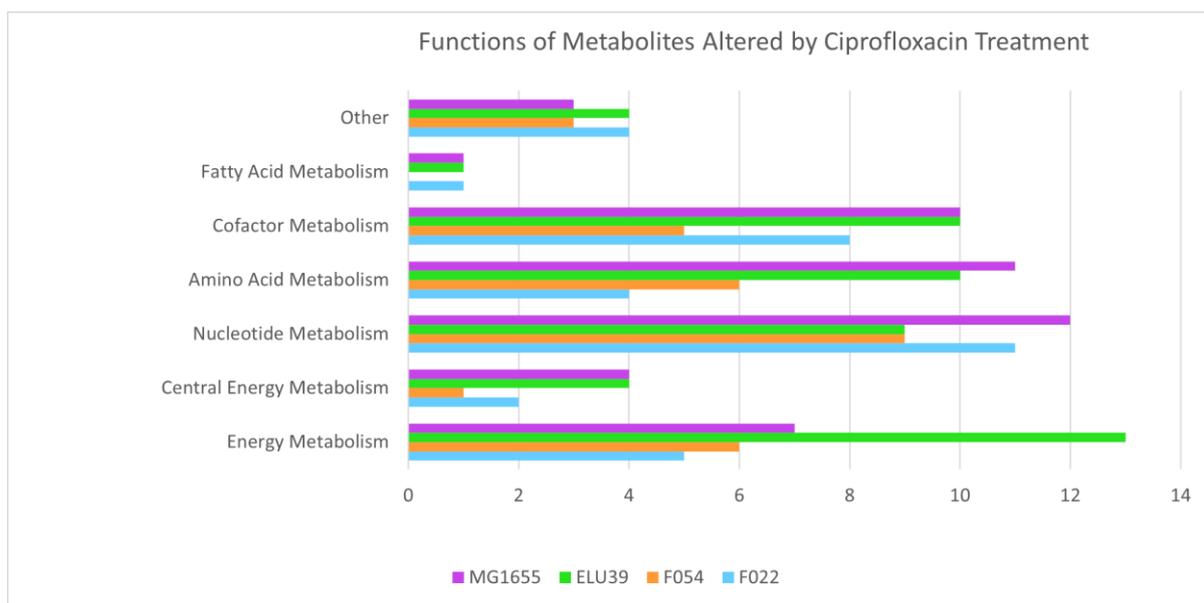


Figure 5.5. A bar chart of the functional categories of identified metabolites significantly altered by the ciprofloxacin treatment, coloured by strain: purple MG1655, green ELU39, orange F054 and blue F022.

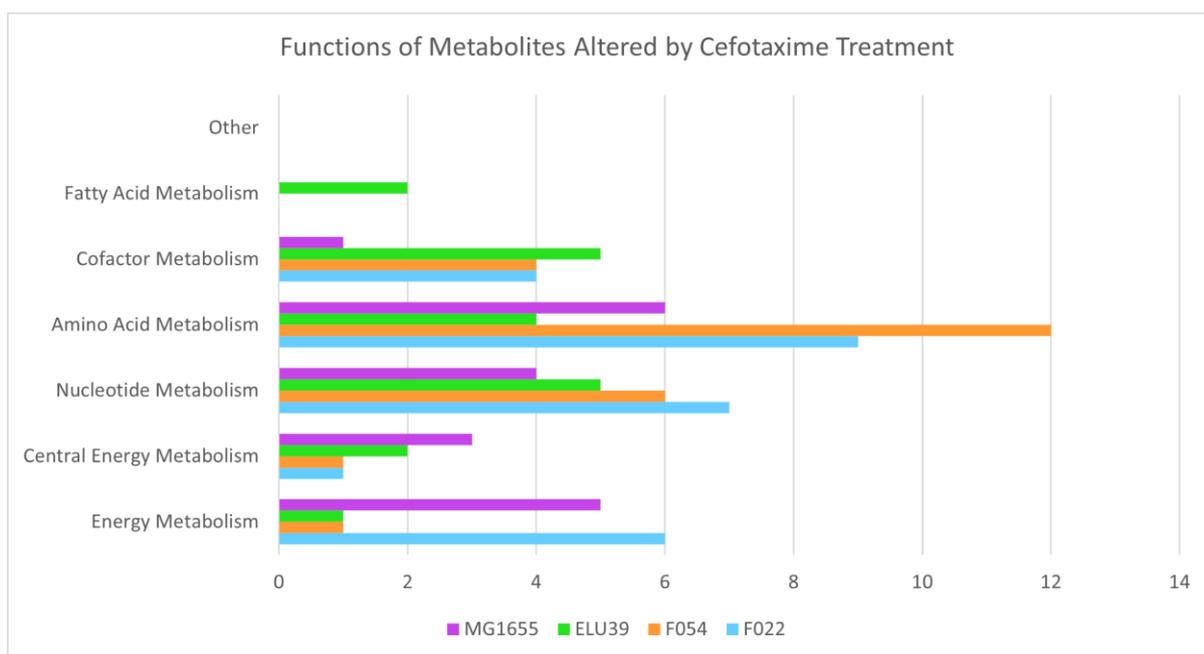


Figure 5.6. A bar chart of the functional categories of identified metabolites significantly altered by the cefotaxime treatment, coloured by strain: purple MG1655, green ELU39, orange F054 and blue F022.

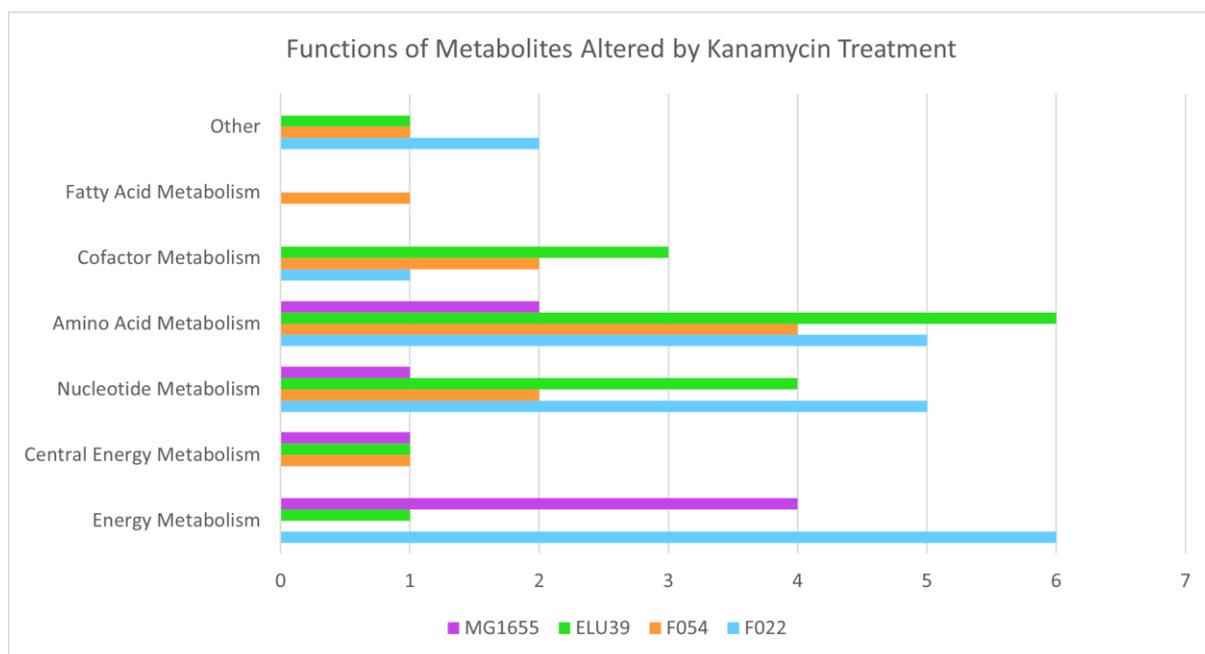


Figure 5.7. A bar chart of the functional categories of identified metabolites significantly altered by the kanamycin treatment, coloured by strain: purple MG1655, green ELU39, orange F054 and blue F022.

5.5 Discussion

Trends in antibiotic disruption

The metabolic effects of antibiotic exposure were assessed in 4 *E.coli* strains from different ecological backgrounds carrying the MDR plasmid pLL35. For most of the strains, the ciprofloxacin treatment caused the largest disruption to the metabolic network, followed by the cefotaxime treatment and the kanamycin treatment, which correlates to the number of plasmid-encoded resistance genes for each antibiotic class. The clinical strains diverged from this trend in different treatments. Strain F022 displayed less than half the disruption in response to ciprofloxacin than the other strains. Clinical strain F054 displayed a 10.71% disruption in response to kanamycin, which is 2-3 times the proportion altered by this treatment in the other strains.

Of the recorded metabolomes in the identifiable range, 19.01-33.33% of significant metabolites were assigned a putative identification. Despite extensive advancements in metabolomics techniques over recent years, full identification of microbial metabolomes is not yet possible (Kok *et al.*, 2022). The identifiable metabolome is reflective of an energetic demand, perturbations to amino acid, nucleotide and cofactor metabolism and the activity of bacterial resistance mechanisms.

Parallel functions

Within each treatment, the metabolic functions altered by antibiotic stress displayed a high degree (10-37%) of parallelism between strains. Since all the strains studied here belong to the same bacterial species this is not unexpected. Additionally, broad functions such as amino acid metabolism, energy metabolism, nucleotide metabolite and cofactors were common to all 3 treatment groups. This illustrates general stress responses, rather than a response to a specific antibiotic. Some of these functions were also present in the plasmid acquisition metabolome (Chapter 3) and in the metabolomes of strains evolved under cefotaxime selection (Chapter 4). The extent of disruption caused by antibiotic exposure is generally higher, up to 12.42% of the metabolome, compared to plasmid acquisition, up to 1.36 % or coevolution with a plasmid, up to 2.6% (Table 5.2, Chapters 3 and 4). However, there remains a high

degree of strain specificity when assessing the pathways and metabolites affected, and the extent and direction in which they are affected.

Amino acid metabolism is a function affected by antibiotic exposure in this data in every strain and treatment, and it was also a consistent feature of plasmid acquisition and co-evolution (Chapter 3 and Chapter 4). Lysine is one of the few amino acids that features throughout this thesis. In the plasmid acquisition data, lysine degradation is upregulated in MG1655 and downregulated in one of the clinical strains, F104. In the clinical strain F022 lysine metabolism is downregulated as a result of long-term plasmid carriage and upregulated after long term exposure to cefotaxime. In the same strain, F022, lysine biosynthesis is upregulated in response to cefotaxime. This suggests lysine is important in maintenance of a plasmid and responses to some stressors. Further investigations would be needed to fully understand its role.

The GABA biosynthesis pathway was downregulated in all strains in the ciprofloxacin treatment, which is indicative of a demand on the end product to combat ciprofloxacin. In metabolomics, a lower intensity can mean that function is in demand and therefore the free form of the metabolite is decreased as it is transformed through use (Laviña *et al.*, 2020). However, in a different treatment, the same metabolite in ELU39 was upregulated instead. In the post evolution data, a metabolite in the same pathway was altered in the same direction in the same strain (Chapter 4). In combination with the data from Chapter 4, this instead suggests that this pathway in the strain ELU39 may be more associated with plasmid carriage than antibiotic stress.

Evidence of an energetic demand

The metabolomes of all strains in each treatment displayed evidence of an energetic demand. This was evidenced by differentially expressed amounts of metabolites belonging to the TCA cycle, aerobic respiration and the biosynthesis of coenzymes which facilitate those reactions. Aerobic respiration, production of ATP through glycolysis and the TCA cycle is the most efficient energy production pathway in bacterial metabolism (Passalacqua, Charbonneau and O’Riordan, 2016). This pathway is evidenced by cofactors such as PLP and biotin, which have roles in energy production, biosynthesis of CoA and malic acid, which are key to the TCA, and finally

the by-product of ATP consumption. In addition, energy production was a function impacted by the acquisition of a plasmid, implying increased energetic demands (Chapter 3). Energy is needed to overcome stressors, so it follows that the bacteria increased the production of energy to combat both the antibiotics and the costs of plasmid acquisition (Zhao *et al.*, 2019).

While all treatments displayed significant alteration in these metabolites, the cefotaxime and ciprofloxacin treatments seemed particularly affected, for example PLP and biotin are unique to these two treatment groups. Some caution is required with this inference however, as kanamycin also had the lowest proportion of assigned identifications. Pyridoxal-P, (PLP) is an important cofactor for regulating energy production from glycogen and amino acids because it is essential for phosphorylytic cleavage of glycogen and amino acid degradation. The biosynthesis of PLP is highlighted by downregulation of PLP precursors in ELU39 and F054 strain in the cefotaxime treatment, and upregulation of the same metabolite in ELU39, F022 and MG1655 in the ciprofloxacin treatment. Further confirmation of this identification would be necessary however because there are alternative identifications related to amino biosynthesis that are equally likely per this analysis.

Ciprofloxacin induced upregulation of biotin which can suggest energy is being derived from fatty acid metabolism. The cofactor biotin is essential for some carboxylases involved in fatty acid metabolism, amino acid metabolism and carbohydrate metabolism (Tong, 2013; Sirithanakorn and Cronan, 2021). Acquisition of the plasmid induced mutations in genes involved in anaerobic metabolism, of which fatty acid metabolism is a part (Dunn *et al.*, 2021). This was reflected in the plasmid acquisition metabolomics of 3 strains (F054, F104 and MG1655) (Chapter 3). Coevolution with a plasmid also resulted in mutations in anaerobic metabolism and fatty acid metabolism (Carrilero, Dunn and Moran, 2023) although this was not reflected in the metabolome (Chapter 4).

Ubiquinol and ubiquinone are integral parts of the electron transport chain, which both facilitates aerobic respiration and maintains redox homeostasis within the cell (Berg *et al.*, 2019). Ubiquinone is consistently featured throughout the data in this thesis as a

molecule important to *E. coli* for acclimating to plasmid carriage upon acquisition and over evolutionary time (Chapter 3 and 4). The active reaction is clear in strain MG1655 in the ciprofloxacin treatment group which saw significant downregulation of ubiquinol because it is reduced and transformed into ubiquinone, which was upregulated over 100-fold. It can be inferred that other changes of these molecules in the same direction are indicative of this metabolic reaction. Further, it confirms that downregulation of a metabolite is indicative of a demand on the process of product that requires its alteration, and is therefore consumed, agreeing with Laviña et al (2020).

Intracellular conditions are reduced and proteins are kept in the reduced state to facilitate DNA synthesis, energy production and protein repair (Arnér and Holmgren, 2000). Furthermore, redox stress is both a consequence of antibiotics and a factor that increases lethality (Brynildsen *et al.*, 2013; Dwyer *et al.*, 2014). Metabolically active cells experience increased bactericidal effects from fluoroquinolones due to the formation of reactive oxygen species (ROS) during oxidative phosphorylation (Gutierrez *et al.*, 2017). Additionally, impeding cell wall synthesis and initiating a futile cycle of peptidoglycan degradation, beta-lactams induce an energy demand and systemic cellular toxicity associated with ROS (Cho, Uehara and Bernhardt, 2014; Dwyer *et al.*, 2014; Adolfsen and Brynildsen, 2015).

Although many of the energy production pathways are common between strains, strain specificity remains prominent. For example, strain F022 has alterations to the amounts of metabolites that supply glycolysis, which is a consistent feature for this strain only in every treatment group. Many of the metabolites related to glycolysis have numerous putative identifications that are indistinguishable from each other, but the broader function of glycolysis supply is agreed upon, and is different than for example energy production from amino acid catabolism and therefore changes unique to this strain to the production of energy within the cell can be inferred. The clinical strains often respond differently to the same stressor throughout the dataset. While there are two strains that come from a similar background of clinical bacteremia, these environments could still have been vastly different and therefore the way the metabolic network adapted to respond to the same stressors is different. Interestingly, only F022 displayed an increased growth rate upon acquisition of the plasmid and after

coevolution under cefotaxime selection (Dunn *et al.*, 2021; Carrilero, Dunn and Moran, 2023).

There is a high level of parallelism of the functions impacted when *E.coli* is exposed to antibiotics, but strain specificity is present at the pathway and metabolite level in the extent and direction in which they are changed. Such diversity in response, even within the same species and plasmid pairing, emphasizes the challenge faced by the research of novel combination therapies. So far, this avenue of research is promising, and aims to use metabolites to manage resistant infections and re-sensitize bacteria to antibiotics (Peng *et al.*, 2015; Meylan, Andrews and Collins, 2018; Crabbé *et al.*, 2019). This data demonstrates the importance of building a metabolomic database to develop clearer trends and allow hypotheses to be formed for treatment research (Kok *et al.*, 2022).

Bacterial resistance occurs alongside plasmid-encoded genes

There is evidence in the metabolome of active bacterial resistance mechanisms, which are being upregulated by the bacteria alongside plasmid encoded resistance. This demonstrates that bacterial resistance mechanisms are not redundant upon acquisition of an MDR plasmid.

Biofilm formation is a bacterial strategy that reduces sensitivity to antibiotics (Crabbé *et al.*, 2019). Metabolomics revealed metabolites repeatedly identified as components of the cell wall and biofilm regulator molecules or their direct derivatives. In the cefotaxime treatment all four strains have a higher level of a cyclic AMP (cAMP) derivative than the controls, potentially indicating higher cAMP use. The common response in the clinical and environmental strains as well as MG1655 after evolution under cefotaxime selection (Chapter 4) was also an upregulation of cAMP.

Cyclic AMP has multiple regulatory roles in *E.coli* including virulence, biofilm formation, type II secretion and carbon metabolism (Berg *et al.*, 2019). Biofilm formation has been shown to be influenced by plasmid carriage in *E.coli* (Shin and Ko, 2015; Schaufler *et al.*, 2016) and is a bacterial resistance mechanism (Jolivet-Gougeon and Bonnaure-Mallet, 2014) that is impaired by beta lactamase production (Gallant *et al.*, 2005). This would explain an increase of cAMP in response to cefotaxime but not the other

antibiotics, where the same metabolites were present in some of the strains. F022 in particular had the cAMP derivative upregulated in every treatment. Selection induced upregulation of cAMP in MG1655, F022 and ELU39, showing this altering of the metabolome remained as a consequence of long-term cefotaxime exposure, as these data were measured in the absence of an antibiotic (Chapter 4).

Peptidoglycan biosynthesis was also upregulated in the Kanamycin and ciprofloxacin treatments, evidenced by upregulation of precursor metabolites. N-Acetyl-D-muramoate was the most dramatically upregulated of these metabolites, approximately 400-800 fold by 3 of the 4 strains. This is by far the largest change to a single metabolite seen throughout this thesis. F054 stood apart in the ciprofloxacin treatment with just a 3-fold upregulation of this metabolite. Since ciprofloxacin does not target the cell wall, the strength of this response was perhaps surprising. N-Acetyl-D-muramoate is a precursor molecule of the two repetitive backbone units of peptidoglycan, NAM. The second of these, NAG is formed by N-Acetyl-D-Glucosamine, which is not in the data except as an alternative identification for a mass that had more likely identifications based on ppm (Table S5.3). While present in all strains, these masses were also not upregulated to the same extent as N-Acetyl-D-muramoate. F022 and MG1655 had this mass upregulated 2.6-fold and 4.6-fold respectively, and it was downregulated in ELU39 and F054 0.43-fold and 0.46-fold respectively. Further research would be necessary to understand the dynamics of the cell wall components in this setting.

Many of the significantly altered metabolites seen in this chapter are intermediates, precursors and derivatives of essential metabolites. This is similar to previous chapters in this thesis and is unsurprising because metabolite levels are highly conserved in *E.coli* even when faced with perturbations the flexible metabolic network can use alternative pathways to minimise disruption (Ishii *et al.*, 2007; Bennett *et al.*, 2010). This is one of the qualities which allow *E. coli* to acclimate to the difference in conditions in different areas of the human body (Mann *et al.*, 2017). Additionally in this chapter, the amounts, relative to the control groups, of some essential metabolites were altered. Examples of these metabolites include L-tryptophan, cytidine and biotin. This is reflective of the larger metabolic disruption caused by antibiotic exposure.

This study examined *E.coli* in the mid-log phase, which was necessary to compare this data with that of previous chapters. It has been suggested that the stationary phase is the optimal growth phase for observation of changes in the metabolome in response to different stress treatments (Szymanski, Jędrzej Józefczuk *et al.*, 2009). However, this study was focused on temperature, oxidation and lactose diauxic stressors rather than antibiotics (Szymanski, Jędrzej Józefczuk *et al.*, 2009).

This data shows that bacteria carrying resistance plasmids still deploy chromosomally encoded resistance mechanisms to mitigate the impact of the antibiotic. Gain of resistance through plasmid transfer is not an evolutionary endpoint as the bacteria will contribute to respond to the stresses of antibiotics and selection. This helps to explain why resistance mutations continue to evolve in the chromosome after plasmid gain (Bottery, Wood and Brockhurst, 2019) and how antibiotics fuel a cycle of mutation and horizontal transfer that allows resistance gene accumulation and therefore the evolution of multidrug resistance (Coluzzi *et al.*, 2023).

The *E.coli* strains F022, F054, ELU39 and MG1655, carrying the MDR plasmid pLL35 were exposed to antibiotic treatments of 3 different classes; beta-lactam, aminoglycoside and fluoroquinolone. The metabolome of the strains was analysed using untargeted LC-MS metabolomics, which revealed a multi-layered response to antibiotic stress that evidences cellular disruption, energetic demands and bacterial resistance mechanisms.

While many functions, pathways and specific metabolites within these pathways are commonly affected in the strains, the extent and direction of the perturbation is strain specific. This demonstrates the complexity of a metabolic network and the nuance in the response to a common stressor. This has implications for future research on innovative combination therapies.

Chapter 6 : General Discussion

6.1 Overview

Antimicrobial resistant infections are predicted to claim 10 million lives a year by 2050 (*Antimicrobial Resistance*, 2021). Understanding the evolutionary mechanisms of AMR emergence is important if we are to create actionable interventions to the AMR crisis. Increased use of antimicrobial drugs is not a viable long-term strategy because it increases the selection pressure for the evolution of resistance (Cantón, González-Alba and Galán, 2012). New mechanisms of resistance evolve continuously, such as the New Dehli metallo-B-lactamase (NDM) which is able to hydrolyze almost all B-lactams and is already found on plasmids carried by *Escherischia coli* and *Klebsiella pneumoniae* (Wu *et al.*, 2019). Moreover, the rate of discovery of novel antibiotics has slowed in recent decades, as antibiotic discovery platforms are largely spent, and no longer producing many antibiotics that reach approval succeeding clinical trials (da Cunha, Fonseca and Calado, 2019).

This global health crisis is exacerbated by the horizontal spread of AMR genes by plasmids. In particular, *E.coli* carrying multidrug resistance plasmids commonly from the IncF family have emerged as some of the most difficult to treat infections in humans and animals globally, driving wider use of last resort antibiotics in hospitals. Plasmids accelerate the evolution of multidrug resistance (MDR) because they can carry multiple resistance genes often against different antibiotic classes and transfer these horizontally between bacterial lineages (Hall, Brockhurst and Harrison, 2017). Thus, new strategies that minimize or even block plasmid transmission between lineages are being investigated. This could involve methods such as plasmid curing or interrupting the process of plasmid dissemination (Buckner, Ciusa and Piddock, 2018; Ragheb *et al.*, 2019). However, the development of interventions, from individual treatments to limiting the spread of resistance on a global scale, require a full understanding of the relationship dynamics between plasmids and their bacterial hosts (Ragheb *et al.*, 2019; Krell and Matilla, 2022).

Metabolomics is proving a useful tool for gaining a systems level understanding of how bacterial cells respond to stressors, including antibiotics. For example, uncovering physiological strategies used by bacteria to tolerate antibiotics with a view of exploiting them to increase the efficacy of existing antibiotics (Peng *et al.*, 2015; Meylan, Andrews and Collins, 2018). To date, metabolomics has been under utilized to understand how bacteria respond to gaining MDR plasmids and subsequently how this is altered by evolution and exposure to antibiotics.

This thesis has explored the metabolic responses of *E. coli* strains carrying the IncF MDR plasmid pLL35 across important stages in the life history of the bacteria-plasmid relationship; upon acquisition of the plasmid (Chapter 3), following subsequent coevolution (Chapter 4) and during exposure to antibiotics (Chapter 5). This work builds on previous work describing the transcriptome and evolutionary responses to pLL35 acquisition, together building a comprehensive multi-omics picture of bacterial-plasmid interactions in a clinically important system. Untargeted metabolomics was used throughout, as a method to obtain a comprehensive picture of all the metabolites in a biological system. In the following sections I briefly summarise the key findings of this thesis:

6.2 Chapter 2 | Metabolomics methodology development

To achieve this, I collaborated on the development of a new analytical pathway which brought together multiple tools into one streamlined methodology and removed the dependence of users on proprietary software (Parker *et al.*, 2023). An online, open access guide was created to help other researchers use untargeted metabolomics. This informs and aids decision making from wet lab protocols, obtaining metabolomic data and the processing and analysis that follows. This technology is valuable for understanding biochemical underpinnings of phenotype as the metabolism is the biological level most closely associated with function (Goodacre *et al.*, 2004; Ryan and Robards, 2006). Using this toolkit, I have discovered novel aspects of bacteria-plasmid interactions as mediated through metabolism.

6.3 Chapter 3 | Strain specific metabolic responses of diverse *E.coli* lineages to the acquisition of a multi-drug resistant plasmid.

Chapter 3 conducted untargeted metabolomics on 9 *E.coli* strains from clinical (6), environmental (2) and lab (1) backgrounds to determine their immediate response to pLL35 acquisition. These bacterial hosts had previously been shown to have different relationships with the plasmid ranging from costly to beneficial and involving different degrees of transcriptomic alteration. By comparing plasmid free and plasmid carrying clones, the analysis revealed the metabolic impact of the acquisition of the MDR plasmid pLL35 on the bacterial metabolome. The results showed subtle effects, impacting <2% of the metabolome of each strain. Of the identified metabolites, commonly affected pathways emerged across strains, associated with ubiquinone biosynthesis, energy production, and amino acid metabolism. Yet, the direction of change was not consistent, meaning the precise ways in which even these commonly affected pathways were altered remained strain specific. The results suggest slight adjustments of the bacterial metabolism upon plasmid acquisition to conserve or produce energy and thus mitigate the impact on the cell. These findings suggest that metabolic adaptation can negate significant costs to plasmid carriage and the necessity for genetic amelioration through mutations. Taken together with previous studies on these strains, this reveals a picture of a multi-layered response to plasmid acquisition. Understanding the flexibility of the bacterial metabolic network at the regulatory level to a new genetic element is thus shown to be important and worthy of study in other systems.

6.4 Chapter 4 | Strain specific responses to coevolution and antibiotic selection of diverse *E.coli* lineages with a multidrug resistant plasmid.

In Chapter 4 a subset of these strains were studied to understand the metabolic response to bacteria-plasmid coevolution. Evolved clones of 3 strains - F022 (clinical), ELU39 (environmental) and MG1655 (lab) - which had evolved for 700 generations either in the absence of the plasmid, carrying the plasmid or with the plasmid under cefotaxime selection (Carrilero, Dunn and Moran, 2023). This previous work observed that compensatory evolution involving genetic and transcriptomic changes targeting

metabolic pathways played a key role in amelioration of fitness costs. Untargeted metabolomics revealed that compared to their ancestors, all evolved clones displayed metabolomic alterations associated with adaptation to lab conditions. Although these were consistent among replicates, the magnitude and metabolites altered varied among strains, with the clinical strain demonstrating the most alterations to its metabolome. Amino acid biosynthesis, glycolysis, and pyrimidine biosynthesis were the functions most commonly altered following lab adaptation.

Bacteria-plasmid coevolution affected the metabolome in strain-specific ways, consistent with the findings of Chapter 3. Key affected functions included amino acid metabolism, nucleotide metabolism and ubiquinone biosynthesis. Moreover, addition of antibiotics during coevolution caused some metabolic changes, including amino acid metabolism, acidity protectants, glycolysis, and ubiquinone biosynthesis. Only the global regulator cyclic AMP (cAMP) was consistently affected across all strains, suggesting a role for regulated downstream functions in response to antibiotic selection.

6.5 Chapter 5 | Metabolic responses of 4 *E.coli* strains carrying a multidrug-resistant plasmid to antibiotic exposure.

In Chapter 5 four of the plasmid-carrying *E. coli* strains - F054, F022, ELU39 and MG1655 - were exposed to 3 different classes of antibiotics against which pLL35 encodes resistance genes: aminoglycosides, beta-lactams and fluoroquinolones represented by kanamycin, cefotaxime and ciprofloxacin. Untargeted metabolomics allowed a comprehensive assessment of the bacterial metabolome under sublethal antibiotic stress. The *E.coli* strains demonstrated a range of stress responses to antibiotic exposure and alterations to metabolites associated with bacterial resistance mechanisms. For the majority of strains, ciprofloxacin caused the most metabolic disruption, followed by cefotaxime and kanamycin, but the clinical strains deviated from these trends at times. There was also considerable functional overlap among strains and between treatments, but the extent and direction of change in particular pathways and metabolites were strain specific. Commonly affected pathways include amino acid and nucleotide metabolism, energy production pathways and cofactor biosynthesis. Interestingly, ubiquinone and ubiquinol were affected in the ciprofloxacin

treatment, suggesting bacterial management of redox stress. Together these data suggest that even after gaining resistance genes, bacteria still experience cell stress upon exposure to a range of antibiotics and use shifts in their metabolic network to mitigate these detrimental effects.

6.6 Common themes across chapters

One of the main themes in this thesis was that changes in the metabolome in response to plasmid acquisition and coevolution were strain specific. Fitness effects of plasmids on bacteria have been shown to vary by genotype, so a reflection of this in the metabolome is expected (Humphrey *et al.*, 2012). The variation in bacterial communities caused by plasmid carriage and the strain specific interactions of genes demonstrates further the properties of plasmids as facilitators of bacterial evolution, because this variation provides an evolutionary landscape for selection to act upon. This may be part of the reason the 'plasmid paradox' defies theory; the reasoning that due to the costs they incur, plasmids would be lost from bacterial populations through purifying selection.

A low impact on the host, such as is seen in the plasmid carriage data in this thesis, would prevent plasmid-carrying bacteria suffering from competitive exclusion (Fischer *et al.*, 2019), especially important for the clinically relevant, narrow host range ESBL plasmids. These plasmids sometimes induce no significant fitness cost at all in the native hosts (Palkovicova *et al.*, 2021).

In chapters 4 and 5 of this thesis the impacts of plasmid acquisition and coevolution have had relatively small effects on the metabolome ($\leq 2.6\%$). There may be several explanations for this.

The majority of absolute concentrations of metabolites in a bacterial metabolome at any given time is very low. In an analysis that measured the absolute amounts of 103 metabolites, the 10 most abundant metabolites make up 77% of the observed metabolome. The least abundant half makes up just 1.3% of the metabolome, glutamate, glutathione, fructose, ATP, UDP, hexose, UTP, GTP, dTTP, and aspartate. This includes most of the pathways highlighted in the results of this thesis, therefore

any of these as fold changes would be a major change in their normal proportion in the metabolome. Thus changes that appear small in terms of fold change may have a large biological impact (Bennett *et al.*, 2010).

Biosynthesis and catabolism pathways of amino acids were highlighted consistently throughout the chapters in this thesis. Amino acids comprise a large fraction of the metabolome, just under half, followed by nucleotides and central carbon (Bennett *et al.*, 2010). Amounts of nucleotides and amino acids are tightly regulated, even in response to a variety of conditions and stressors, they vary consistently less than 30%, which is 0.3 in fold change terms (Radoš *et al.*, 2022). *E. coli* metabolism can adjust to deletion of even major biosynthetic pathways to continue producing essential molecules (Cotton *et al.*, 2020). This suggests that differentially expressed amino acids may be indicative of a bigger metabolic response than previously thought.

In a community setting fitness variation in bacterial hosts may be advantageous for the plasmids, by creating a stable pool of resistance in some strains or species to persist and coevolve intensively with their bacterial hosts, one of the ecological solutions to the paradox (Clarke *et al.*, 2020; Brockhurst and Harrison, 2022; Newbury *et al.*, 2022). Given this stability in communities, limiting infection spread may be an increasing component of the clinical management strategy rather than a cessation of antibiotic use, because alone this will not promote re-susceptibility, thanks to genetic linkage of resistance genes and stable plasmids (Enne *et al.*, 2004; Peirano and Pitout, 2019; Newbury *et al.*, 2022).

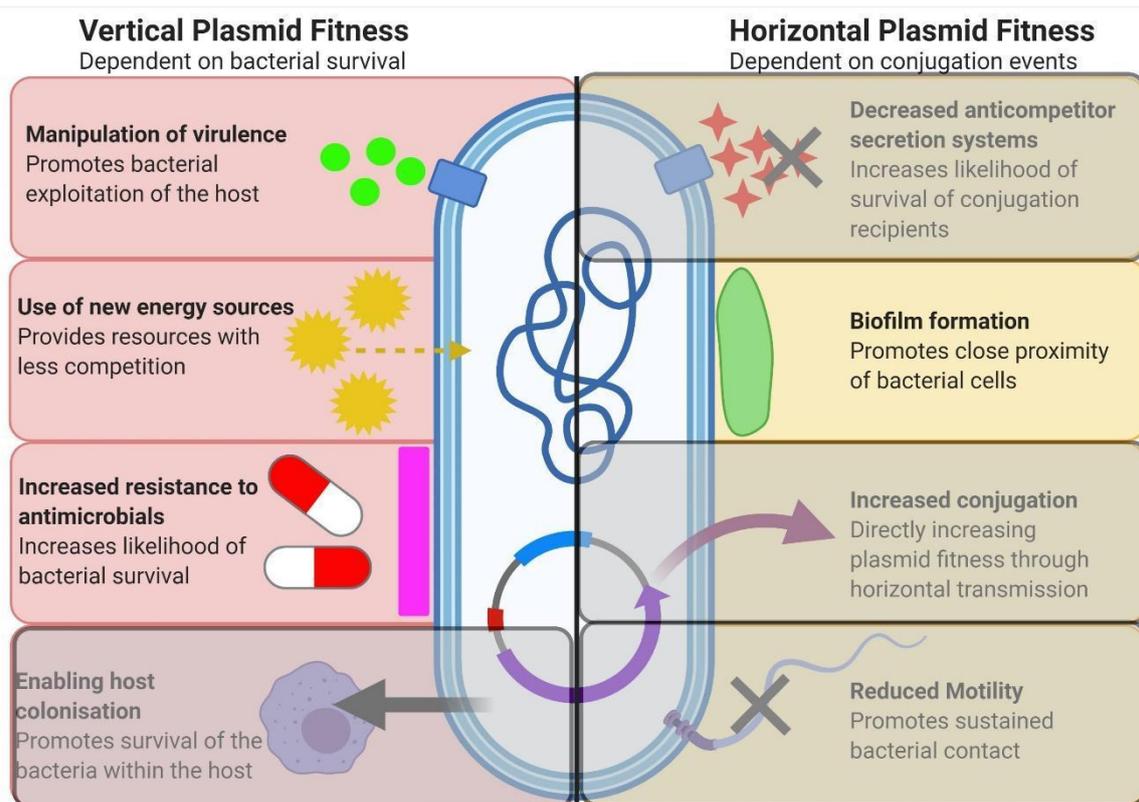


Figure 6.1 An edited version of the figure published in (Billane *et al.*, 2022), summarising common phenotypic effects of plasmids on their bacterial hosts, and how this may relate to different replication strategies of the plasmid. Greyed out are the functions not reflected in the metabolomics of this thesis.

This work extends previous work interrogating these strains which again, show subtle strain specific changes - but the targets differ (Dunn *et al.*, 2021; Carrilero, Dunn and Moran, 2023). This shows a multi-layered, complex pattern of bacterial responses. Plasmid mediated alterations to the metabolome without antibiotic selection fall in line with some of the phenotypic effects common to diverse plasmid and bacteria relationships (Covered Chapter 1, Billane *et al.*, 2022).

In the metabolome, plasmid carriage was associated with use of alternative energy sources, such as catabolism of amino acids, indicators of anaerobic metabolism and aromatic compounds (Chapter 3). Coevolution with a plasmid and exposure to antibiotics induced expression of bacterial resistance mechanisms in tandem with resistance genes harboured on the plasmid (Chapters 4 & 5). This also saw the

alteration of the global regulator cAMP which has multiple regulatory functions within the bacteria, including virulence factors and biofilm formation (McDonough and Rodriguez, 2012) (Chapter 4). An increased rate of aerobic respiration is indicated by ubiquinone and ubiquinol, which facilitate oxidative phosphorylation and are molecules highlighted in every chapter of this thesis, suggesting it's importance for *E.coli* acclimating to plasmid carriage and subsequent coevolution, and combatting antibiotic stress (Chapter 3, 4 & 5).

E.coli undergo metabolic changes within a human host from non-virulent high nutrient and sugar intestinal niche to the low-nutrient urinary tract, where it switches to catabolism of amino acids and peptides and increased virulence (Alteri and Mobley, 2012). Different amino acid utilisation and synthesis pathways are activated at different infection stages. Metabolites associated with virulence include pathogenicity carbon metabolism, gluconeogenesis, amino acid metabolism B-galactosidase, sorbitol. Purine and pyrimidine synthesis are also vital for host colonisation and growth (Mann *et al.*, 2017).

Together, the functional impacts of plasmid carriage complement the other phenotypic impacts observed in myriad pairings of plasmids and bacteria, observed through the lens of omic technologies (Figure 1).

6.7 Hypothesis Generation

The metabolome reveals different information than genomics and transcriptomics. This may be important to consider for future studies as historically transcriptomics are used to predict phenotype (Long *et al.*, 2019). However, metabolic disruption is not necessarily reflective of the transcriptomic disruption due to redundancy at different functional levels (Ishii *et al.*, 2007; Jozefczuk *et al.*, 2010). It is increasingly appreciated that combined multi-omic studies are the best way to obtain a comprehensive understanding of a biological system. Different levels reveal different information, and one cannot necessarily be used to predict another. Untargeted metabolomics is often used as a hypothesis generating tool (Di Minno *et al.*, 2021).

Based on the results of this thesis, further research should investigate the following:

Could plasmid impact include adaptive effects?

Future work should aim to untangle the effects of stress from either the plasmid or antibiotic and explore the possibility that plasmid carriage may induced a preparatory effect within the bacterial metabolome. The biosynthesis and use of the cellular regulator cAMP was featured repeatedly in the metabolome of *E.coli* that had coevolved with a plasmid with and without antibiotic selection and in the metabolomes of the ancestral plasmid carriers upon exposure to antibiotics. cAMP levels were altered in all strains studied in Chapter 4 and indicated that plasmid carriage induced downregulation, when compared to plasmid free strains, and cefotaxime exposure induced upregulation, compared to plasmid free strains. Chapter 5 showed that cAMP was important in the bacterial response to antibiotic stress. The plasmid has an impact on cAMP that is opposite to the cAMP response when the bacteria encounters antibiotics. Further assessment of this pathway should test if this represents an adaptive effect, where the plasmid alters some aspects of the bacterial metabolome to conserve or accumulate metabolic products required for the survival of the bacteria during exposure to stressors.

A number of pathways highlighted in the untargeted metabolomics throughout this thesis are candidates for further, targeted analysis.

Examining the dynamics of these pathways may further our understanding of plasmid persistence and the continued evolution of resistance. In response to plasmid acquisition, co-evolution and antibiotic exposure, significant changes are consistently seen throughout this thesis to ubiquinone and ubiquinol, and metabolism of amino acids such as lysine, arginine and tryptophan. Targeted analysis to examine the full pathway would determine whether these are potential targets for metabolite based combination therapies (Meylan, Andrews and Collins, 2018; Crabbé *et al.*, 2019).

Additionally, the upregulation of NAM, part of peptidoglycan biosynthesis in response to ciprofloxacin was the largest metabolic alternation to a single metabolite in multiple strains in the data. While untargeted metabolomics facilitated this discovery, understanding of this remains incomplete. The gain of resistance does not mean that the bacteria cease to respond to selection. Instead the compounding tolerance by plasmid acquisition and chromosomal resistance mutations each facilitate the other (Coluzzi *et al.*, 2023). This contributes to the evolution of multidrug resistance and so understanding key metabolic pathways that change during bacterial-plasmid coevolution, or plasmid carrier responses to survive antibiotics, are important to aid the mitigation of the dissemination of multi-drug resistance.

6.8 Conclusions

This thesis demonstrates, by use of untargeted metabolomics, that plasmid carriage induced strain specific responses in *E. coli* which have broad functional commonalities. Adjusting the metabolism of amino acids, production of energy including utilisation of alternative sources or a switch to anaerobic metabolism. These functions continued to be affected by evolution, which impacted amino acid, energy and nucleotide metabolism as the strains adapted to laboratory conditions. Coevolution with a plasmid had a varied scale of impact among the strains when compared to plasmid free controls and affected amino acid metabolism, nucleotide metabolism and energy production in the clinical strains, while the environmental strain altered fatty acid metabolism and acid protectants to adapt to plasmid carriage. Antibiotic selection induced a range of stress responses, highlighted by the altered regulation of energy production, amino acid metabolism and nucleotide metabolism. Bacterial resistance mechanisms are also induced, highlighted by altered regulation of cell wall and biofilm component biosynthesis and redox stress regulation, which has roles in both stress and resistance responses of the bacteria. This shows the bacteria continue to respond to stressors in tandem with plasmid-encoded resistance genes. Thus *Escherichia coli* adapt to plasmid carriage and antibiotic exposure by subtle alterations of its metabolism on a network wide scale.

References

Adolfson, K.J. and Brynildsen, M.P. (2015) 'Futile cycling increases sensitivity toward oxidative stress in *Escherichia coli*', *Metabolic Engineering*, 29, pp. 26–35. doi:10.1016/j.ymben.2015.02.006.

Agrawal, S. *et al.* (2017) 'A genome-wide screen in *Escherichia coli* reveals that ubiquinone is a key antioxidant for metabolism of long-chain', *Journal of Biological Chemistry*, 292(49), pp. 20086–20099. doi:10.1074/jbc.M117.806240.

Aldridge, B.B. and Rhee, K.Y. (2014) 'Microbial metabolomics: Innovation, application, insight', *Current Opinion in Microbiology*. Elsevier Ltd, pp. 90–96. doi:10.1016/j.mib.2014.06.009.

Alhashash, F. *et al.* (2013) 'Multidrug-resistant *Escherichia coli* bacteremia', *Emerging Infectious Diseases*, 19(10), pp. 1699–1701. doi:10.3201/eid1910.130309.

Allwood, J.W. *et al.* (2021) 'Unravelling plant responses to stress—the importance of targeted and untargeted metabolomics', *Metabolites*, 11(8), pp. 1–24. doi:10.3390/metabo11080558.

Alonso-del Valle, A. *et al.* (2021) 'Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities', *Nature Communications*, 12(1), pp. 1–14. doi:10.1038/s41467-021-22849-y.

Alteri, C.J. and Mobley, H.L.T. (2012) '*Escherichia coli* physiology and metabolism dictates adaptation to diverse host microenvironments', *Current Opinion in Microbiology*, 15(1), pp. 3–9. doi:10.1016/j.mib.2011.12.004.

Antimicrobial Resistance (2021) *World Health Organization*. Available at: <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance> (Accessed: 23 February 2023).

Arense, P. *et al.* (2010) 'Metabolic adaptation of *Escherichia coli* to long-term exposure to salt stress', *Process Biochemistry*, 45(9), pp. 1459–1467. doi:10.1016/j.procbio.2010.05.022.

Ares-arroyo, M. *et al.* (2022) 'Genomics, Transcriptomics, and Metabolomics Reveal That Minimal Modifications in the Host Are Crucial for the Compensatory Evolution of ColE1-Like Plasmids', *mSphere*, 7(6), p. 1. Doi: 10.1128/msphere.00184-22.

Arias-cartin, R. *et al.* (2023) 'Role of the *Escherichia coli* ubiquinone-synthesizing UbiUVT pathway in adaptation to changing respiratory conditions', *mBio*, 14(4), pp. 1–24. doi: 10.1128/mbio.03298-22

Arnér, E.S.J. and Holmgren, A. (2000) 'Physiological functions of thioredoxin and thioredoxin reductase', *European Journal of Biochemistry*, 267(20), pp. 6102–6109. doi:10.1046/j.1432-1327.2000.01701.x.

Aussel, L. *et al.* (2014) 'Biosynthesis and physiology of coenzyme Q in bacteria', *Biochimica et Biophysica Acta*, pp. 1004–1011. doi:10.1016/j.bbabi.2014.01.015

Von Bargen, K. *et al.* (2009) 'Rhodococcus equi virulence-associated protein A is required for diversion of phagosome biogenesis but not for cytotoxicity', *Infection and Immunity*, 77(12), pp. 5676–5681. doi:10.1128/IAI.00856-09.

Bennett, B.D. *et al.* (2010) 'Absolute Metabolite Concentrations and Implied Enzyme Active Site Occupancy in *Escherichia coli*', *Nature chemical biology*, 5(8), pp. 593–599. doi:10.1038/nchembio.186.Absolute.

Benz, F. *et al.* (2021) 'Plasmid- and strain-specific factors drive variation in ESBL-plasmid spread in vitro and in vivo', *ISME Journal*, 15(3), pp. 862–878. doi:10.1038/s41396-020-00819-4.

Berg, J.M. *et al.* (2019) *Biochemistry*. 9th edition, Macmillan, New York.

Bevan, E.R., Jones, A.M. and Hawkey, P.M. (2017) 'Global epidemiology of CTX-M β -lactamases: Temporal and geographical shifts in genotype', *Journal of Antimicrobial Chemotherapy*, 72(8), pp. 2145–2155. doi:10.1093/jac/dkx146.

Billane, K. *et al.* (2022) 'Why do plasmids manipulate the expression of bacterial phenotypes?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1842), pp. 1–9. doi:10.1098/rstb.2020.0461.

Blaženović, I. *et al.* (2018) 'Software tools and approaches for compound identification of LC-MS/MS data in metabolomics', *Metabolites*. MDPI AG, pp. 1–23. doi:10.3390/metabo8020031.

Bottery, M.J., Wood, A.J. and Brockhurst, M.A. (2017) 'Adaptive modulation of antibiotic resistance through intragenomic coevolution', *Nature Ecology and Evolution*, 1(9), pp. 1364–1369. doi:10.1038/s41559-017-0242-3.

Bottery, M.J., Wood, A.J. and Brockhurst, M.A. (2019) 'Temporal dynamics of bacteria-plasmid coevolution under antibiotic selection', *The ISME Journal*, 13, pp. 559–562. doi:10.1038/s41396-018-0276-9.

Brockhurst, M.A. and Harrison, E. (2022) 'Ecological and evolutionary solutions to the plasmid paradox', *Trends in Microbiology*, 30(6), pp. 534–543. doi.org/10.1016/j.tim.2021.11.001

Brynildsen, M.P. *et al.* (2013) 'Potentiating antibacterial activity by predictably enhancing endogenous microbial ROS production', *Nature Biotechnology*, 31(2), pp. 160–165. doi:10.1038/nbt.2458.

Buckner, M.M.C., Ciusa, M.L. and Piddock, L.J.V. (2018) 'Strategies to combat antimicrobial resistance: Anti-plasmid and plasmid curing', *FEMS Microbiology Reviews*, 42(6), pp. 781–804. doi:10.1093/femsre/fuy031.

Canton, R. and Coque, T.M. (2006) 'The CTX-M β -lactamase pandemic', *Current Opinion in Microbiology*, 9, pp. 466–475. doi:10.1016/j.mib.2006.08.011.

Cantón, R., González-Alba, J.M. and Galán, J.C. (2012) 'CTX-M enzymes: Origin and diffusion', *Frontiers in Microbiology*, 3(APR), pp. 1–19. doi:10.3389/fmicb.2012.00110.

Carattoli, A. (2013) 'Plasmids and the spread of resistance', *International Journal of Medical Microbiology*, 303(6–7), pp. 298–304. doi:10.1016/j.ijmm.2013.02.001.

Carrilero, L., Dunn, S.J. and Moran, R.A. (2023) 'Evolutionary Responses to Acquiring a Multidrug Resistance Plasmid Are Dominated by Metabolic Functions across Diverse *Escherichia coli* Lineages', *mSystems*, 8(1), pp. 1–15. doi: 10.1128/msystems.00713-22 1.

Carroll, A.C. and Wong, A. (2018) 'Plasmid persistence: costs, benefits, and the plasmid paradox', *Canadian Journal of Microbiology*, 64(5), pp. 293–304. doi:10.1139/cjm-2017-0609.

Cayley, S., Lewis, B.A. and Record, M.T. (1992) 'Origins of the osmoprotective properties of betaine and proline in *Escherichia coli* K-12', *Journal of Bacteriology*, 174(5), pp. 1586–1595. doi:10.1128/jb.174.5.1586-1595.1992.

Chaleckis, R. *et al.* (2019) 'Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics', *Current Opinion in Biotechnology*. Elsevier Ltd, pp. 44–50. doi:10.1016/j.copbio.2018.07.010.

Che, Y. *et al.* (2021) 'Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes', *PNAS*, 118(6), pp. 1–12. doi:10.1073/pnas.2008731118/-/DCSupplemental.

Cho, H., Uehara, T. and Bernhardt, T.G. (2014) 'Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery', *Cell*, 159(6), pp. 1300–1311. doi:10.1016/j.cell.2014.11.017.

Clarke, L. *et al.* (2020) 'The effect of environmental heterogeneity on the fitness of antibiotic resistance mutations in *Escherichia coli*', *Evolutionary Ecology*, 34(3), pp. 379–390. doi:10.1007/s10682-019-10027-y.

Coluzzi, C. *et al.* (2023) 'Chance Favors the Prepared Genomes: Horizontal Transfer Shapes the Emergence of Antibiotic Resistance Mutations in Core Genes', *Molecular Biology and Evolution*, 40(10), pp. 1–20. doi:10.1093/molbev/msad217.

Cotton, C.A.R. *et al.* (2020) 'Underground isoleucine biosynthesis pathways in e. Coli', *eLife*, 9, pp. 1–25. doi:10.7554/ELIFE.54207.

Coulson, G.B. *et al.* (2015) 'Transcriptome reprogramming by plasmid-encoded transcriptional regulators is required for host niche adaption of a macrophage pathogen', *Infection and Immunity*, 83(8), pp. 3137–3145. doi:10.1128/IAI.00230-15.

Crabbé, A. *et al.* (2019) 'Antimicrobial Tolerance and Metabolic Adaptations in Microbial Biofilms', *Trends in Microbiology*, 27(10), pp. 850–863.

doi:10.1016/j.tim.2019.05.003.

da Cunha, B.R., Fonseca, L.P. and Calado, C.R.C. (2019) 'Antibiotic discovery: Where have we come from, where do we go?', *Antibiotics*, 8(2). doi:10.3390/antibiotics8020045.

Deter, H.S., Hossain, T. and Butzin, N.C. (2021) 'Antibiotic tolerance is associated with a broad and complex transcriptional response in *E. coli*', *Scientific Reports*, 11(1), pp. 1–15. doi:10.1038/s41598-021-85509-7.

Dimroth, P. (1987) 'Sodium Ion Transport Decarboxylases and Other Aspects of Sodium Ion Cycling in Bacteria', *Microbiological Reviews*, 51(3), pp. 320–340.

Dionisio, F., Zilhão, R. and Gama, J.A. (2019) 'Interactions between plasmids and other mobile genetic elements affect their transmission and persistence', *Plasmid*, 102(February), pp. 29–36. doi:10.1016/j.plasmid.2019.01.003.

Dolan, S.K. *et al.* (2020) 'Contextual flexibility in pseudomonas aeruginosa central carbon metabolism during growth in single carbon sources', *mBio*, 11(2). doi:10.1128/mBio.02684-19.

Dudzik, D. *et al.* (2018) 'Quality assurance procedures for mass spectrometry untargeted metabolomics. a review', *Journal of Pharmaceutical and Biomedical Analysis*. Elsevier B.V., pp. 149–173. doi:10.1016/j.jpba.2017.07.044.

Dunn, S. *et al.* (2021) 'Limited and Strain-Specific Transcriptional and Growth Responses to Acquisition of a Multidrug Resistance Plasmid in Genetically Diverse *Escherichia coli* Lineages', *mSystems*, 6(2), pp. 1–14. doi:10.1128/mSystems.00083-21

Dunn, S.J., Connor, C. and McNally, A. (2019) 'The evolution and transmission of multi-drug resistant *Escherichia coli* and *Klebsiella pneumoniae*: the complexity of clones and plasmids', *Current Opinion in Microbiology*, 51, pp. 51–56. doi:10.1016/j.mib.2019.06.004.

Dwyer, D.J. *et al.* (2014) 'Antibiotics induce redox-related physiological alterations as part of their lethality', *Proceedings of the National Academy of Sciences of the United*

States of America, 111(20), pp. E2100–E2109. doi:10.1073/pnas.1401876111.

Enne, V.I. *et al.* (2004) 'Enhancement of host fitness by the *sul2*-coding plasmid p9123 in the absence of selective pressure', *Journal of Antimicrobial Chemotherapy*, 53(6), pp. 958–963. doi:10.1093/jac/dkh217.

Fischer, E.A.J. *et al.* (2019) 'Competition between *Escherichia coli* Populations with and without Plasmids Carrying a Gene Encoding Extended-Spectrum Beta-Lactamase in the Broiler Chicken Gut ', *Applied and Environmental Microbiology*, 85(17), pp. 1–12. doi:10.1128/aem.00892-19.

Gallant, C. V. *et al.* (2005) 'Common β -lactamases inhibit bacterial biofilm formation', *Molecular Microbiology*, 58(4), pp. 1012–1024. doi:10.1111/j.1365-2958.2005.04892.x.

Garneau-Tsodikovaa, S. and Labby, K.J. (2016) 'Mechanisms of resistance to aminoglycoside antibiotics: overview and perspectives', *Med. Chem. Commun.*, 7, pp. 11–27. doi:10.1039/C5MD00344J

De Gelder, L. *et al.* (2007) 'Stability of a promiscuous plasmid in different hosts: No guarantee for a long-term relationship', *Microbiology*, 153(2), pp. 452–463. doi:10.1099/mic.0.2006/001784-0.

Gomez-Duarte, O.G. and Kaper, J.B. (1995) 'A plasmid-encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*', *Infection and Immunity*, 63(5), pp. 1767–1776. doi:10.1128/iai.63.5.1767-1776.1995.

Goodacre, R. *et al.* (2004) 'Metabolomics by numbers: Acquiring and understanding global metabolite data', *Trends in Biotechnology*, pp. 245–252. doi:10.1016/j.tibtech.2004.03.007.

Guo AC, J. *et al.* (2012) *ECMDB: The E. coli Metabolome Database.*, *S. Nucleic Acids Res.* Available at: <https://ecmdb.ca/>.

Gutierrez, A. *et al.* (2017) 'Understanding and Sensitizing Density-Dependent Persistence to Quinolone Antibiotics', *Molecular Cell*, 68(6), pp. 1147-1154.e3. doi:10.1016/j.molcel.2017.11.012.

Hall, J.P.J. *et al.* (2016) 'Source-sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities', *Proceedings of the National Academy of Sciences*, 113(29), pp. 8260–8265. doi:10.1016/j.physa.2005.09.069.

Hall, J.P.J. *et al.* (2021) *Plasmid fitness costs are caused by specific genetic conflicts enabling resolution by compensatory mutation*, *PLoS Biology*. doi:10.1371/journal.pbio.3001225.

Hall, J.P.J., Brockhurst, M.A. and Harrison, E. (2017) 'Sampling the mobile gene pool: Innovation via horizontal gene transfer in bacteria', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1735), pp. 1–10. doi:10.1098/rstb.2016.0424.

Harrison, E. *et al.* (2015) 'Parallel Compensatory Evolution Stabilizes Plasmids across the Parasitism-Mutualism Continuum', *Current Biology*, 25(15), pp. 2034–2039. doi:10.1016/j.cub.2015.06.024.

Harrison, E. and Brockhurst, M.A. (2012) 'Plasmid-mediated horizontal gene transfer is a coevolutionary process', *Trends in Microbiology*, 20(6), pp. 262–267. doi:10.1016/j.tim.2012.04.003.

Hasik, A.Z. and Siepielski, A.M. (2022) 'Parasitism shapes selection by drastically reducing host fitness and increasing host fitness variation', *Biology letters*, 18(11), pp. 2–6. doi:10.1098/rsbl.2022.0323.

Hawkey, P.M. and Jones, A.M. (2009) 'The changing epidemiology of resistance', *Journal of Antimicrobial Chemotherapy*, 64(SUPPL.1), pp. i3–i10. doi:10.1093/jac/dkp256.

Hernando-amado, S. *et al.* (2017) 'Fitness costs associated with the acquisition of antibiotic resistance', *Essays in Biochemistry*, 61(March), pp. 37–48. doi:10.1042/EBC20160057.

Hoerr, V. *et al.* (2016) 'Characterization and prediction of the mechanism of action of antibiotics through NMR metabolomics', *BMC Microbiology*, 16(1), pp. 1–14. doi:10.1186/s12866-016-0696-5.

Huang, C. *et al.* (2020) 'A novel incompatibility group X3 plasmid carrying bla_{NDM-1} encodes a small RNA that regulates host fucose metabolism and biofilm formation', *RNA Biology*, 17(12), pp. 1767–1776. doi:10.1080/15476286.2020.1780040.

Humphrey, B. *et al.* (2012) 'Fitness of *Escherichia coli* strains carrying expressed and partially silent IncN and IncP1 plasmids', *BMC Microbiology*, 12. doi:10.1186/1471-2180-12-53.

Ishii, N. *et al.* (2007) 'Multiple High-Throughput Analyses Monitor the Response of *E. coli* to Perturbations', *Science*, 316(April), pp. 593–596. doi:10.1126/science.1136031.

Jiang, X. *et al.* (2017) 'The CTX-M-14 plasmid pHK01 encodes novel small RNAs and influences host growth and motility', *FEMS microbiology ecology*, 93(7), pp. 1–10. doi:10.1093/femsec/fix090.

Jindrova, E. *et al.* (2002) 'Bacterial Aerobic Degradation of Benzene, Toluene, Ethylbenzene and Xylene', *Folia Microbiol.* Available at: doi:10.1007/BF02817664

Johnson, T.J. *et al.* (2015) 'In Vivo transmission of an incA/C plasmid in *Escherichia coli* depends on tetracycline concentration, and acquisition of the plasmid results in a variable cost of fitness', *Applied and Environmental Microbiology*, 81(10), pp. 3561–3570. doi:10.1128/AEM.04193-14.

Jolivet-Gougeon, A. and Bonnaure-Mallet, M. (2014) 'Biofilms as a mechanism of bacterial resistance', *Drug Discovery Today: Technologies*, 11(1), pp. 49–56. doi:10.1016/j.ddtec.2014.02.003.

Jozefczuk, S. *et al.* (2010) 'Metabolomic and transcriptomic stress response of *Escherichia coli*', *Molecular Systems Biology*, 6(364), pp. 1–16. doi:10.1038/msb.2010.18.

KEGG Kanehisa Laboratories. Available at: <https://www.genome.jp/kegg/>.

Kok, M. *et al.* (2022) 'Unraveling antimicrobial resistance using metabolomics', *Drug Discovery Today*, 27(6), pp. 1774–1783. doi:10.1016/j.drudis.2022.03.015.

Kotra, L.P., Haddad, J. and Mobashery, S. (2000) 'Aminoglycosides: Perspectives on

mechanisms of action and resistance and strategies to counter resistance', *Antimicrobial Agents and Chemotherapy*, 44(12), pp. 3249–3256. doi:10.1128/AAC.44.12.3249-3256.2000.

Krell, T. and Matilla, M.A. (2022) 'Antimicrobial resistance: progress and challenges in antibiotic discovery and anti-infective therapy', *Microbial Biotechnology*, 15(1), pp. 70–78. doi:10.1111/1751-7915.13945.

Lahlaoui, H., Ben Haj Khalifa, A. and Ben Moussa, M. (2014) 'Epidemiology of Enterobacteriaceae producing CTX-M type extended spectrum β -lactamase (ESBL)', *Medecine et Maladies Infectieuses*, 44(9), pp. 400–404. doi:10.1016/j.medmal.2014.03.010.

Lang, K.S. and Johnson, T.J. (2015) 'Transcriptome modulations due to A/C2 plasmid acquisition', *Plasmid*, 80, pp. 83–89. doi:10.1016/j.plasmid.2015.05.005.

Laviña, W.A. *et al.* (2020) 'Metabolomics analysis reveals global metabolic changes in the evolved E. Coli strain with improved growth and 1-butanol production in minimal medium', *Metabolites*, 10(192), pp. 1–13. doi:10.3390/metabo10050192.

Lee Ventola, C. (2015) 'The Antibiotic Resistance Crisis', *Comprehensive Biochemistry*, 40(4), pp. 277–283. doi:10.1016/B978-1-4831-9711-1.50022-3.

Lopatkin, A.J. *et al.* (2017) 'Persistence and reversal of plasmid-mediated antibiotic resistance', *Nature Communications*, 8(1689), pp. 1–10. doi:10.1038/s41467-017-01532-1.

López-Sámamo, M. *et al.* (2020) 'A novel way to synthesize pantothenate in bacteria involves β -alanine synthase present in uracil degradation pathway', *MicrobiologyOpen*, 9(4), pp. 1–25. doi:10.1002/mbo3.1006.

Long, D. *et al.* (2019) 'Phenotypical profile and global transcriptomic profile of Hypervirulent *Klebsiella pneumoniae* due to carbapenemase-encoding plasmid acquisition', *BMC Genomics*, 20(480), pp. 1–13. doi:10.1186/s12864-019-5705-2.

MacLean, R.C. and San Millan, A. (2015) 'Microbial Evolution: Towards Resolving the Plasmid Paradox', *Current Biology*, 25(17), pp. R764–R767.

doi:10.1016/j.cub.2015.07.006.

MacLean, R.C. and San Millan, A. (2019) 'The Evolution of Antibiotic Resistance', *SCIENCE Microbiology*, 365(6458), pp. 1082–1083. doi:10.1016/B978-0-12-799942-5.00012-3.

Mann, R. *et al.* (2017) 'Metabolic adaptations of Uropathogenic *E. coli* in the urinary tract', *Frontiers in Cellular and Infection Microbiology*, 7(JUN). doi:10.3389/fcimb.2017.00241.

Mardegan, V. *et al.* (2021) 'Untargeted and targeted metabolomic profiling of preterm newborns with early onset sepsis: A case-control study', *Metabolites*, 11(115), pp. 1–15. doi:10.3390/metabo11020115.

Maser, A. *et al.* (2020) 'Amino acids are key substrates to *Escherichia coli* BW25113 for achieving high specific growth rate', *Research in Microbiology*, 171, pp. 185–193. doi:10.1016/j.resmic.2020.02.001

Mathers, A.J., Peirano, G. and Pitout, J.D.D. (2015) 'The role of epidemic resistance plasmids and international high- risk clones in the spread of multidrug-resistant Enterobacteriaceae', *Clinical Microbiology Reviews*, 28(3), pp. 565–591. doi:10.1128/CMR.00116-14.

Matsumoto, A. *et al.* (1998) 'Plaque formation by and plaque cloning of *Chlamydia trachomatis* biovar *trachoma*', *Journal of Clinical Microbiology*, 36(10), pp. 3013–3019. doi:10.1128/jcm.36.10.3013-3019.1998.

McDonough, K.A. and Rodriguez, A. (2012) 'The myriad roles of cyclic AMP in microbial pathogens: From signal to sword', *Nature Reviews Microbiology*, 10(1), pp. 27–38. doi:10.1038/nrmicro2688.

Metris, A. *et al.* (2014) 'Metabolic shift of *Escherichia coli* under salt stress in the presence of glycine betaine', *Applied and Environmental Microbiology*, 80(15), pp. 4745–4756. doi:10.1128/AEM.00599-14.

Meylan, S. *et al.* (2017) 'Carbon Sources Tune Antibiotic Susceptibility in *Pseudomonas aeruginosa* via Tricarboxylic Acid Cycle Control', *Cell Chemical*

Biology, 24(2), pp. 195–206. doi:10.1016/j.chembiol.2016.12.015.

Meylan, S., Andrews, I.W. and Collins, J.J. (2018) 'Targeting Antibiotic Tolerance, Pathogen by Pathogen', *Cell*, 172(6), pp. 1228–1238. doi:10.1016/j.cell.2018.01.037.

Di Minno, A. *et al.* (2021) 'The evolving landscape of untargeted metabolomics', *Nutrition, Metabolism and Cardiovascular Diseases*, pp. 1645–1652. doi:10.1016/j.numecd.2021.01.008.

Misra, B.B. (2018) 'New tools and resources in metabolomics: 2016–2017', *Electrophoresis*. pp. 909–923. doi:10.1002/elps.201700441.

Newbury, A. *et al.* (2022) 'Fitness effects of plasmids shape the structure of bacteria–plasmid interaction networks', *Proceedings of the National Academy of Sciences of the United States of America*, 119(22), pp. 1–8. doi:10.1073/pnas.2118361119.

Nitzschke, A. and Bettenbrock, K. (2018) 'All three quinone species play distinct roles in ensuring optimal growth under aerobic and fermentative conditions in *E. coli* K12', *Plos ONE*, pp. 1–18. Available at: <https://doi.org/10.1371/journal.pone.0194699>.

Norman, A., Hansen, L.H. and Sørensen, S.J. (2009) 'Conjugative plasmids: Vessels of the communal gene pool', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1527), pp. 2275–2289. doi:10.1098/rstb.2009.0037.

O'Loughlin, J. *et al.* (2021) 'The Antibacterial Drug Candidate SBC3 is a Potent Inhibitor of Bacterial Thioredoxin Reductase', *ChemBioChem*, 22(6), pp. 1093–1098. doi:10.1002/cbic.202000707.

Ojkic, N. *et al.* (2020) 'A Roadblock-and-Kill Mechanism of Action Model for the DNA-Targeting Antibiotic Ciprofloxacin', *Antimicrobial Agents and Chemotherapy*, 64(9), pp. 1–17. doi:10.1128/AAC.02487-19.

Oliver, S.G. *et al.* (1998) 'Systematic functional analysis of the yeast genome', *Tibtech*, 16(September), pp. 373–378.

Pacheco, J.O. *et al.* (2017) 'Metabolic compensation of fitness costs is a general outcome for antibiotic-resistant *Pseudomonas aeruginosa* mutants overexpressing

efflux pumps', *mBio*, 8(4). doi:10.1128/mBio.00500-17.

Padda, I.S. and Nagalli, S. (2022) *Cefotaxime*. StatPearls Publishing, Treasure Island (FL). Available at: <http://europepmc.org/books/NBK560653>.

Palkovicova, Jana. *et al.* Fitness effects of *bla*_{CTX-M-15}-harbouring F2:A1:B- plasmids on their native *Escherichia coli* ST131 H30Rx hosts, *Journal of Antimicrobial Chemotherapy*, Volume 77, Issue 11, November 2022, Pages 2960–2963, doi : 10.1093/jac/dkac250

Parashar, V. *et al.* (2013) 'A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence', *Journal of Bacteriology*, 195(10), pp. 2437–2448. doi:10.1128/JB.02030-12.

Parker, E.J. *et al.* (2023) 'Untangling the complexities of processing and analysis for untargeted LC-MS data using open-source tools', *Metabolites*, 13(4), pp. 436–448. doi:10.3390/metabo13040463.

Parsons, J.B. and Rock, C.O. (2013) 'Bacterial lipids: Metabolism and membrane homeostasis', *Progress in Lipid Research*, 52(3), pp. 249–276. doi:10.1016/j.plipres.2013.02.002.

Passalacqua, K.D., Charbonneau, M.E. and O'Riordan, M.X.D. (2016) 'Bacterial metabolism shapes the host-pathogen interface', *Virulence Mechanisms of Bacterial Pathogens*, pp. 15–41. doi:10.1128/9781555819286.ch2.

Patton, M.J. *et al.* (2018) 'Plasmid Negative Regulation of CPAF Expression Is Pgp4 Independent and Restricted to Invasive *Chlamydia trachomatis* Biovars', *mBio*, 9(1), pp. 1–14. doi: 10.1128/mBio.02164-17

Peirano, G. and Pitout, J.D.D. (2019) 'Extended-Spectrum β -Lactamase-Producing Enterobacteriaceae: Update on Molecular Epidemiology and Treatment Options', *Drugs*, 79(14), pp. 1529–1541. doi:10.1007/s40265-019-01180-3.

Peña-Miller, R. *et al.* (2015) 'Evaluating the effect of horizontal transmission on the stability of plasmids under different selection regimes', *Mobile Genetic Elements*, 5(3), pp. 29–33. doi:10.1080/2159256X.2015.1045115.

Peng, B. *et al.* (2015) 'Exogenous Alanine and/or Glucose plus Kanamycin Kills Antibiotic-Resistant Bacteria', *Cell Metabolism*, 21(2), pp. 249–262. doi:10.1016/j.cmet.2015.01.008.

Plumbridge, J. (2015) 'Regulation of the Utilization of Amino Sugars by *Escherichia coli* and *Bacillus subtilis*: Same Genes, Different Control', *Journal of Molecular Microbiology and Biotechnology*, 25(2–3), pp. 154–167. doi: 10.1159/000369583

Poole, K. (2004) 'Resistance to β -lactam antibiotics', *Cellular and Molecular Life Sciences*, 61(17), pp. 2200–2223. doi:10.1007/s00018-004-4060-9.

Porse, A. *et al.* (2016a) 'Survival and Evolution of a Large Multidrug Resistance Plasmid in New Clinical Bacterial Hosts', *Molecular Biology and Evolution*, 33(11), pp. 2860–2873. doi:10.1093/molbev/msw163.

Prensky, H. *et al.* (2021) 'Conjugation dynamics depend on both the plasmid acquisition cost and the fitness cost', *Molecular Systems Biology*, 17(3), pp. 1–15. doi:10.15252/msb.20209913.

Radoš, D. *et al.* (2022) 'Homeostasis of the biosynthetic *E. coli* metabolome', *iScience*, 25(7). doi:10.1016/j.isci.2022.104503.

Ragheb, M.N. *et al.* (2019) 'Inhibiting the Evolution of Antibiotic Resistance', *Molecular Cell*, 73(1), pp. 157–165. doi:10.1016/j.molcel.2018.10.015.

Ranjan, A. *et al.* (2018) 'ESBL-plasmid carriage in *E. coli* enhances in vitro bacterial competition fitness and serum resistance in some strains of pandemic sequence types without overall fitness cost', *Gut Pathogens*, 10(1), pp. 1–9. doi:10.1186/s13099-018-0243-z.

Rodríguez-Beltrán, J. *et al.* (2021) 'Beyond horizontal gene transfer: the role of plasmids in bacterial evolution', *Nature Reviews Microbiology*. Nature Research, pp. 347–359. doi:10.1038/s41579-020-00497-1.

Rodríguez-Beltrán, J. *et al.* (2022) 'Translational demand is not a major source of plasmid-associated fitness costs', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1842). doi:10.1098/rstb.2020.0463.

Ronin, I. *et al.* (2017) 'A long-term epigenetic memory switch controls bacterial virulence bimodality', *eLife*, 6(19599), pp. 1–26. doi:10.7554/eLife.19599.

Rui, B. *et al.* (2010) 'A systematic investigation of *Escherichia coli* central carbon metabolism in response to superoxide stress.', *BMC systems biology*, 4, p. 122. doi:10.1186/1752-0509-4-122.

Ryan, D. and Robards, K. (2006) 'Metabolomics: The greatest omics of them all?', *Analytical Chemistry*, pp. 7954–7958. doi:10.1021/ac0614341.

San Millan, A. *et al.* (2014) 'Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids', *Nature Communications*, 5(5208), pp. 1–11. doi:10.1038/ncomms6208.

San Millan, A. (2018) 'Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context', *Trends in Microbiology*, 26(12), pp. 978–985. doi:10.1016/j.tim.2018.06.007.

San Millan, A. *et al.* (2018) 'Integrative analysis of fitness and metabolic effects of plasmids in *Pseudomonas aeruginosa* PAO1', *ISME Journal*, 12(12), pp. 3014–3024. doi:10.1038/s41396-018-0224-8.

San Millan, A., Heilbron, K. and MacLean, R.C. (2014) 'Positive epistasis between co-infecting plasmids promotes plasmid survival in bacterial populations', *ISME Journal*, 8(3), pp. 601–612. doi:10.1038/ismej.2013.182.

San Millan, A. and MacLean, R.C. (2017) 'Fitness Costs of Plasmids: a Limit to Plasmid Transmission', *Microbiology Spectrum*, 5(5). doi:10.1128/microbiolspec.mtbp-0016-2017.

Santos-Lopez, A. *et al.* (2017) 'A naturally occurring single nucleotide polymorphism in a multicopy plasmid produces a reversible increase in antibiotic resistance', *Antimicrobial Agents and Chemotherapy*, 61(2), pp. 1–7. doi:10.1128/AAC.01735-16.

Sarasa, S.B. *et al.* (2020) 'A Brief Review on the Non-protein Amino Acid, Gamma-amino Butyric Acid (GABA): Its Production and Role in Microbes', *Current Microbiology*, 77(4), pp. 534–544. doi:10.1007/s00284-019-01839-w.

Schaufler, K. *et al.* (2016) 'Carriage of extended-spectrum beta-lactamase-plasmids does not reduce fitness but enhances virulence in some strains of pandemic *E. coli* lineages', *Frontiers in Microbiology*, 7(MAR), pp. 1–12. doi:10.3389/fmicb.2016.00336.

Schrimpe-Rutledge, A.C. *et al.* (2016) 'Untargeted Metabolomics Strategies—Challenges and Emerging Directions', *Journal of the American Society for Mass Spectrometry*, 27(12), pp. 1897–1905. doi:10.1007/s13361-016-1469-y.

Sezmis, A.L. *et al.* (2023) 'Horizontal Gene Transfer, Fitness Costs and Mobility Shape the Spread of Antibiotic Resistance Genes into Experimental Populations of *Acinetobacter Baylyi*', *Molecular biology and evolution*, 40(3), pp. 1–11. doi:10.1093/molbev/msad028.

Shimizu, K. (2013) 'Metabolic Regulation of a Bacterial Cell System with Emphasis on *Escherichia coli* Metabolism', *ISRN Biochemistry*, 2013, pp. 1–47. doi:10.1155/2013/645983.

Shin, J. and Ko, K.S. (2015) 'Effect of plasmids harbouring blaCTX-M on the virulence and fitness of *Escherichia coli* ST131 isolates', *International Journal of Antimicrobial Agents*, 46, pp. 214–218. doi:10.1016/j.ijantimicag.2015.04.012.

Shintani, M. *et al.* (2010) 'Response of the *Pseudomonas* host chromosomal transcriptome to carriage of the IncP-7 plasmid pCAR1', *Environmental Microbiology*, 12(6), pp. 1413–1426. doi:10.1111/j.1462-2920.2009.02110.x.

Silva, F., Queiroz, J.A. and Domingues, F.C. (2012) 'Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*', *Biotechnology Advances*, 30(3), pp. 691–708. doi:10.1016/j.biotechadv.2011.12.005.

Da Silva, G.J. and Domingues, S. (2017) 'Interplay between colistin resistance, virulence and fitness in *Acinetobacter baumannii*', *Antibiotics*, 6(4), pp. 1–11. doi:10.3390/antibiotics6040028.

Sirithanakorn, C. and Cronan, J.E. (2021) 'Biotin, a universal and essential cofactor: Synthesis, ligation and regulation', *FEMS Microbiology Reviews*, 45(4), pp. 1–18. doi:10.1093/femsre/fuab003.

Smillie, C. *et al.* (2010) 'Mobility of Plasmids', *Microbiology and Molecular Biology Reviews*, 74(3), pp. 434–452. doi:10.1128/mnbr.00020-10.

Song, L. *et al.* (2013) '*Chlamydia trachomatis* plasmid-encoded *pgp4* is a transcriptional regulator of virulence-associated genes', *Infection and Immunity*, 81(3), pp. 636–644. doi:10.1128/IAI.01305-12.

[EcoCyc23] Karp, P.D., Paley, S., Caspi, R., Kothari, A., Krummenacker, M., Midford, P., Moore, L.R., Subhraveti, P., Gama-Castro, S., Tierrafria, V., Paloma, L., Muñiz-Rascado, L., Bonavides-Martinez, C., Santos-Zavaleta, A., Mackie, A., Sun, G., Ahn-Horst, T.A., Choi, H., Covert, M.W., Collado-Vides, J. and Paulsen, I., [The EcoCyc database in 2023](#)

Stoesser, N. *et al.* (2016) 'Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131', *mBio*, 7(2), pp. 1–15. doi:10.1128/mBio.02162-15.

Szymanski, Jędrzej Jozefczuk, S. *et al.* (2009) 'Stability of Metabolic Correlations under Changing Environmental Conditions in *Escherichia coli* – A Systems Approach', *PLOS ONE* 4, 4(10). doi: 10.1371/journal.pone.0007441

Takahashi, Y. *et al.* (2015) 'Modulation of primary cell function of host *Pseudomonas* bacteria by the conjugative plasmid pCAR1', *Environmental Microbiology*, 17(1), pp. 134–155. doi:10.1111/1462-2920.12515.

Tong, L. (2013) 'Structure and function of biotin-dependent carboxylases', *Cellular and Molecular Life Sciences*, 70(5), pp. 863–891. doi:10.1007/s00018-012-1096-0.

Tweeddale, H., Notley-McRobb, L. and Ferenci, T. (1998) 'Effect of Slow Growth on Metabolism of *Escherichia coli*, as Revealed by Global Metabolite Pool ("Metabolome") Analysis', *JOURNAL OF BACTERIOLOGY*, 180(19), pp. 5109–5116. Available at: <https://journals.asm.org/journal/jb>.

Valle, J. *et al.* (2008) 'The amino acid valine is secreted in continuous-flow bacterial biofilms', *Journal of Bacteriology*, 190(1), pp. 264–274. doi:10.1128/JB.01405-07.

Vasileva, D. *et al.* (2018) 'Proteome and acylome analyses of the functional interaction

network between the carbazole-degradative plasmid pCAR1 and host *Pseudomonas putida* KT2440', *Environmental Microbiology Reports*, 10(3), pp. 299–309. doi:10.1111/1758-2229.12639.

Venanzio, G. Di *et al.* (2019) 'Multidrug-resistant plasmids repress chromosomally encoded T6SS to enable their dissemination', *Proceedings of the National Academy of Sciences of the United States of America*, 116(4), pp. 1378–1383. doi:10.1073/pnas.1812557116.

Vincent, I.M. *et al.* (2016) 'Untargeted metabolomics to ascertain antibiotic modes of action', *Antimicrobial Agents and Chemotherapy*, 60(4), pp. 2281–2291. doi:10.1128/AAC.02109-15.

Wagner, S. *et al.* (2007) 'Consequences of membrane protein overexpression in *Escherichia coli*', *Molecular and Cellular Proteomics*, 6(9), pp. 1527–1550. doi:10.1074/mcp.M600431-MCP200.

Wang, Z. *et al.* (2016) 'GC-MS-Based Metabolome and Metabolite Regulation in Serum-Resistant *Streptococcus agalactiae*', *Journal of Proteome Research*, 15(7), pp. 2246–2253. doi:10.1021/acs.jproteome.6b00215.

Want, E.J., Cravatt, B.F. and Siuzdak, G. (2005) 'The expanding role of mass spectrometry in metabolite profiling and characterization', *ChemBioChem*, pp. 1941–1951. doi:10.1002/cbic.200500151.

Wei, Y. *et al.* (2021) 'Early Breast Cancer Detection Using Untargeted and Targeted Metabolomics', *J. Proteome Res*, 20, pp. 3124–3133. doi:10.17632/kc9g8ybk45.1.

Wu, W. *et al.* (2019) 'NDM Metallo-beta-Lactamases and Their Bacterial Producers in Health Care Settings', *Clinical microbiology reviews*, 32(2), pp. 1–45. 8. doi: 10.1128/CMR.00115-18.

Xu, Y.J. *et al.* (2014) 'Recent developments and applications of metabolomics in microbiological investigations', *TrAC - Trends in Analytical Chemistry*. Elsevier B.V., pp. 37–48. doi:10.1016/j.trac.2013.12.009.

Zhang, L., Alfano, J.R. and Becker, D.F. (2015) 'Proline metabolism increases katG

expression and oxidative stress resistance in *Escherichia coli*, *Journal of Bacteriology*, 197(3), pp. 431–440. doi:10.1128/JB.02282-14.

Zhao, L. *et al.* (2019) 'Comparison of metabolic response between the planktonic and air-dried *Escherichia coli* to electrolysed water combined with ultrasound by ¹H NMR spectroscopy', *Food Research International*, 125(108607), pp. 1–13. doi:10.1016/j.foodres.2019.108607.

Supplementary Materials

Supplementary Chapter 3 S3 – Data Tables

Table S3.1 – Mass values (Da) with a Fold Change of 2.0 or greater and putative Identifications where possible for 9 E.coli strains when the plasmid carrying profiles were compared against plasmid free controls.

Strain	Mass	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	Fold Change	log2(FC)
F054	70.005	Methylamine	C0021 8	5	M+K	70.0054	31.0422	CH5N	2.0299	1.0214
F037	136.0858	-							2.4148	1.2719
F104	136.0858								10.959	3.4541
ELU39	136.0858								11.741	3.5534
GU15	136.0858								2.399	1.2624
F048	136.13162	Creatinine	C0079 1	22	M+Na	136.0481	113.0589	C4H7N3O	0.11428	-3.1294
F054	136.13162	Creatinine	C0079 1	22	M+Na	136.0481	113.0589	C4H7N3O	15.621	3.9654
F104	136.13162	Creatinine	C0079 1	22	M+Na	136.0481	113.0589	C4H7N3O	2.7381	1.4532
ELU39	136.13162	Creatinine	C0079 1	22	M+Na	136.0481	113.0589	C4H7N3O	0.28535	-1.8092
GU15	136.13162	Creatinine	C0079 1	22	M+Na	136.0481	113.0589	C4H7N3O	0.085993	-3.5396
F054	146.15521								3.11F054	1.6394
F022	147.06438	Pyrazinic acid	C1991 5	10	M+Na	147.0165	124.0273	C5H4N2O2	0.42509	-1.2342
		Benzyl-alcohol	C0055 6	39	M+K	147.0207	108.0575	C7H8O		
		4-Cresol	C0146 8	39	M+K	147.0207	108.0575	C7H8O		

F104	211.1012	N6-Acetyl-L-Lysine	C0272 7	25	M+Na	211.1053	188.1161	C8H16N2O 3	0.49985	-1.0004
		Decanoate	C0157 1	37	M+K	211.1077	171.1385	C10H19O2		
MG165 5	211.1012	N6-Acetyl-L-Lysine	C0272 7	25	M+Na	211.1053	188.1161	C8H16N2O 3	2.1908	1.1315
		Decanoate	C0157 1	37	M+K	211.1077	171.1385	C10H19O2		
F047	233.1467								0.47428	-1.0762
F104	233.14876								0.24024	-2.0575
F054	247.12517								0.33153	-1.5928
F104	247.12517								0.4166	-1.2633
F054	248.12694	Tetradecanoate		11	M+Na	248.1747	225.1855	C14H25O2	0.4625	-1.1125
F047	251.06503	D-Glyceric acid	C0025 8	36	M+H	251.0074	250.0002	C6H10CaO 8	0.4477	-1.1594
F047	266.18066								0.29143	-1.7788
F047	275.12669	4-Hydroxy-3-polyprenylbenzoate	C0584 8	36	M+H	275.1642	274.1569	C17H22O3	0.37122	-1.4297
F104	275.12766	4-Hydroxy-3-polyprenylbenzoate	C0584 8	36	M+H	275.1642	274.1569	C17H22O3	2.9549	1.5631
F054	275.13172	4-Hydroxy-3-polyprenylbenzoate	C0584 8	36	M+H	275.1642	274.1569	C17H22O3	9.473	3.2438
F022	275.13489	4-Hydroxy-3-polyprenylbenzoate	C0584 8	36	M+H	275.1642	274.1569	C17H22O3	2.2185	1.1496
MG165 5	280.95598								11.716	3.5505
F054	286.14576								2.0658	1.0467
F047	296.95104								2.6433	1.4024
GU15	296.95127								0.29292	-1.7714
F104	296.95246								0.46937	-1.0912

F037	307.22445								0.49981	-1.0006
GU15	308.14917								3.793	1.9234
ELU39	308.14924								4.6182	2.2073
F047	313.27234								0.48324	-1.0492
GU15	313.27234								2.2325	1.1587
GU15	313.39844								2.2256	1.1542
GU15	314.27759								2.0088	1.0064
ELU39	353.32385								2.3799	1.2509
F047	354.40425								0.45746	-1.1283
F047	356.28549								0.40282	-1.3118
F054	365.06347								2.1756	1.1214
F047	371.26914								0.41866	-1.2562
F054	415.16229								0.4571	-1.1294
MG165 5	481.2796								0.48633	-1.04
MG165 5	525.31009								0.39106	-1.3545
GU15	540.74636								2.7599	1.4646
F047	607.97027								0.47358	-1.0783
MG165 5	613.36561								0.3985	-1.3274
MG165 5	629.2655								0.4906	-1.0274
ELU39	662.51522								2.1303	1.0911
ELU39	664.52067								2.0728	1.0516
F047	683.57686								0.38913	-1.3617
ELU39	688.53377								2.3932	1.259
ELU39	690.54459								3.3654	1.7508
ELU39	691.54931								2.5932	1.3747
ELU39	692.55674								2.0569	1.0405
F022	693.24188								4.3015	2.1048

ELU39	701.58616								0.39151	-1.3529
ELU39	704.56329								2.8312	1.5014
ELU39	705.56621								2.2565	1.1741
F047	711.61334								0.39569	-1.3376
ELU39	712.53027								2.5051	1.3249
ELU39	716.55932								2.4449	1.2897
ELU39	717.56305								3.3521	1.745
ELU39	718.57616								3.0266	1.5977
ELU39	719.57945								2.2581	1.1751
ELU39	744.57516								2.1704	1.1179
GU15	746.55976								2.6058	1.3817

S3 – Principle Component Analysis

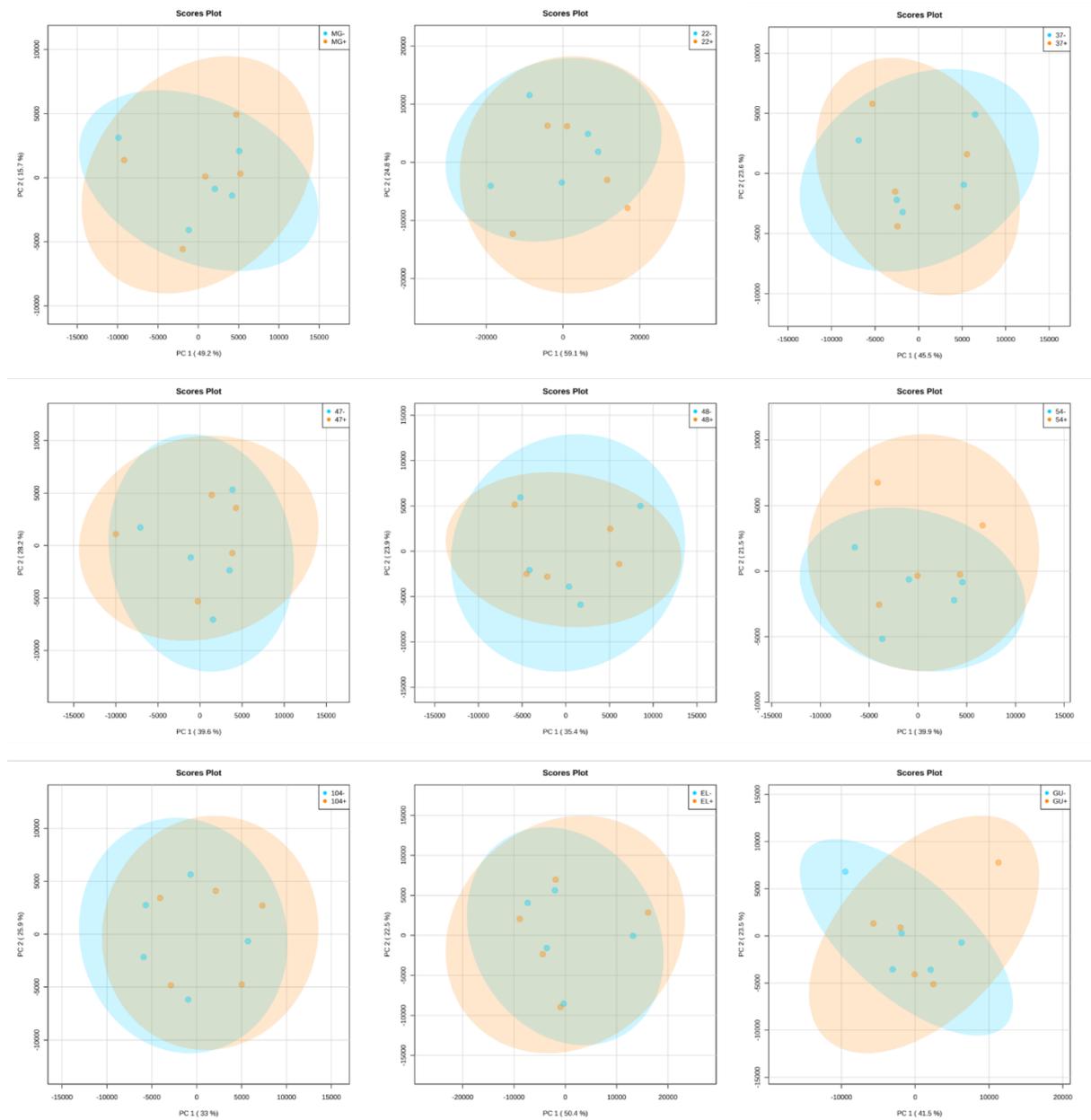


Figure S3.1. Principle component analysis for each of 9 *E. coli* strains comparing the plasmid free metabolome to the plasmid carrying metabolome.

S3 - Mass Spectrometry Parameters

Parameters for C:\MassLynx\Heather.PRO\ACQUDB\esi_pos_sens.EXP

Created by 4.2 SCN983

Lock Spray Configuration:

Tuning on Analyte

Temperature Correction:

Temperature Correction Disabled

Instrument Configuration:

Lteff	1800.0
Veff	7199.60
Resolution	10000
Min Points in Peak	2
Acquisition Device	WatersADC
Acquisition Algorithm	ADC Mode
ADC Trigger Threshold (V)	1.00
ADC Input Offset (V)	-1.50
Average Single Ion Intensity	24
ADC Amplitude Threshold	3
ADC Centroid Threshold	-1
ADC Ion Area Threshold	3
ADC Ion Area Offset	10
ADC Pushes Per IMS Increment	1
EDC Delay Coefficient	1.4100
EDC Delay Offset	0.4000

Experimental Instrument Parameters

Instrument

Parameter Filename

C:\MassLynx\Heather.PRO\ACQUDB\190221.IPR (MODIFIED)

Polarity	ES+
Capillary (kV)	3.6000
Source Temperature (°C)	100
Sampling Cone	80.0000
Source Offset	20.0000
Source Gas Flow (mL/min)	0.00
Desolvation Temperature (°C)	280
Cone Gas Flow (L/Hr)	0.0
Desolvation Gas Flow (L/Hr)	500.0
Nebuliser Gas Flow (Bar)	2.5
LM Resolution	4.4
HM Resolution	15.0
Aperture 1	0.0
Pre-filter	2.0
Ion Energy	1.0
Manual Trap Collision Energy	FALSE
Trap Collision Energy	4.0
Manual Transfer Collision Energy	FALSE
Transfer Collision Energy	2.0
Manual Gas Control	FALSE
Trap Gas Flow (mL/min)	2.00
HeliumCellGasFlow	180.00
IMS Gas Flow (mL/min)	90.00
Detector	3400
DetectorCache	0
Sample Infusion Flow Rate (µL/min)	5
Sample Flow State	LC
Sample Fill Volume (µL)	250
Sample Reservoir	Wash
LockSpray Infusion Flow Rate (µL/min)	10
LockSpray Flow State	Infusion

LockSpray Reservoir		B
LockSpray Capillary (kV)	3.0	
Use Manual LockSpray Collision Energy		FALSE
Collision Energy	4.0	
Acceleration1		70.0
Acceleration2		200.0
Aperture2	40.0	
Transport1	70.0	
Transport2	70.0	
Steering	-0.75	
Tube Lens	72	
Pusher		1900.0
Pusher Offset		-0.30
Puller	1370.0	
Pusher Cycle Time (μ s)		Automatic
Pusher Width (μ s)		Automatic
Collector	50	
Collector Pulse		10.0
Stopper		10
Stopper Pulse		20.0
Entrance	62	
Static Offset	180	
Puller Offset	0.00	
Reflectron Grid (kV)	1.470	
Flight Tube (kV)	10.00	
Reflectron (kV)		3.780
Use Manual Trap DC		FALSE
Trap DC Entrance	1.0	
Trap DC Bias		2.0
Trap DC		-2.0
Trap DC Exit	0.0	
Use Manual IMS DC		FALSE
IMS DC Entrance		-20.0

Helium Cell DC	1.0	
Helium Exit	-20.0	
IMSBias	2.0	
IMS DC Exit	20.0	
USe Manual Transfer DC	FALSE	
Transfer DC Entrance	5.0	
Transfer DC Exit	15.0	
Trap Manual Control	OFF	
Trap Wave Velocity (m/s)	300	
Trap Wave Height (V)	0.5	
IMS Manual Control	OFF	
IMS Wave Velocity (m/s)	300	
IMS Wave Height (V)	0.0	
Transfer Manual Control	OFF	
Transfer Wave Velocity (m/s)	247	
Transfer Wave Height (V)	0.2	
Step Wave 1 In Manual Control	OFF	
Enable Reverse Operation	OFF	
Step Wave 1 In Velocity (m/s)	300.0	
Step Wave 1 In Height	15.0	
Step Wave 1 Out Manual Control	OFF	
Step Wave 1 Out Velocity (m/s)	300.0	
Step Wave 1 Out Height	15.0	
Step Wave 2 Manual Control	OFF	
Step Wave 2 Velocity (m/s)	300.0	
Step Wave 2 Height	1.0	
Use Manual Step Wave DC		OFF
Step Wave TransferOffset	25.0	
Step Wave DiffAperture1	3.0	
Step Wave DiffAperture2	0.0	
Use Automatic RF Settings	TRUE	
StepWave1RFOffset	300.0	
StepWave2RFOffset	350.0	

Target Enhancement Enabled	FALSE
Target Enhancement Mode	EDC
Target Enhancement Mass	785.0
Target Enhancement Trap Height (V)	4.0
Target Enhancement Extract Height (V)	15.0
Mobility Trapping Manual Release Enabled	FALSE
Mobility Trapping Release Time (μ s)	500
Mobility Trap Height (V)	15.0
Mobility Extract Height (V)	0.0
Trag Gate LUT table enabled	FALSE
TriWave Trap Gate LookUp Table	
Using Drift Time Trimming	FALSE
Drift Time Bins	0
Using Mobility Delay after Trap Release	TRUE
IMS Wave Delay (μ s)	1000
Variable Wave Height Enabled	FALSE
Wave Height Ramp Type	Linear
Wave Height Start (V)	10.0
Wave Height End (V)	40.0
Wave Height Using Full IMS	TRUE
Wave Height Ramp (%)	100.0
Wave Height Look Up Table	
Variable Wave Velocity Enabled	FALSE
Wave Velocity Ramp Type	Linear
Wave Velocity Start (m/s)	1000.0
Wave Velocity End (m/s)	300.0
Wave Velocity Using Full IMS	TRUE
Wave Velocity Ramp (%)	100.0
Wave Velocity Look Up Table	
Backing	2.92e0
Source	7.25e-3
Sample Plate	1.19e3
Trap	8.67e-3

Helium Cell	9.97e-5
IMS	9.84e-5
Transfer	7.66e-3
TOF	4.45e-7
IMSRFOffset	300
IMSMobilityRFOffset	250
TrapRFOffset	300
Use Automatic RF Settings	TRUE
AutoStepWave1RFOffset	300
AutoStepWave2RFOffset	300
TransferRFOffset	350
MS Profile Type	Auto P
MSProfileMass1	100
MSProfileDwellTime1	20
MSProfileRampTime1	20
MSProfileMass2	300
MSProfileDwellTime2	20
MSProfileRampTime2	40
MSProfileMass3	500
PusherInterval	54.000000
PusherOffset	0.250000
LockMassValidSigma	5

Acquisition mass range

Start mass	50.000
End mass	1200.000

Calibration mass range

Start mass	0.000
End mass	0.000

Experiment Reference Compound Name: N/A

Function Parameters - Function 1 - TOF MS FUNCTION

Scan Time (sec)	1.000
Interscan Time (sec)	0.014
Start Mass	50.0
End Mass	1200.0
Start Time (mins)	0.00
End Time (mins)	3.00
Data Format	Continuum
Analyser	Sensitivity Mode
ADC Sample Frequency (GHz)	3.0
ADC Pusher Frequency (μ s)	54.0
ADC Pusher Width (μ s)	1.50
Use Tune Page Cone Voltage	YES
Using Auto Trap Collision Energy (eV)	4.000000
Using Auto Transfer Collision Energy (eV)	2.000000
Sensitivity	Normal
Dynamic Range	Normal
Save Collapsed Retention Time Data	No
Use Rule File Filtering	No
FragmentationMode	CID
Calibration	Dynamic 2

S3 - XCMS online Parameters

Feature Detection method	Centwave
Ppm	30
Minimum peak width	10
Maximum peak width	60
Retention Time Correction method	Obiwarp
Alignment	
Bw	5
Minfrac	0.5
Mzwid	0.025

S3 - Normalisation of Data

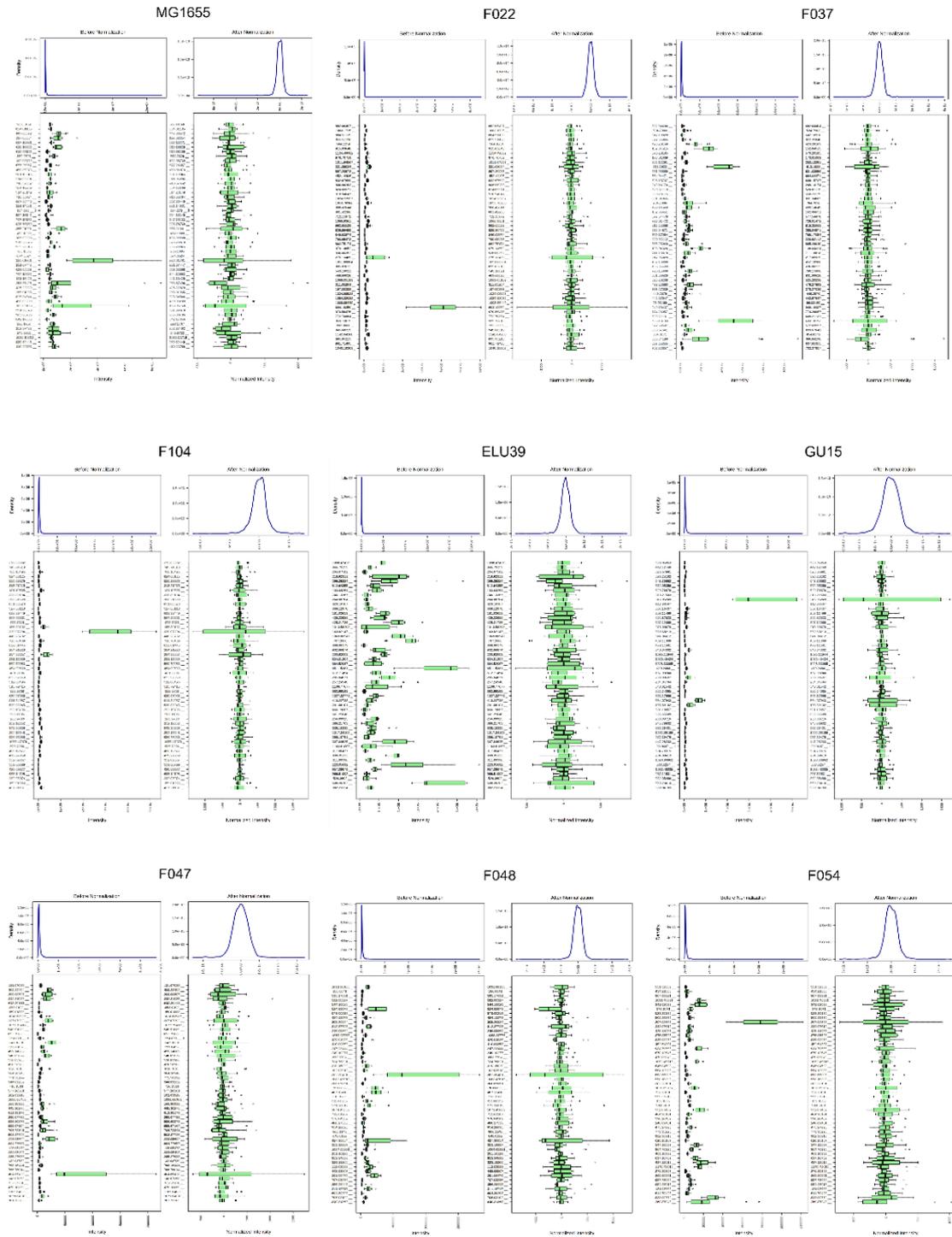


Figure S3.2 – Graphs showing the density and intensity of the data before and after normalization. Normalized by Pareto scaling to retain the shape of the data.

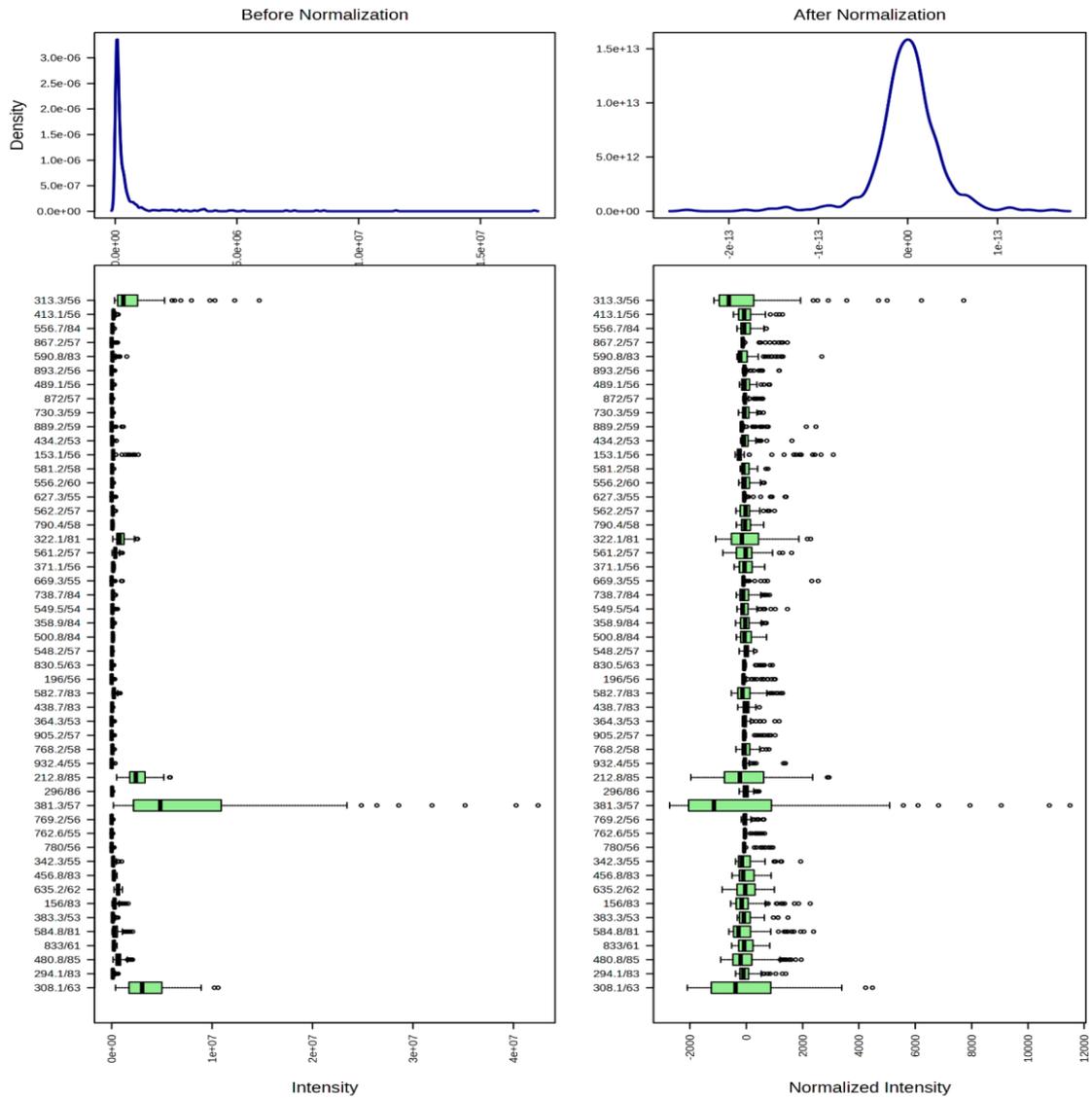


Figure S3.3 – Graphs showing the density and intensity of the data before and after normalization. Normalized by Pareto scaling to retain the shape of the data.

Supplementary Chapter 4

S4 – Data Tables

Table S4.1 Fold change summary of the total masses (Da) that cross the fold change threshold of +/- 2.0 when comparing the plasmid carrying profiles against the plasmid free profiles by strain. Metabolites were identified using KEGG and ECMDDB.

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	log2(FC)	Function
F022	96.07998								0.43264	-1.2088	
F022	118.09047	Betaine	C00719	31	M+H	118.0868	118.0868	C5H12NO2	0.36273	-1.463	Serine and threonine biosynthesis
		L-Valine	C00183	36	M+H	118.0863	117.079	C5H12NO2			Branched chain amino acid metabolism valine-isoleucine metabolism
F022	119.09351								0.39633	-1.3352	
F022	140.07101	Betaine	C00719	16	M+Na		118.0868	C5H12NO2	0.36083	-1.4706	Serine and threonine biosynthesis
		L-Valine	C00183	20	M+Na		117.079	C5H12NO2			Branched chain amino acid metabolism
F022	153.04427	Xanthine	C00385	23	M+H				0.46548	-1.1032	Adenine and guanine metabolism
		2-Keto-3-methyl-valerate	C00671	48	M+Na	153.0517	129.0552	C6H9O3			Branched chain amino acid metabolism valine-isoleucine metabolism
F022	153.07936	Ribitol	C00474	24	M+H	153.0757	152.0685	C5H12O5	0.47956	-1.0602	
		L-Arabitol	C00532	24	M+H	153.0757	152.0685	C5H12O5			
F022	156.04794	Betaine	C00719	34	M+K	156.0427	118.0868	C5H12NO2	0.30117	-1.7313	Serine and threonine biosynthesis
		L-Valine	C00183	37	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism
F022	213.02325	3-Carboxy-3-hydroxy-isocaproate	C02504	34	M+K	213.016	174.0528	C7H10O5	0.48593	-1.0412	Branched chain amino acid metabolism leucine biosynthesis
		2-Isopropyl-3-oxosuccinate	C04236	34	M+K	213.016	174.0528	C7H10O5			Branched chain amino acid metabolism leucine biosynthesis
		Shikimic acid	C00493	34	M+K	213.016	174.0528	C7H10O5			Shikimate pathway
ELU39	217.11175	gamma-Glutamyl-gamma-butyraldehyde	C15700	30	M+H	216.111	217.1183	C9H16N2O4	2.544	1.3471	GABA biosynthesis
ELU39	226.9613	2,5-Dichloro-4-oxohex-2-enedioate	C12835	46	M+H	226.9509	225.9436	C6H4Cl2O5	2.2061	1.1415	

F022	230.24096								0.39979	-1.3227	
ELU39	230.24292								3.7446	1.9048	
MG1655	250.88692								0.43583	-1.1982	
ELU39	254.84663								3.5313	1.8202	
F022	254.84667								2.2911	1.196	
F022	260.90304								0.49388	-1.0178	
MG1655	261.1383								0.47497	-1.0741	
F022	262.87619								0.44241	-1.1765	
MG1655	262.87666								2.2873	1.1936	
F022	270.15899	Ubiquinol-1	C00390	41	M+NH4	270.17	252.1362	C14H20O4	0.33108	-1.5948	Ubiquinone biosynthesis
MG1655	277.08735	Nicotinamide riboside	C03150	26	M+Na	277.08	255.0981	C11H15N2O5	0.44779	-1.1591	
		D-Galactosamine 6-phosphate	C06377	28	M+NH4	277.0795	259.0457	C6H14NO8P			
		alpha-D-Glucosamine 1-phosphate	C06156	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	28	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
F022	329.07394	N-Succinyl-L,L-2,6-diaminopimelate	C04421	2	M+K	329.0746	290.1114	C11H18N2O7	0.43099	-1.2143	Lysine metabolism
		Argininosuccinic acid	C03406	36	M+K	329.0858	290.1226	C10H18N4O6			Arginine biosynthesis
		2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	C16519	39	M+H	329.0867	328.0794	C14H16O9			Menaquinone biosynthesis
F022	347.05896								0.12	-3.0589	
F022	347.08038	Adenosine 2',3'-cyclic phosphate	C02353	17	M+NH4	347.0863	329.0525	C10H12N5O6P	0.42755	-1.2258	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	17	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F022	347.58478								0.49322	-1.0197	
ELU39	362.94504								2.157	1.109	
F022	385.07204								0.45366	-1.1403	
ELU39	404.93512								2.5603	1.3563	
ELU39	412.77255								2.1846	1.1273	
ELU39	426.84985								2.0265	1.019	
ELU39	430.936								2.0084	1.0061	
F022	499.15677								0.44123	-1.1804	
F022	520.11624								0.41277	-1.2766	
MG1655	520.13006								0.35585	-1.4907	
F022	520.63378								0.4467	-1.1626	
F022	521.13348								0.4868	-1.0386	

F022	541.1								0.41977	-1.2523
ELU39	546.76873								2.141	1.0983
F022	578.14658								0.495	-1.0145
ELU39	590.75225								2.0206	1.0148
F022	596.15631								0.4947	-1.0154
MG1655	638.6573								0.44215	-1.1774
MG1655	693.16966								0.20921	-2.257
F022	693.17004								0.44041	-1.1831
ELU39	693.17201								2.1898	1.1308
F022	694.17386								0.44444	-1.1699
F022	695.17316								0.40985	-1.2868
ELU39	742.7783								2.2841	1.1916
ELU39	764.75922								2.0853	1.0602
ELU39	796.68474								2.5129	1.3293
F022	808.54628								0.40686	-1.2974
F022	815.86261								0.39291	-1.3477
F022	823.86119								0.49104	-1.0261
F022	866.21927								0.3994	-1.3241
F022	888.17153								0.49332	-1.0194
F022	1039.2673								0.26872	-1.8958
F022	1040.26905								0.33133	-1.5936
F022	1062.25633								0.49898	-1.003

Table S4.2 Fold change summary of the total masses (Da) that cross the fold change threshold of +/- 2.0 when comparing the antibiotic selection profiles against the plasmid free profiles by strain. Metabolites were identified using KEGG and ECMDB

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	Log2(FC)	Function
F022	118.09047	Betaine	C00719	31	M+H	118.0868	118.0868	C5H12NO2	0.39376	-1.3446	Serine and threonine biosynthesis
		L-Valine	C00183	36	M+H	118.0863	117.079	C5H11NO2			Branched chain amino acid metabolism
F022	156.04794	Betaine	C00719	34	M+K	156.0427	118.0868	C5H12NO2	0.34037	-1.5548	Serine and threonine biosynthesis
		L-Valine	C00183	37	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism
ELU39	166.0853	D-Phenylalanine	C02265	6	M+H	166.0863	165.079	C9H11NO2	0.48577	-1.0417	Phenylalanine biosynthesis
		L-Phenylalanine	C00079	6	M+H	166.0863	165.079	C9H11NO2			Phenylalanine biosynthesis
MG1655	174.94267								0.4903	-1.0283	
MG1655	198.04356	7-Cyano-7-carbaguanine	C15996	25	M+Na	198.0386	175.0494	C7H5N5O	0.41098	-1.2829	Folate biosynthesis
ELU39	217.10864	gamma-Glutamyl-gamma-butyraldehyde	C15700	44	M+H	217.1183	216.111	C9H16N2O4	4.7016	2.2332	GABA biosynthesis
MG1655	220.0286	L-Tyrosine	C00082	38	M+K	220.0371	181.0739	C9H11NO3	0.47746	-1.0666	Tyrosine biosynthesis, thiamine biosynthesis
ELU39	230.24152								2.0834	1.0589	
ELU39	242.04925								0.41289	-1.2762	
F022	247.06603	-Phosphoribosylamine	C03090	12	M+NH4	247.069	229.0351	C5H12NO7P	0.35436	-1.4967	Denovo puring biosynthesis
F022	254.84738								2.8627	1.5174	
F022	270.15827	Ubiquinol-1	C00390	43	M+NH4	270.17	252.1362	C14H20O4	0.38333	-1.3834	Ubiquinone biosynthesis
F022	338.89928								2.1589	1.1103	
F022	347.05936								0.47002	-1.0892	
ELU39	347.08318	Adenosine 2',3'-cyclic phosphate	C02353	9	M+NH4	347.0863	329.0525	C10H12N5O6P	4.7682	2.2534	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	9	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
MG1655	404.80551								0.49128	-1.0254	
ELU39	404.93164								0.31718	-1.6566	
F022	436.87816								2.5865	1.371	
ELU39	660.64963								2.5848	1.3701	
ELU39	682.73257								2.2513	1.1708	
ELU39	693.1525								5.0024	2.3226	
ELU39	715.13273								4.0504	2.0181	
ELU396	764.73754								2.2546	1.1729	

Table S4.3. Total masses (Da) that cross the fold change threshold of +/-2.0 comparing the antibiotic selection profiles against the plasmid carrying profiles by strain. Metabolites were identified using KEGG and ECMDDB

Strain	M/z	ID	KEGG	ppm	Adduct	Adduct +Compound M/Z	Monoisotopic Mass	Chemical Formula	FC	Log2(FC)	Function
F022	99.04636	Propylene glycol	C00583	48	M+Na	99.0416	76.0524	C3H8O2	0.3453	-1.5341	By-product of glycolysis
F022	115.0412	Glycerol	C00116	40	M+Na	115.0366	92.0473	C3H8O3	2.0506	1.036	Feeds into glycolysis
F022	135.069	R)-2,3-Dihydroxy-isovalerate	C04272	28	M+H	135.0652	134.0579	C5H10O4	2.1982	1.1363	Valine/isoleucine branched chain amino acid metabolism
		Deoxyribose	C01801	28	M+H	135.0652	134.0579	C5H10O4			
F022	153.0794	Ribitol	C00474	24	M+H	153.0757	152.0685	C5H12O5	2.0533	1.038	
		L-Arabitol	C00532	24	M+H	153.0757	152.0685	C5H12O5			
MG1655	174.9427								0.47841	-1.0637	
F022	195.0083	Glycerol 3-phosphate	C00093	28	M+Na	195.0029	172.0137	C3H9O6P	2.1478	1.1028	
F022	213.0233	3-Carboxy-3-hydroxy-isocaproate	C02504	34	M+K	213.016	174.0528	C7H10O5	2.2291	1.1565	Branched chain amino acid metabolism leucine biosynthesis
		2-Isopropyl-3-oxosuccinate	C04236	34	M+K	213.016	174.0528	C7H10O5			Branched chain amino acid metabolism leucine biosynthesis
		Shikimic acid	C00493	34	M+K	213.016	174.0528	C7H10O5			Shikimate pathway
ELU39	242.0494								0.35686	-1.4866	
F022	247.1377								0.48772	-1.0359	
ELU39	254.8458								0.46021	-1.1196	
F022	260.9036								2.2249	1.1538	
F022	304.9481								2.3769	1.2491	
MG1655	316.8844								2.5795	1.3671	
F022	329.0711	N-Succinyl-L,L-2,6-diaminopimelate	C04421	11	M+K	329.0746	290.1114	C11H18N2O7	2.2688	1.1819	Lysine metabolism
		Argininosuccinic acid	C03406	45	M+K	329.0858	290.1226	C10H18N4O6			Arginine biosynthesis
		2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	C16519	47	M+H	329.0867	328.0794	C14H16O9			Menaquinone biosynthesis
F022	347.0599								2.8285	1.5001	
ELU39	347.0818	Adenosine 2',3'-cyclic phosphate	C02353	13	M+NH4	347.0863	329.0525	C10H12N5O6P	3.4007	1.7658	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	13	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation

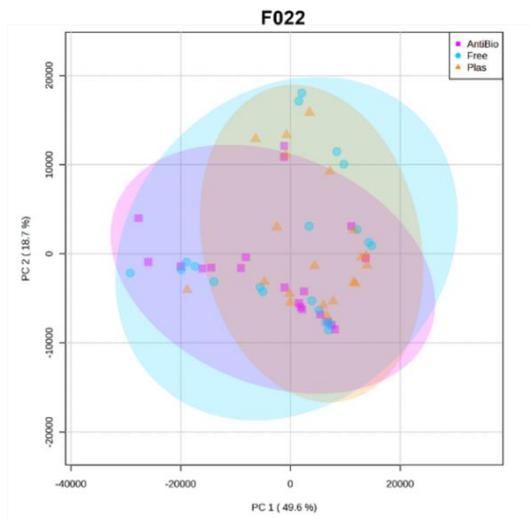
MG1655	347.0851	Adenosine 2',3'-cyclic phosphate	C02353	4	M+NH4	347.0863	329.0525	C10H12N5O6P	2.5321	1.3403	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	4	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F022	347.0879	Adenosine 2',3'-cyclic phosphate	C02353	4	M+NH4	347.0863	329.0525	C10H12N5O6P	2.2155	1.1476	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575	4	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F022	347.5848								2.3	1.2016	
F022	358.0724	S-Formylglutathione	C01031	12	M+Na	358.0679	335.0787	C11H17N3O7S	2.0873	1.0617	Methane metabolism
F022	369.0697								2.5728	1.3634	
ELU39	398.8441								3.1552	1.6577	
ELU39	404.9322								0.11981	-3.0612	
F022	409.1748								2.0805	1.0569	
F022	462.445								2.1938	1.1334	
F022	469.7692								2.1208	1.0846	
F022	499.1608								2.1493	1.1039	
F022	500.1609								2.0636	1.0452	
ELU39	502.85								0.45617	-1.1324	
F022	520.1261								2.285	1.1922	
F022	520.6239								2.2336	1.1594	
F022	521.1231								2.2784	1.188	
F022	531.109								2.4302	1.281	
F022	531.6092								2.3913	1.2578	
ELU39	534.8558								0.49959	-1.0012	
F022	541.0803								2.3724	1.2463	
ELU39	546.8318								0.32046	-1.6418	
F022	561.1086								2.0018	1.0013	
F022	577.8001								2.2354	1.1605	
F022	578.136								2.363	1.2406	
F022	578.5359								2.3221	1.2154	
F022	585.1369								2.2843	1.1918	
F022	585.4745								2.0176	1.0126	
F022	592.4567								2.1904	1.1312	
F022	606.6503								2.1267	1.0886	
F022	606.8986								2.0521	1.0371	
ELU39	660.6517								3.038	1.6031	
F022	675.1639								3.2815	1.7144	
F022	676.1462								2.0865	1.0611	
ELU39	676.8052								0.35836	-1.4805	

ELU39	682.7312								3.2472	1.6992	
ELU39	682.7544								0.41454	-1.2704	
F022	684.164								2.1138	1.0799	
ELU39	693.1537								2.2797	1.1888	
F022	693.1585								2.2049	1.1407	
F022	693.4101								2.0616	1.0437	
F022	693.6755								2.127	1.0888	
F022	694.1806								2.21	1.1441	
F022	695.1583								2.3347	1.2233	
F022	696.1748								2.0056	1.004	
F022	700.484								2.1951	1.1343	
ELU39	700.7594								0.4963	-1.0107	
F022	701.1654								2.8292	1.5004	
F022	704.1633								3.8042	1.9276	
F022	704.6641								2.4489	1.2921	
F022	707.8293								2.2233	1.1527	
F022	708.4742								2.0593	1.0421	
F022	712.1524								2.1136	1.0797	
F022	713.1514								2.0255	1.0183	
ELU39	715.1322								4.5337	2.1807	
F022	715.4912								2.118	1.0827	
F022	732.1231								2.0764	1.0541	
ELU39	742.7792								0.45989	-1.1207	
ELU39	764.7592								0.31643	-1.66	
F022	779.9226								2.3799	1.2509	
F022	785.4189								2.0549	1.0391	
F022	808.5463								2.0878	1.062	
F022	809.2014								2.0247	1.0177	
F022	815.8646								2.7339	1.451	
F022	816.1998								2.2517	1.171	
F022	816.5364								2.2072	1.1422	
F022	828.4881								2.0754	1.0534	
F022	831.7824								2.0436	1.0311	
F022	845.2236								0.48066	-1.0569	
F022	866.2215								2.199	1.1368	
F022	866.4733								2.174	1.1203	
F022	866.7228								2.1086	1.0763	
F022	888.1718								2.2862	1.1929	
F022	888.709								2.1651	1.1144	
F022	889.1985								2.0353	1.0252	

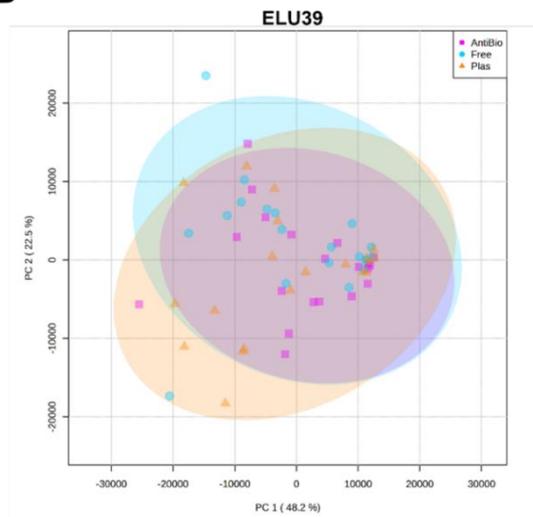
F022	896.1966								2.0443	1.0316	
F022	923.8691								2.1241	1.0869	
F022	931.8987								2.0533	1.0379	
F022	932.8868								2.0026	1.0019	
F022	968.9474								2.0696	1.0493	
F022	1031.287								2.0307	1.022	
F022	1040.274								2.2722	1.1841	
F022	1061.246								2.3537	1.235	
F022	1061.747								2.1563	1.1085	
F022	1062.212								2.0094	1.0068	
F022	1062.249								2.5801	1.3674	

S4 – Principle Component Analysis

A



B



C

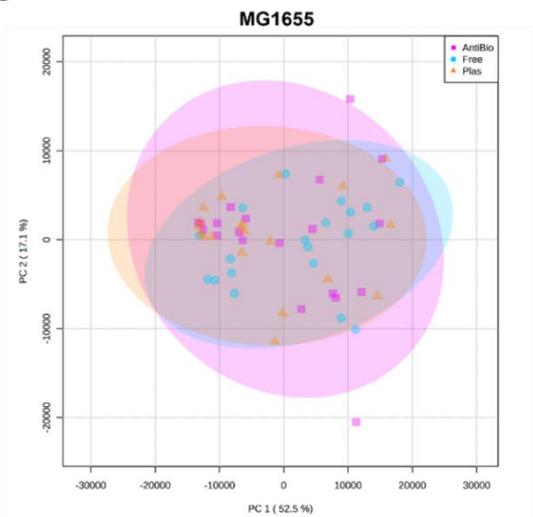
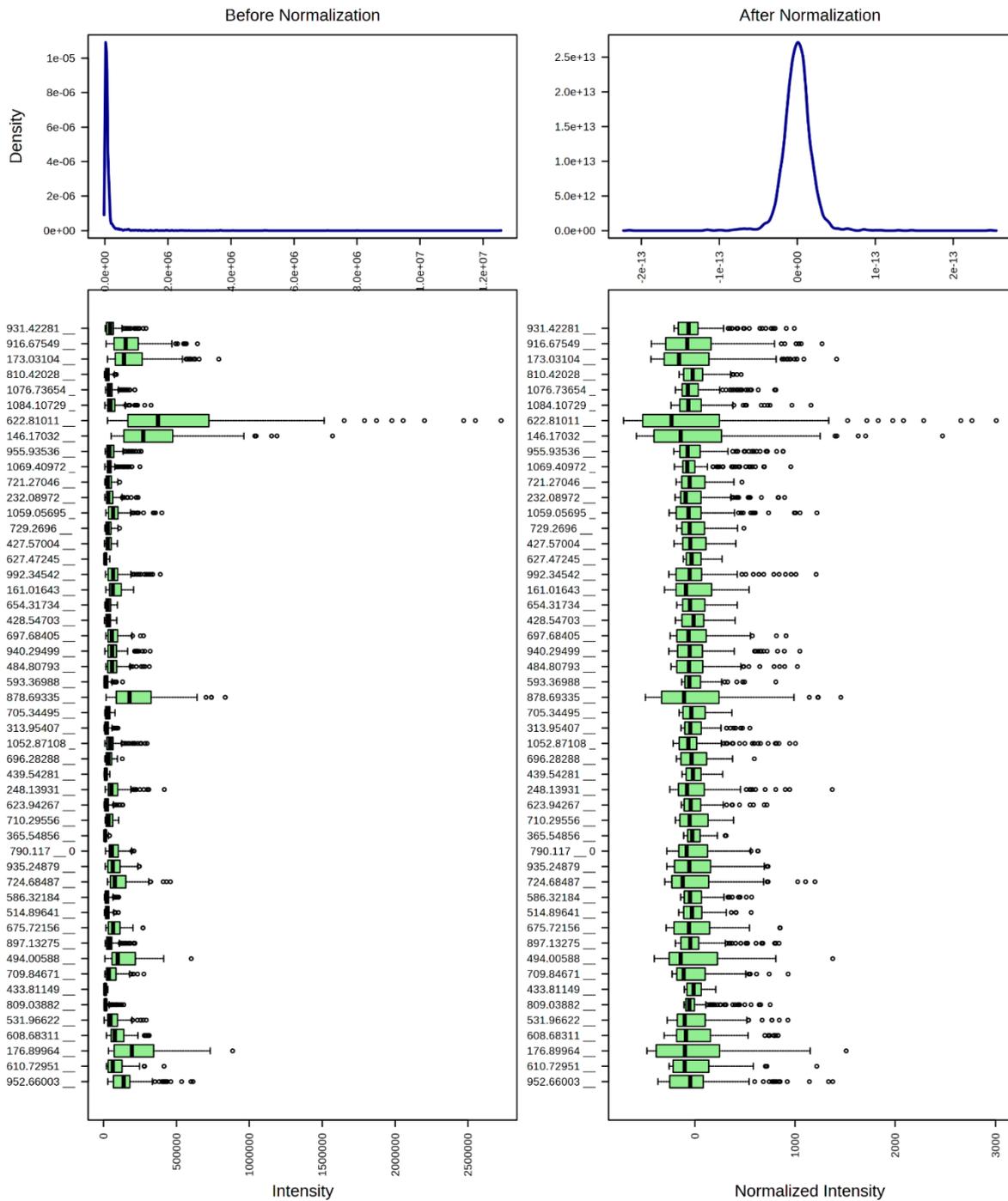


Figure S4.1, Principle component analysis of the metabolic profiles of each strain grouped by evolutionary treatment. A – F022. B- ELU39. C – MG1655. Antibio – plasmid carrying with cefotaxime selection. Plas – plasmid carrying. Free – plasmid free.

S4 – Normalisation of Data



S4. 2 Graphs showing the density and intensity of the data before and after normalization. Normalized by Pareto scaling to retain the shape of the data.

S4 –Mass Spectrometry Parameters

Parameters for C:\MassLynx\Heather.PRO\ACQUDB\esi_pos_sens_3min.EXP

Created by 4.2 SCN983

Lock Spray Configuration:

Tuning on Analyte

Temperature Correction:

Temperature Correction Disabled

Instrument Configuration:

Lteff	1800.0
Veff	7199.60
Resolution	10000
Min Points in Peak	2
Acquisition Device	WatersADC
Acquisition Algorithm	ADC Mode
ADC Trigger Threshold (V)	1.00
ADC Input Offset (V)	-1.50
Average Single Ion Intensity	29
ADC Amplitude Threshold	3
ADC Centroid Threshold	-1
ADC Ion Area Threshold	3
ADC Ion Area Offset	10
ADC Pushes Per IMS Increment	1
EDC Delay Coefficient	1.4100
EDC Delay Offset	0.4000

Experimental Instrument Parameters

Instrument Parameter Filename

C:\MassLynx\Heather.PRO\ACQUDB\esi241121.IPR (MODIFIED)

Polarity	ES+
Capillary (kV)	3.2000

Source Temperature (°C)	100
Sampling Cone	10.0000
Source Offset	40.0000
Source Gas Flow (mL/min)	0.00
Desolvation Temperature (°C)	280
Cone Gas Flow (L/Hr)	0.0
Desolvation Gas Flow (L/Hr)	600.0
Nebuliser Gas Flow (Bar)	7.0
LM Resolution	4.4
HM Resolution	15.0
Aperture 1	0.0
Pre-filter	2.0
Ion Energy	1.0
Manual Trap Collision Energy	FALSE
Trap Collision Energy	4.0
Manual Transfer Collision Energy	FALSE
Transfer Collision Energy	2.0
Manual Gas Control	FALSE
Trap Gas Flow (mL/min)	2.00
HeliumCellGasFlow	180.00
IMS Gas Flow (mL/min)	90.00
Detector	3375
DetectorCache	0
Sample Infusion Flow Rate (µL/min)	5
Sample Flow State	LC
Sample Fill Volume (µL)	250
Sample Reservoir	B
LockSpray Infusion Flow Rate (µL/min)	10
LockSpray Flow State	Infusion
LockSpray Reservoir	B
LockSpray Capillary (kV)	3.0
Use Manual LockSpray Collision Energy	FALSE
Collision Energy	4.0

Acceleration1	70.0
Acceleration2	200.0
Aperture2	40.0
Transport1	70.0
Transport2	70.0
Steering	-0.83
Tube Lens	77
Pusher	1900.0
Pusher Offset	-0.55
Puller	1370.0
Pusher Cycle Time (μ s)	Automatic
Pusher Width (μ s)	Automatic
Collector	50
Collector Pulse	10.0
Stopper	10
Stopper Pulse	20.0
Entrance	68
Static Offset	180
Puller Offset	0.00
Reflectron Grid (kV)	1.469
Flight Tube (kV)	10.00
Reflectron (kV)	3.780
Use Manual Trap DC	FALSE
Trap DC Entrance	1.0
Trap DC Bias	2.0
Trap DC	-2.0
Trap DC Exit	0.0
Use Manual IMS DC	FALSE
IMS DC Entrance	-20.0
Helium Cell DC	1.0
Helium Exit	-20.0
IMSBias	2.0
IMS DC Exit	20.0

Use Manual Transfer DC		FALSE
Transfer DC Entrance		5.0
Transfer DC Exit	15.0	
Trap Manual Control		OFF
Trap Wave Velocity (m/s)	300	
Trap Wave Height (V)		0.5
IMS Manual Control	OFF	
IMS Wave Velocity (m/s)		300
IMS Wave Height (V)		0.0
Transfer Manual Control		OFF
Transfer Wave Velocity (m/s)		247
Transfer Wave Height (V)	0.2	
Step Wave 1 In Manual Control		OFF
Enable Reverse Operation	OFF	
Step Wave 1 In Velocity (m/s)		300.0
Step Wave 1 In Height		15.0
Step Wave 1 Out Manual Control		OFF
Step Wave 1 Out Velocity (m/s)		300.0
Step Wave 1 Out Height		15.0
Step Wave 2 Manual Control		OFF
Step Wave 2 Velocity (m/s)		300.0
Step Wave 2 Height		1.0
Use Manual Step Wave DC		OFF
Step Wave TransferOffset	25.0	
Step Wave DiffAperture1		3.0
Step Wave DiffAperture2		0.0
Use Automatic RF Settings		TRUE
StepWave1RFOffset		300.0
StepWave2RFOffset		350.0
Target Enhancement Enabled		FALSE
Target Enhancement Mode		EDC
Target Enhancement Mass		785.0
Target Enhancement Trap Height (V)	4.0	

Target Enhancement Extract Height (V)	15.0
Mobility Trapping Manual Release Enabled	FALSE
Mobility Trapping Release Time (μs)	500
Mobility Trap Height (V)	15.0
Mobility Extract Height (V)	0.0
Trag Gate LUT table enabled	FALSE
TriWave Trap Gate LookUp Table	
Using Drift Time Trimming	FALSE
Drift Time Bins	0
Using Mobility Delay after Trap Release	TRUE
IMS Wave Delay (μs)	1000
Variable Wave Height Enabled	FALSE
Wave Height Ramp Type	Linear
Wave Height Start (V)	10.0
Wave Height End (V)	40.0
Wave Height Using Full IMS	TRUE
Wave Height Ramp (%)	100.0
Wave Height Look Up Table	
Variable Wave Velocity Enabled	FALSE
Wave Velocity Ramp Type	Linear
Wave Velocity Start (m/s)	1000.0
Wave Velocity End (m/s)	300.0
Wave Velocity Using Full IMS	TRUE
Wave Velocity Ramp (%)	100.0
Wave Velocity Look Up Table	
Backing	2.63e0
Source	6.52e-3
Sample Plate	1.25e3
Trap	8.78e-3
Helium Cell	1.00e-4
IMS	9.84e-5
Transfer	7.71e-3
TOF	4.40e-7

IMSRFOffset	300
IMSMobilityRFOffset	250
TrapRFOffset	300
Use Automatic RF Settings	TRUE
AutoStepWave1RFOffset	300
AutoStepWave2RFOffset	300
TransferRFOffset	350
MS Profile Type	Auto P
MSProfileMass1	100
MSProfileDwellTime1	20
MSProfileRampTime1	20
MSProfileMass2	300
MSProfileDwellTime2	20
MSProfileRampTime2	40
MSProfileMass3	500
PusherInterval	54.000000
PusherOffset	0.250000
LockMassValidSigma	5

Acquisition mass range

Start mass	50.000
End mass	1200.000

Calibration mass range

Start mass	0.000
End mass	0.000

Experiment Reference Compound Name: N/A

Function Parameters - Function 1 - TOF MS FUNCTION

Scan Time (sec)	1.000
Interscan Time (sec)	0.014
Start Mass	50.0
End Mass	1200.0

Start Time (mins)	0.00
End Time (mins)	3.00
Data Format	Continuum
Analyser	Sensitivity Mode
ADC Sample Frequency (GHz)	3.0
ADC Pusher Frequency (μ s)	54.0
ADC Pusher Width (μ s)	1.50
Use Tune Page Cone Voltage	YES
Using Auto Trap Collision Energy (eV)	4.000000
Using Auto Transfer Collision Energy (eV)	2.000000
Sensitivity	Normal
Dynamic Range	Normal
Save Collapsed Retention Time Data	No
Use Rule File Filtering	No
FragmentationMode	CID
Calibration	Dynamic 2

Supplementary Chapter 5

S5 – Normalisation of Data

F022

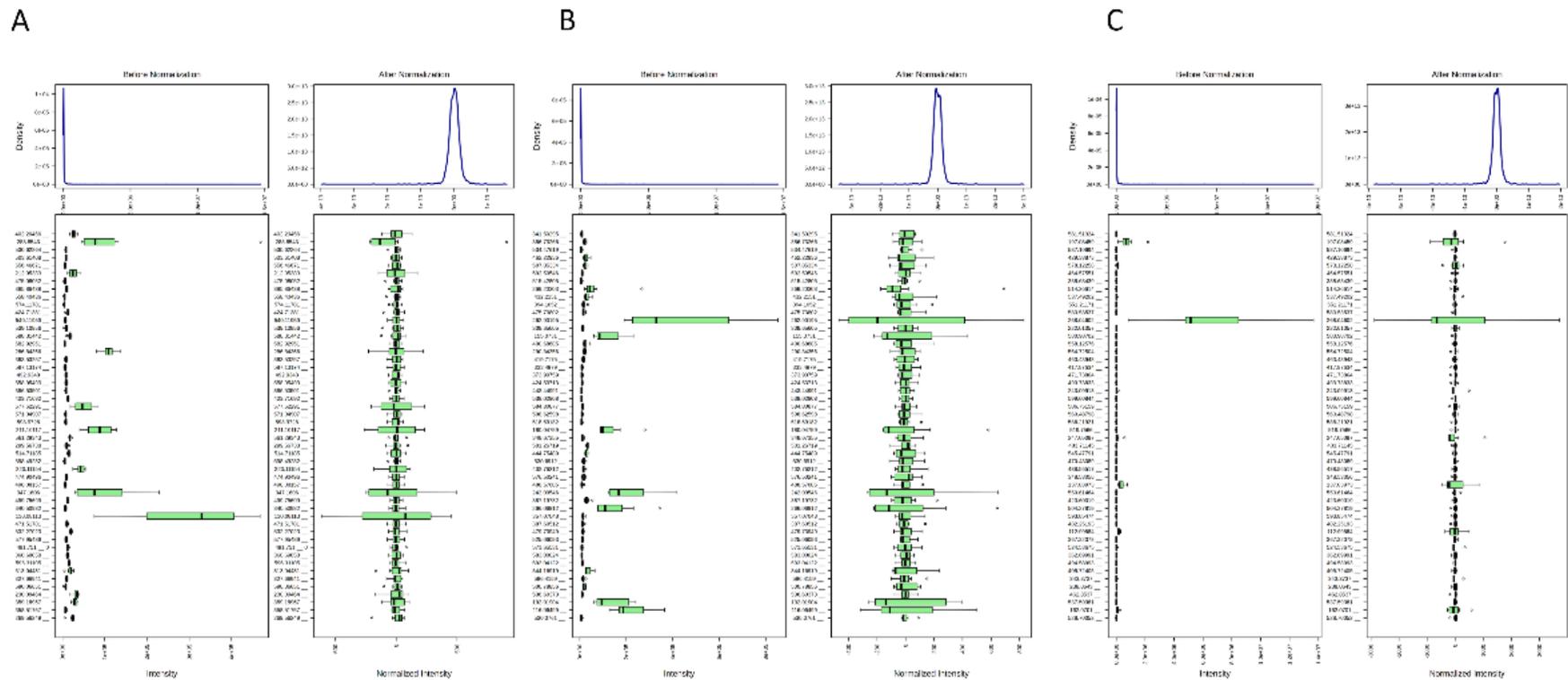


Fig S5.1 Normalisation plots for strain F022 for pairwise comparisons of the 3 treatments against the antibiotic free control. A) ciprofloxacin. B) cefotaxime. C) kanamycin. The data was pareto scaled.

MG1655

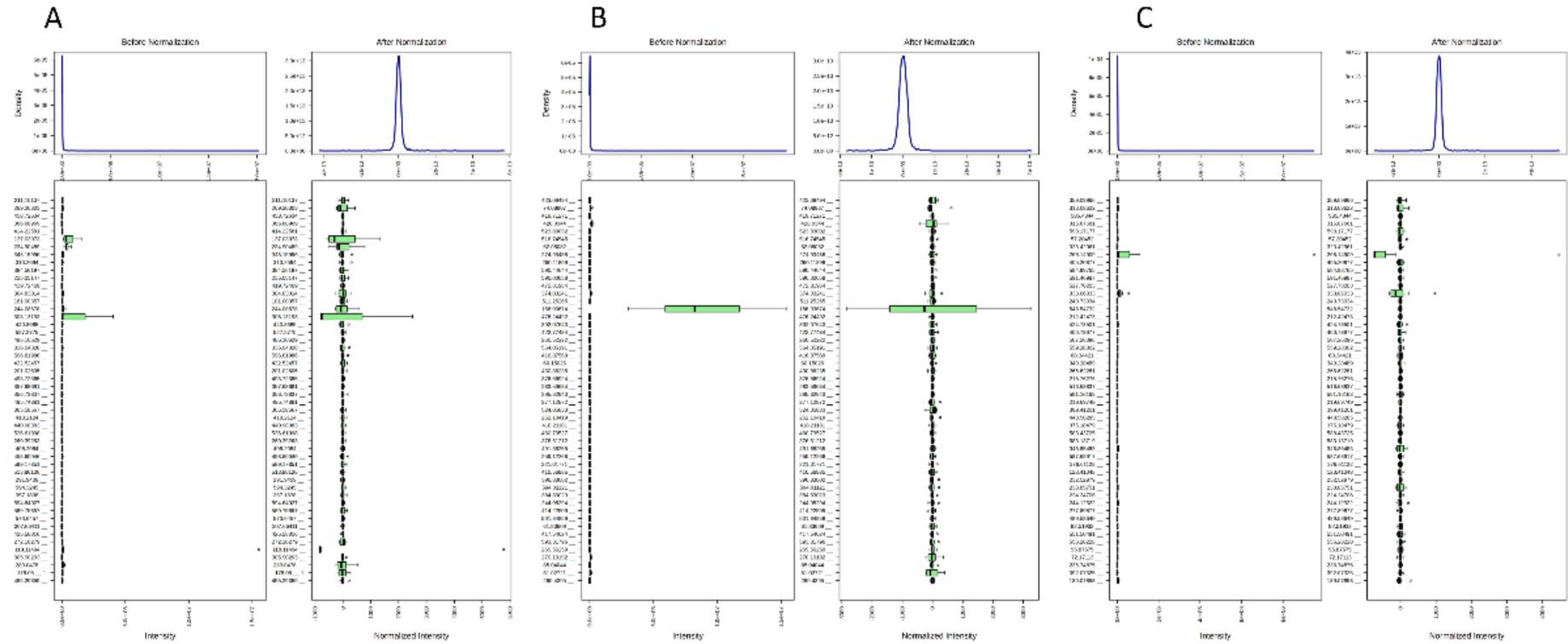


Fig S5.4 Normalisation plots for strain MG1655 for pairwise comparisons of the 3 treatments against the antibiotic free control. A) ciprofloxacin. B) cefotaxime. C) kanamycin. The data was pareto scaled.

S5 – Principle Component Analysis

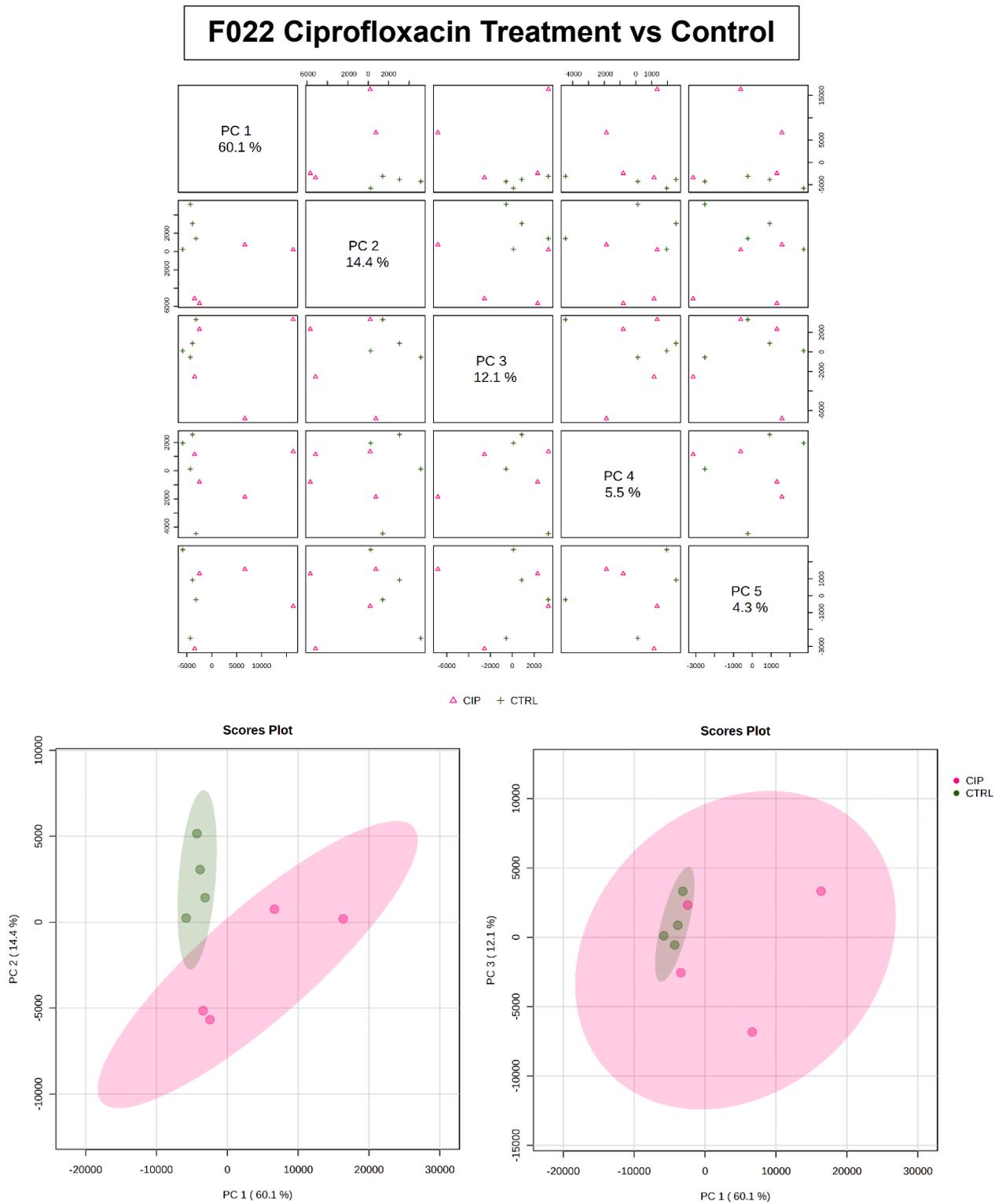


Figure S5.5 A) Principle component analysis (PCA) of strain F022 treated with ciprofloxacin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CIP = ciprofloxacin, in pink. CTRL = control, in green.

F022 Cefotaxime Treatment vs Control

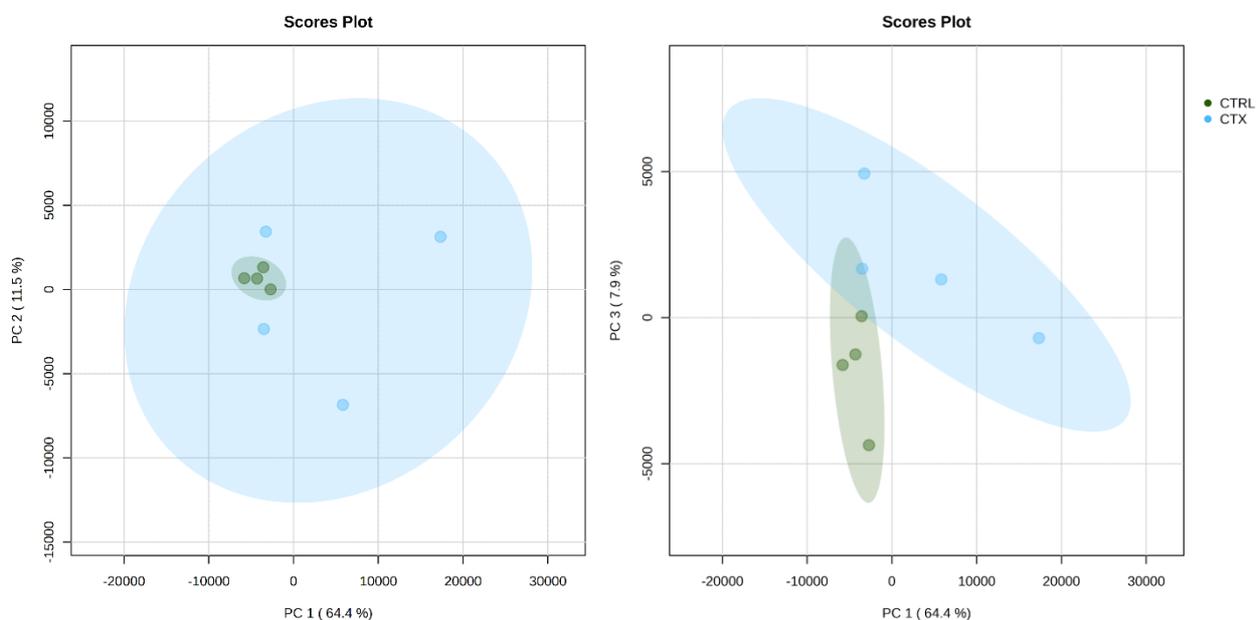
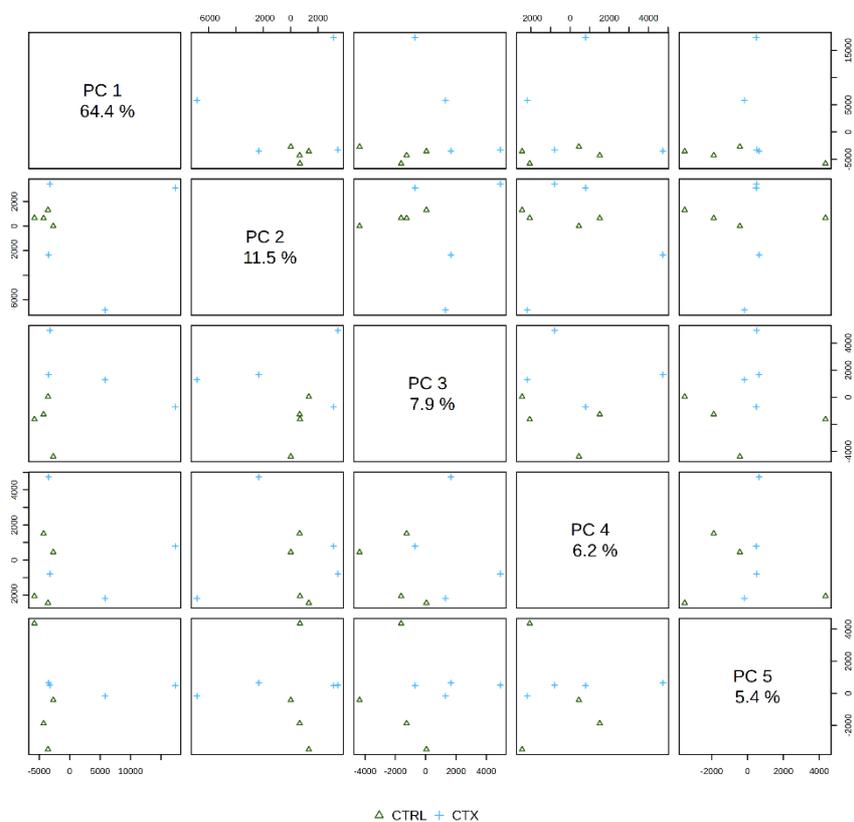


Figure S5.5 B) Principle component analysis (PCA) of strain F022 treated with cefotaxime compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = cefotaxime, in blue. CTRL = control, in green.

F022 Kanamycin Treatment vs Control

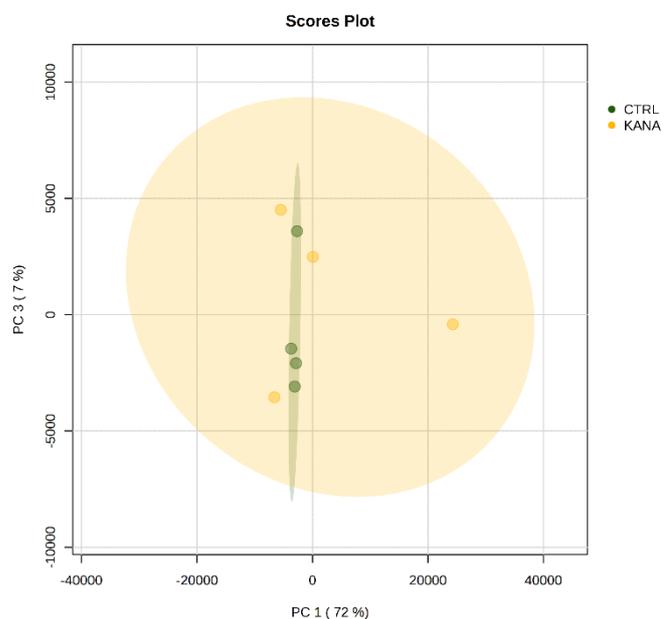
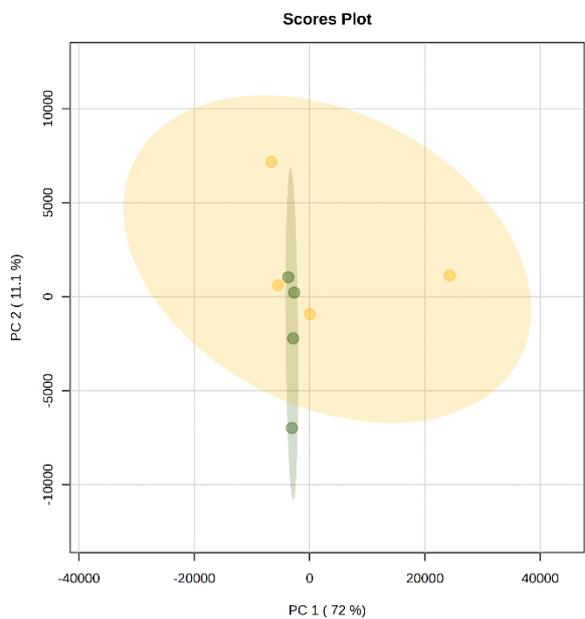
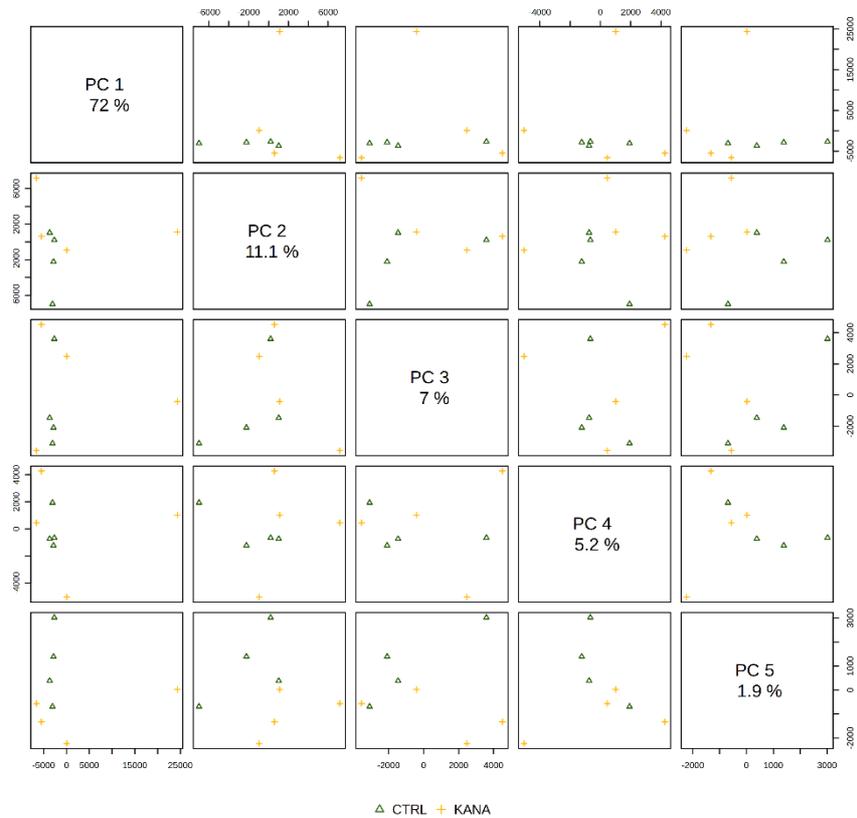


Figure S5.5 C) Principle component analysis (PCA) of strain F022 treated with kanamycin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = kanamycin, in yellow. CTRL = control, in green.

F054 Ciprofloxacin Treatment vs Control

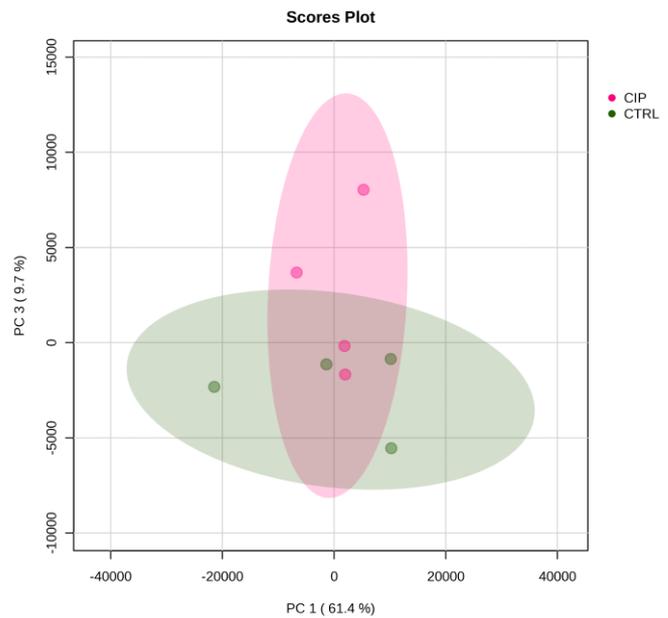
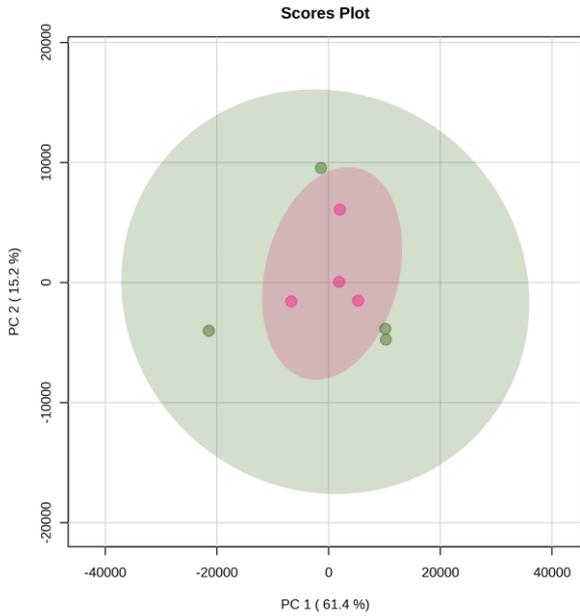
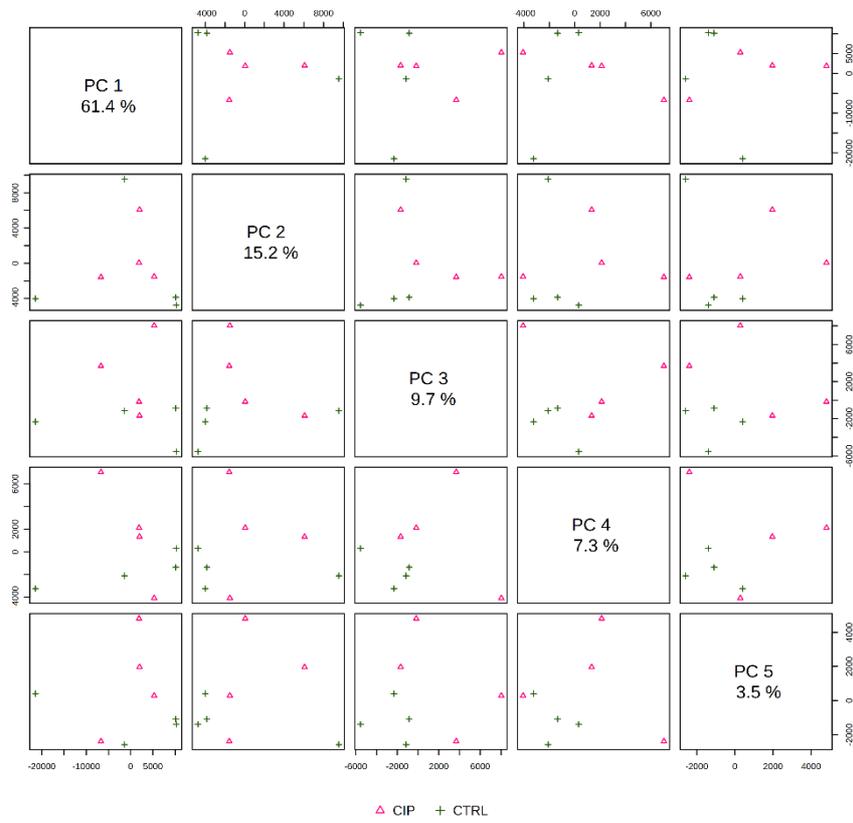
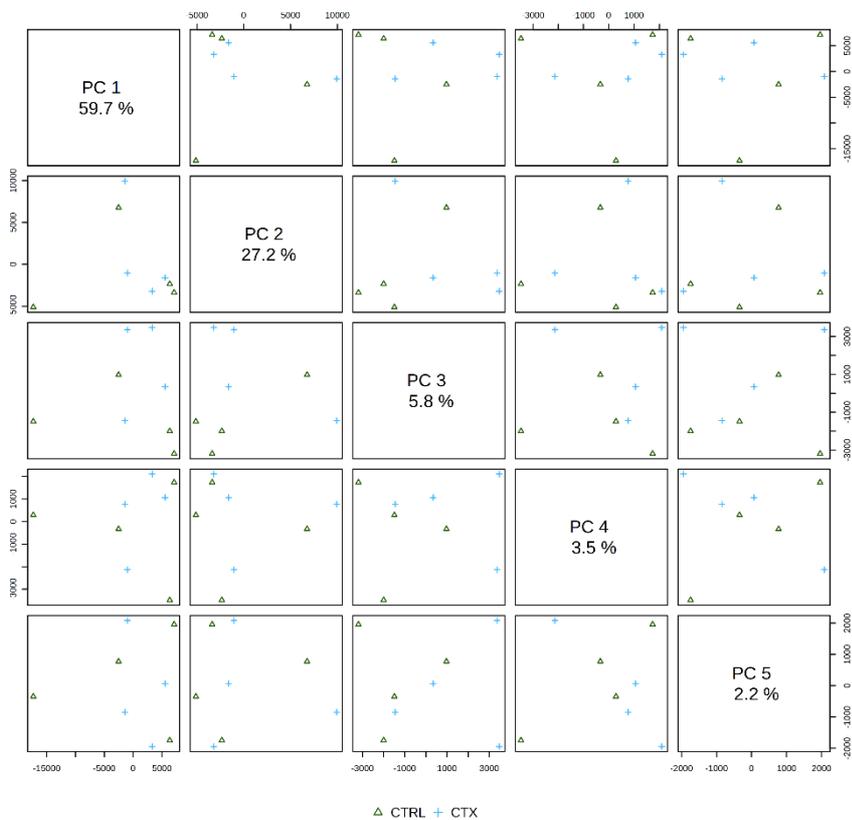


Figure S5.6 A) Principle component analysis (PCA) of strain F054 treated with ciprofloxacin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CIP = ciprofloxacin, in pink. CTRL = control, in green.

F054 Cefotaxime Treatment vs Control



△ CTRL + CTX

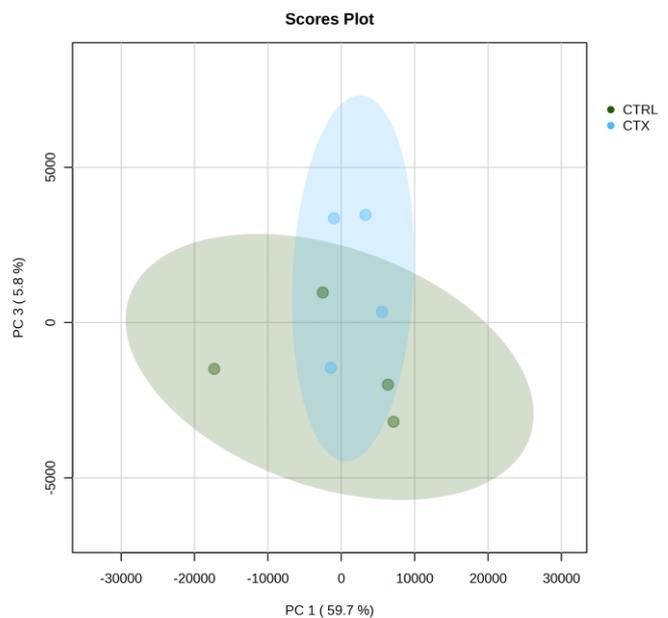
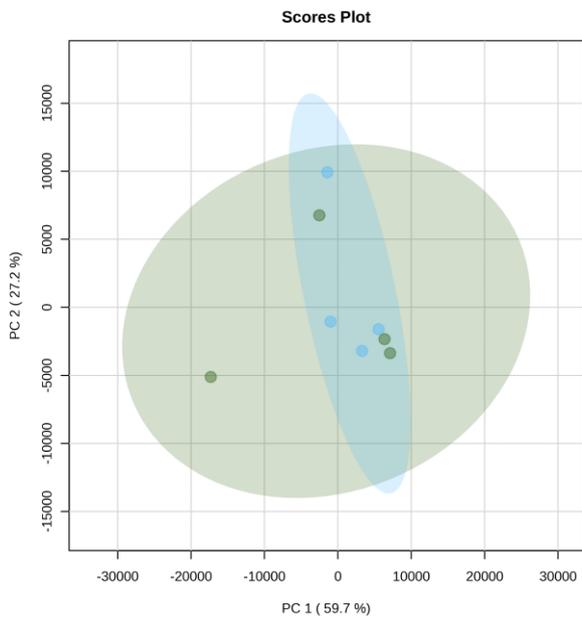


Figure S5.6 B) Principle component analysis (PCA) of strain F054 treated with cefotaxime compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = cefotaxime, in blue. CTRL = control, in green.

F054 Kanamycin Treatment vs Control

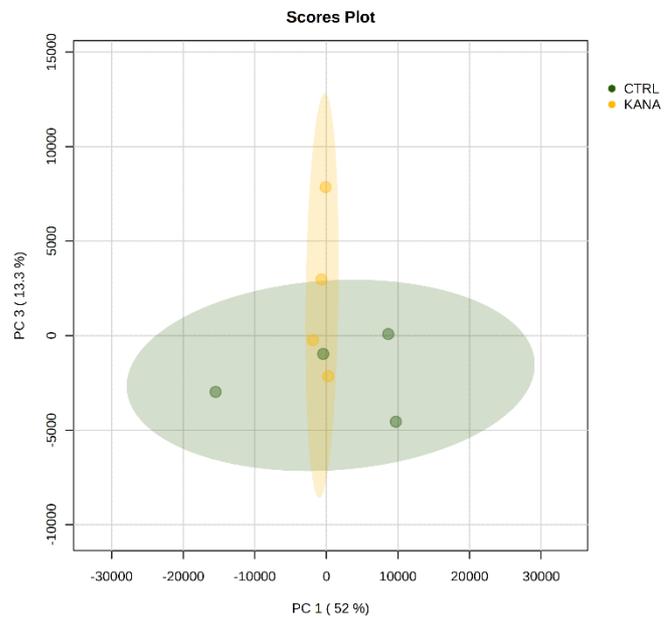
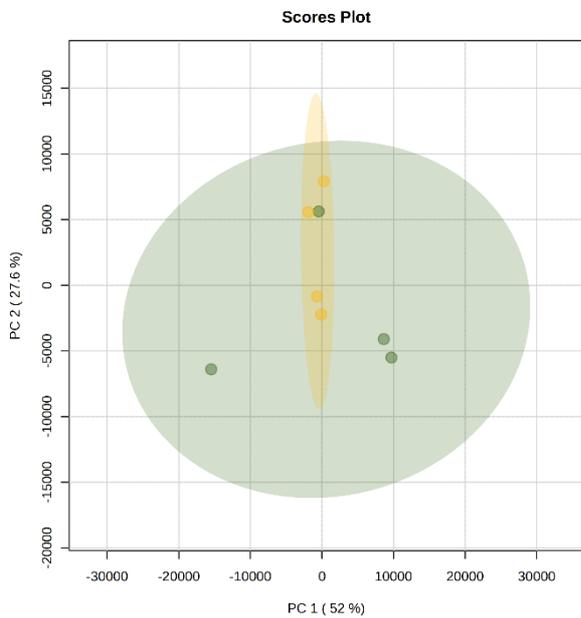
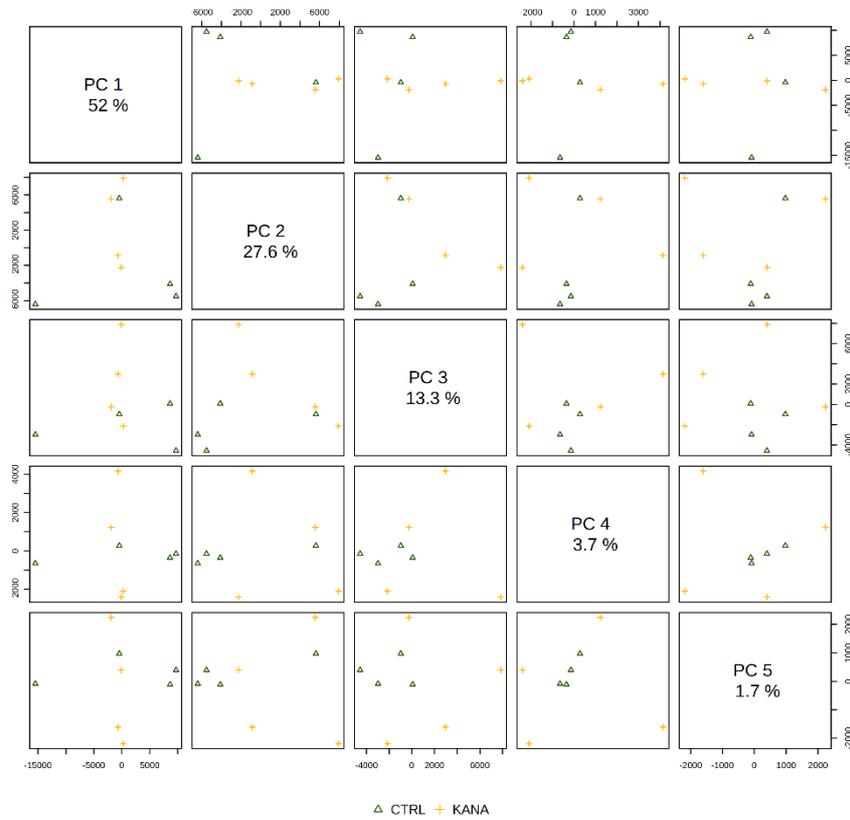


Figure S5.6 C) Principle component analysis (PCA) of strain F054 treated with kanamycin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = kanamycin, in yellow. CTRL = control, in green.

ELU39 Ciprofloxacin Treatment vs Control

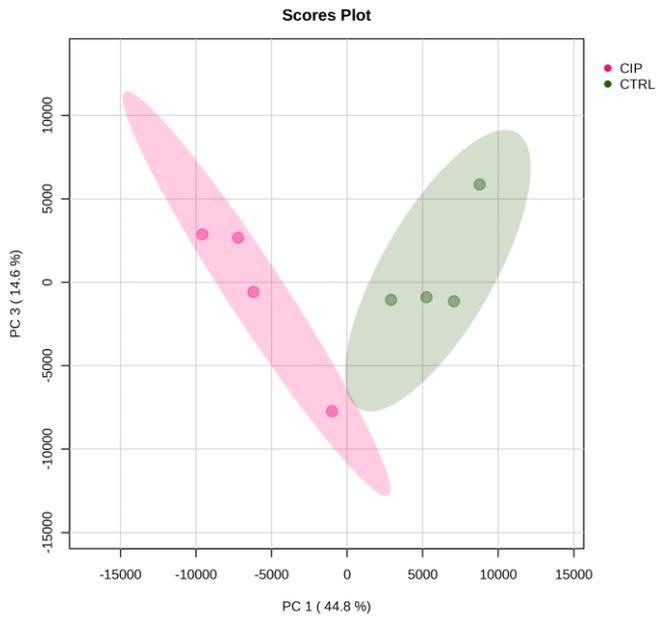
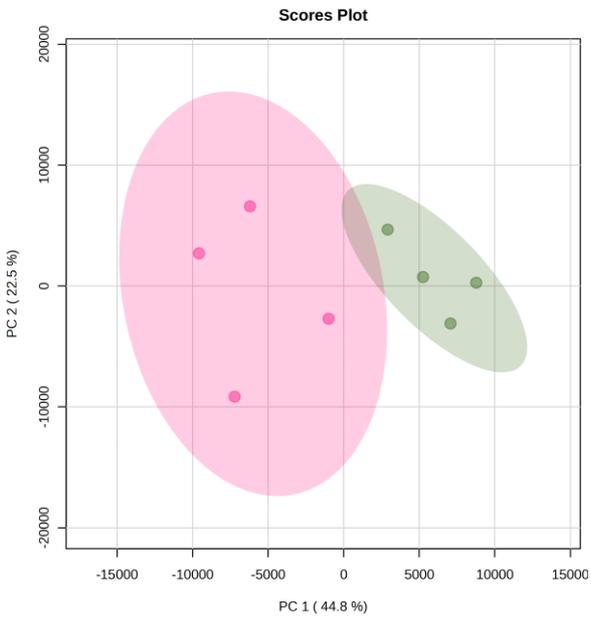
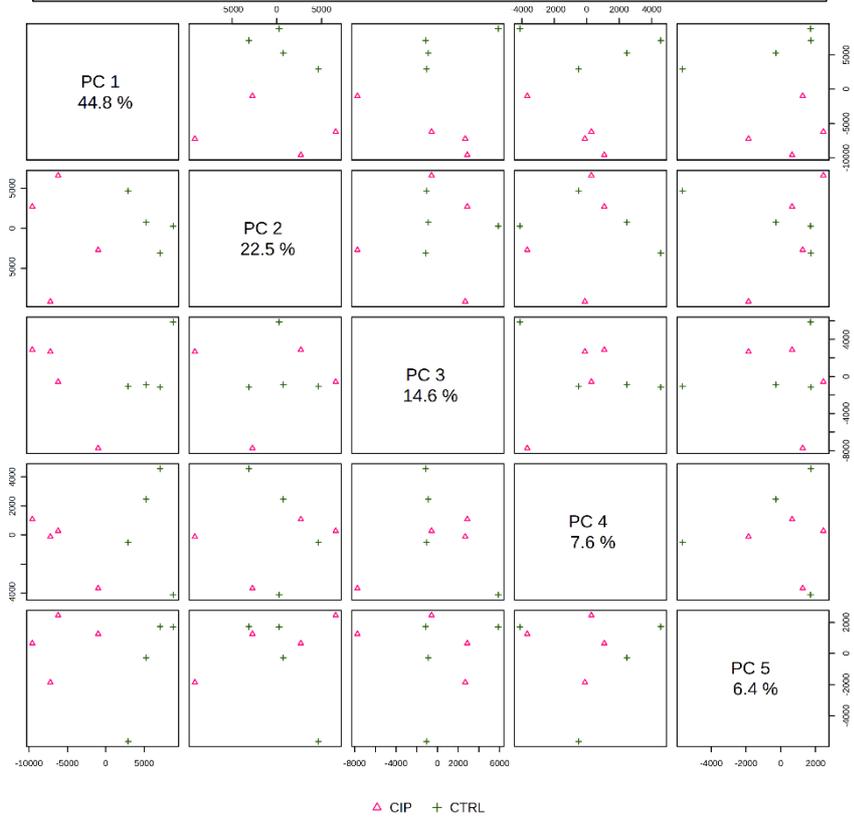


Figure S5.7 A) Principle component analysis (PCA) of strain ELU39 treated with ciprofloxacin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CIP = ciprofloxacin, in pink. CTRL = control, in green.

ELU39 Cefotaxime Treatment vs Control

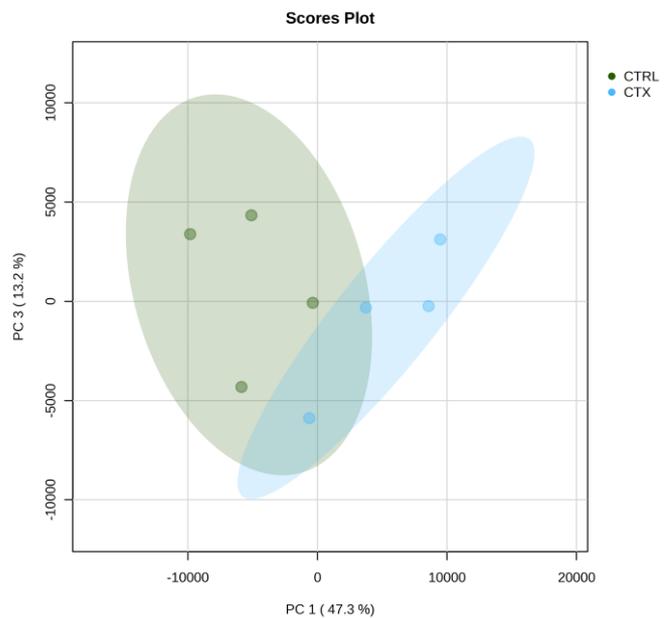
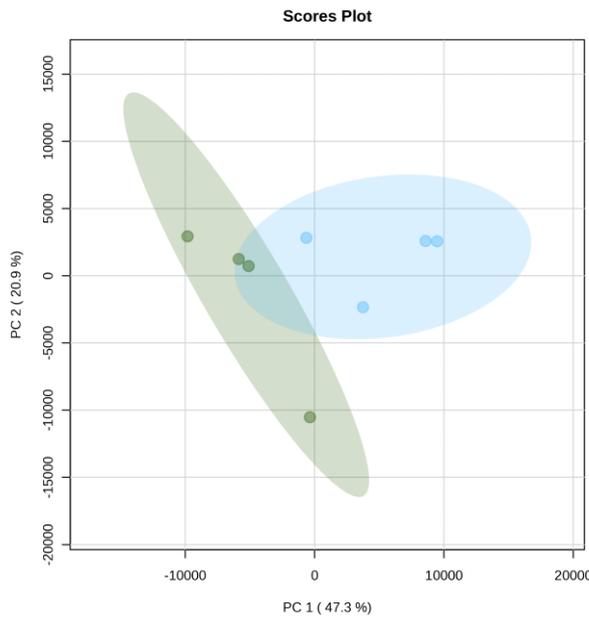
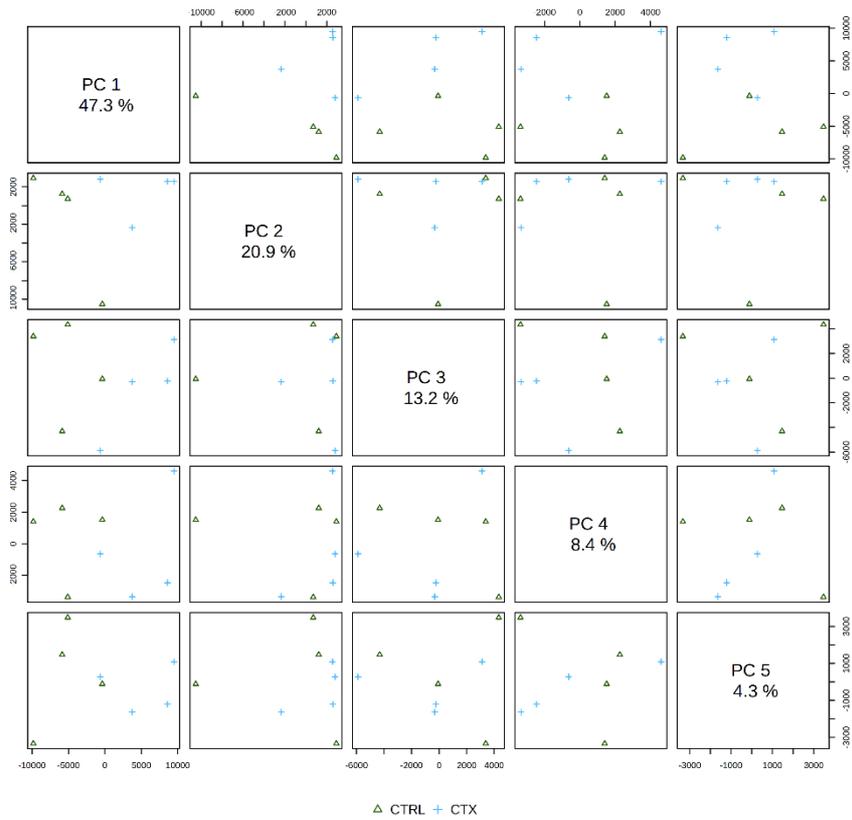


Figure S5.7 B) Principle component analysis (PCA) of strain ELU39 treated with cefotaxime compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = cefotaxime, in blue. CTRL = control, in green.

ELU39 Kanamycin Treatment vs Control

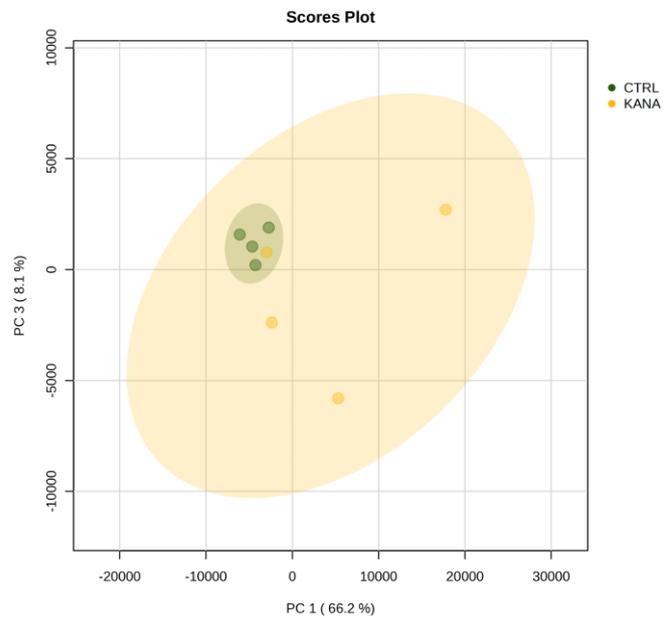
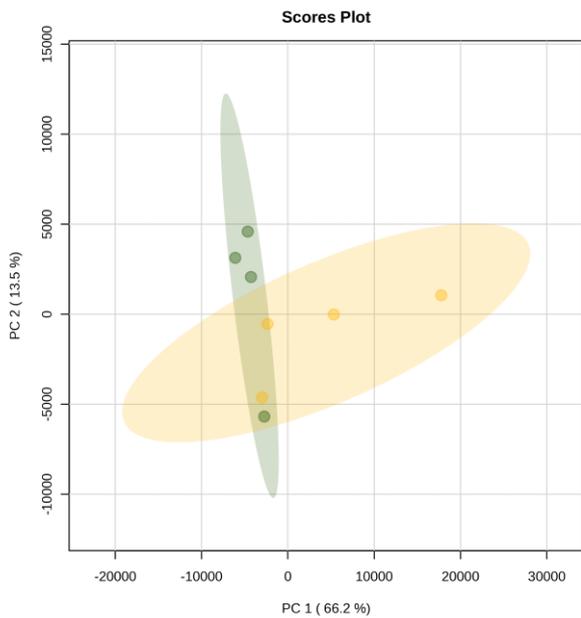
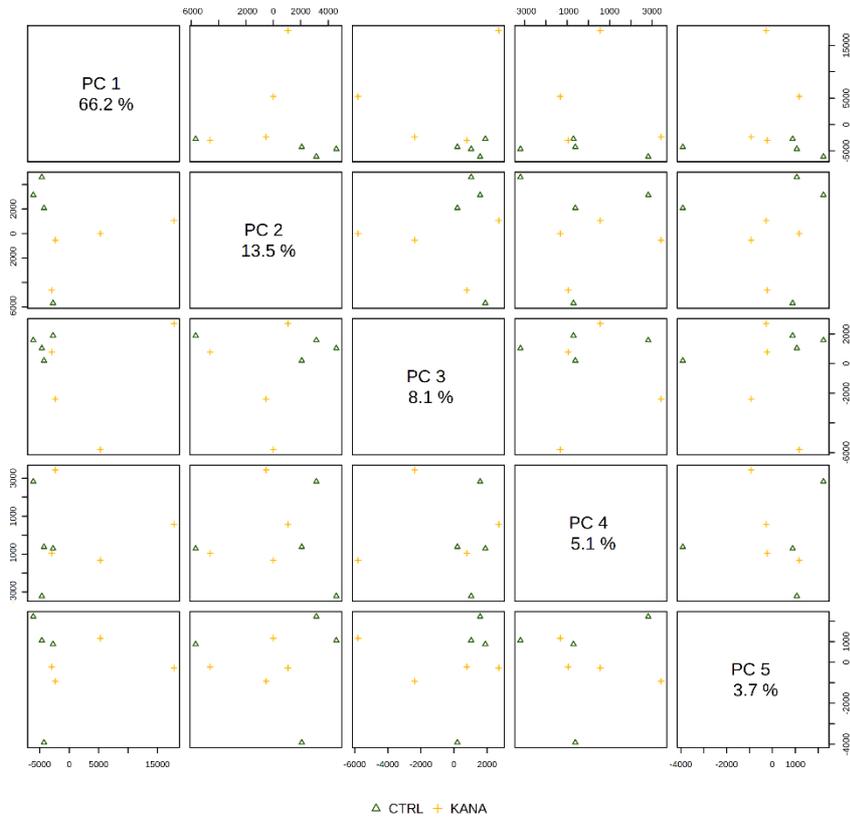


Figure S5.7 C) Principle component analysis (PCA) of strain ELU39 treated with kanamycin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = kanamycin, in yellow. CTRL = control, in green.

MG1655 Ciprofloxacin Treatment vs Control

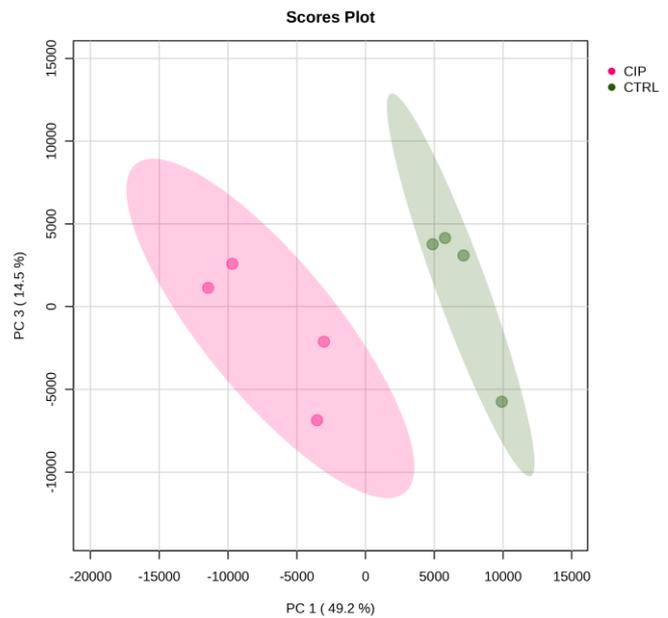
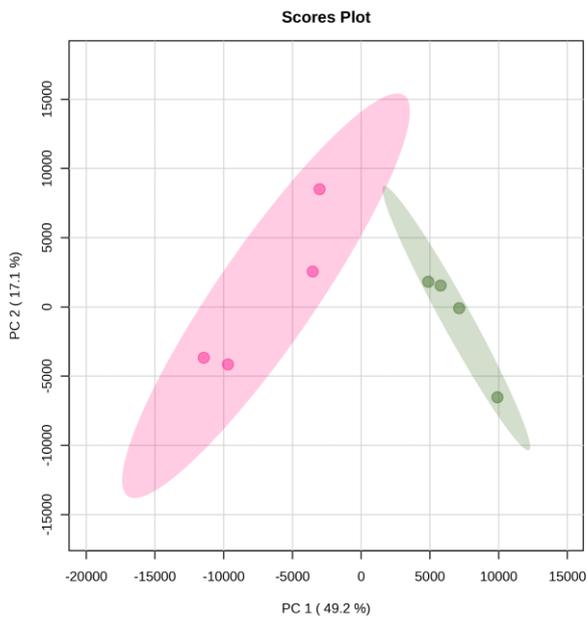
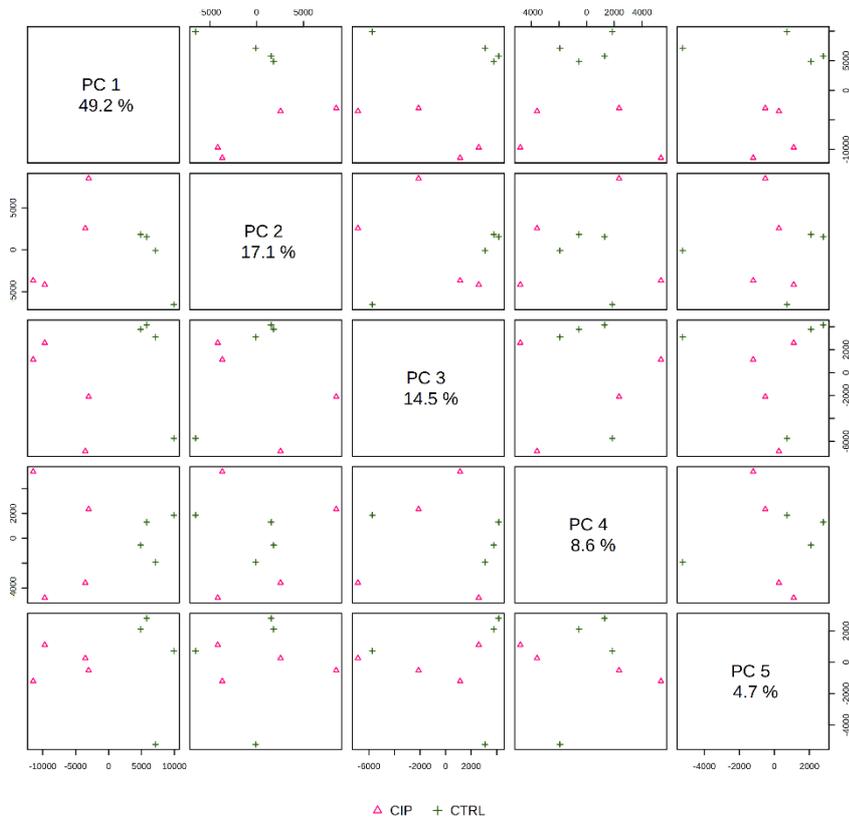


Figure S5.8 A) Principle component analysis (PCA) of strain MG1655 treated with ciprofloxacin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 and PC2 and PC1 vs PC3. CIP = ciprofloxacin, in pink. CTRL = control, in green.

MG1655 Cefotaxime Treatment vs Control

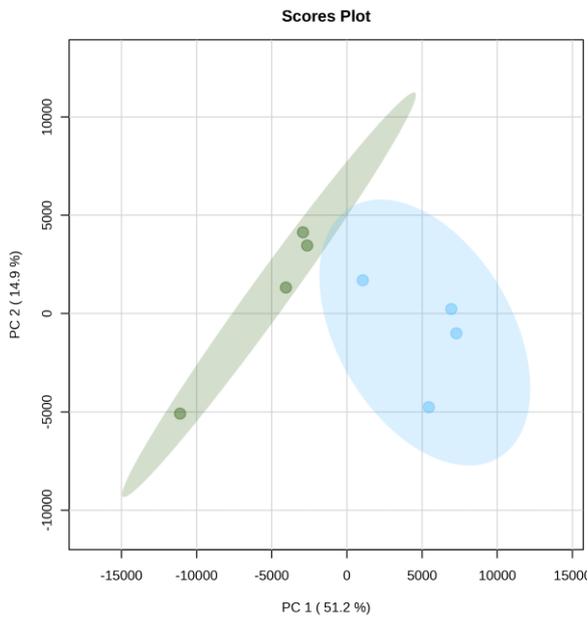
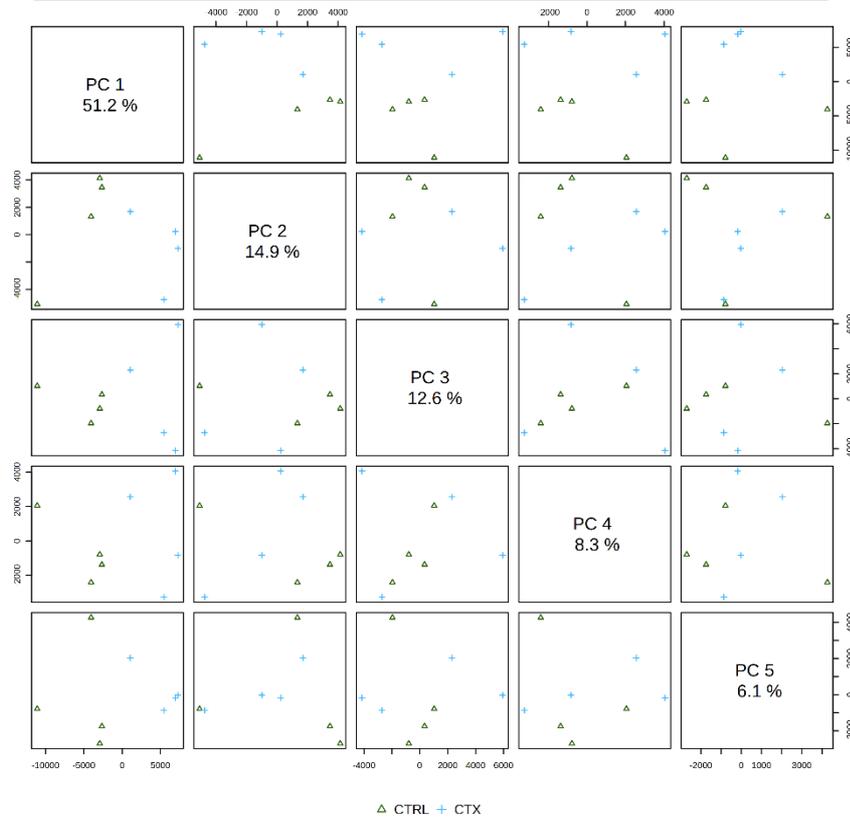


Figure S5.8 B) Principle component analysis (PCA) of strain MG1655 treated with cefotaxime compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = cefotaxime, in blue. CTRL = control, in green.

MG1655 Kanamycin Treatment vs Control

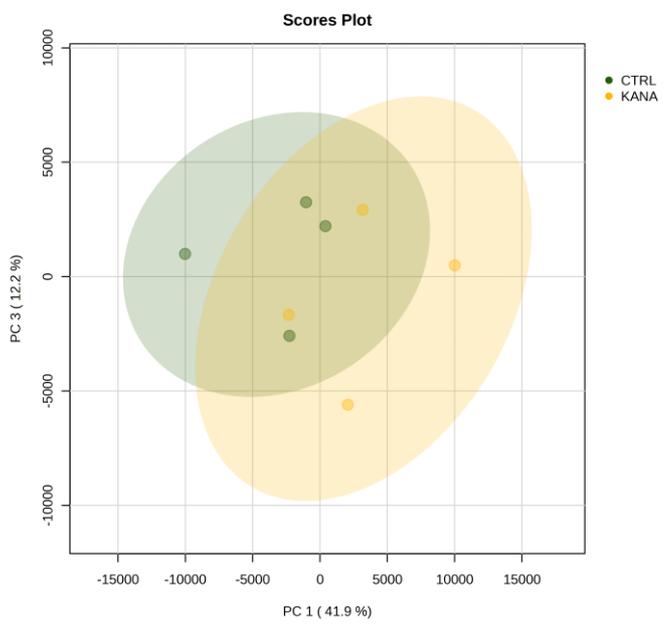
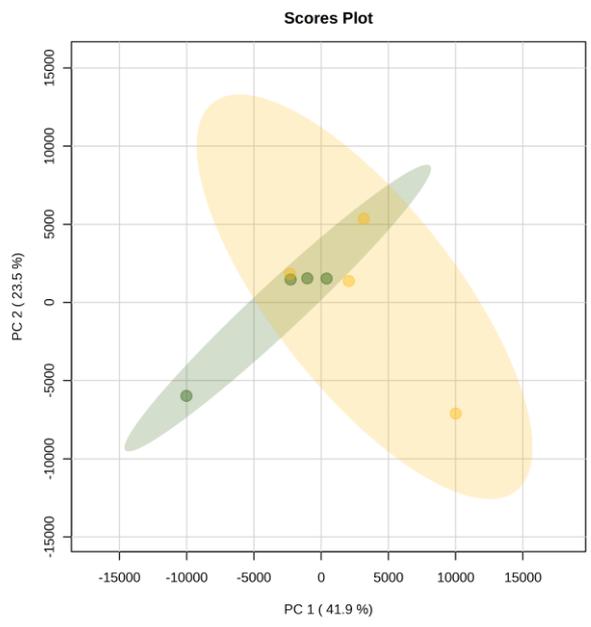
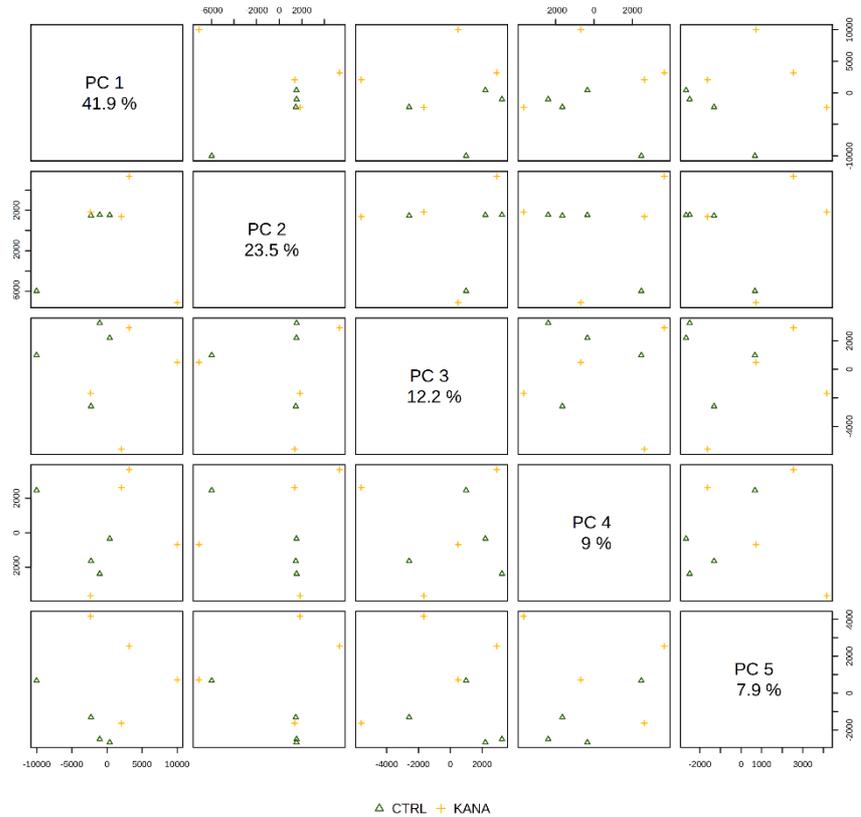


Figure S5.8 C) Principle component analysis (PCA) of strain MG1655 treated with kanamycin compared to the control. Above, loading scores for the first 5 principle components, and below PC1 vs PC2 and PC1 vs PC3. CTX = kanamycin, in yellow. CTRL = control, in green.

S5 – Data Tables

Table S5.1 Identification table of metabolites significantly affected by the kanamycin treatment in the 4 *E.coli* strains. Masses were identified using ECMDB and cross referenced with KEGG metabolism database for *E.coli* K-12 MG1655.

Kanamycin												
Strain	Mass	ID	KEGG	MODULE	ppm	Adduct	Adduct + Compound M/Z	Monoisotopic Mass	Chemical Formula	Fold Change	log2(FC)	Function
MG1655	81.02721	Nitrate	C00244	M00530	28	M+NH4	81.0295	62.9956	HNO3	3.1371	1.6494	Dissimilatory nitrate reduction
ELU39	84.00504	Carbamate	C01563		7	M+NA	84.0056	61.0164	CH3NO2	3.0182	1.5937	Pyrimidine metabolism
F022	98.05503	Trimethylamine N-Oxide	C01104		27	M+Na	98.0576	75.0684	C3H9NO	2.5827	1.3689	Methane metabolism
F022	99.0387	Propylene glycol	C00583		30	M+Na	99.0416	76.0524	C3H8O2	18.923	4.2421	Propanoate metabolism
ELU39	101.00322	Dimethyl sulfoxide	C11143		1	M+Na	101.0032	78.0139	C2H6OS	2.0164	1.0118	Sulfur metabolism
ELU39	112.00041	Oxamate	C01444		1	M+Na	112.0005	89.0113	C2H3NO3	2.1051	1.0739	DEM
F022	113.01904	D-Lactic acid	C00256		17	M+Na	113.0209	90.0317	C3H6O3	2.2133	1.1462	Pyruvate metabolism
		3-Hydroxypropanoate	C01013	M00939	17	M+Na	113.0209	90.0317	C3H6O3			Pyrimidine degradation
		L-Lactic acid	C00186		17	M+Na	113.0209	90.0317	C3H6O3			Pyruvate metabolism
		Glycerone	C00184		17	M+Na	113.0209	90.0317	C3H6O3			Glycerolipid metabolism
ELU39	123.0519	Niacinamide	C00153		28	M+H	123.0553	122.048	C6H6N2O	11.39	3.5097	Precursor to NAD
ELU39	124.03439	Nicotinic acid	C00253	M00115	40	M+H	124.0393	123.032	C6H5NO2	2.0319	1.0228	Precursor to NAD
ELU39	137.04485	Hypoxanthine	C00262	M00958	7	M+H	137.0458	136.0385	C5H4N4O	2.6448	1.4032	Adenine ribonucleotide degradation
F054	142.02567	(S)-b-aminoisobutyric acid	C03284		6	M+K	142.0265	103.0633	C4H9NO2	0.46601	-1.1016	Valine, Leucine, Isoleucine degradation
		gamma-Aminobutyric acid	C00334	M00027 M00136	6	M+K	142.0265	103.0633	C4H9NO2			GABA shunt GABA biosynthesis
ELU39	152.02783	Pyrroline hydroxycarboxylic acid	C04281		26	M+Na	152.0318	129.0426	C5H7NO3	2.1923	1.1325	Arginine and proline metabolism
F054	152.02783	Pyrroline hydroxycarboxylic acid	C04281		26	M+Na	152.0318	129.0426	C5H7NO3	0.16968	-2.5591	Arginine and proline metabolism
F054	152.04936	L-Malic acid	C00149	M00009 M00168	39	M+NH4	152.0553	134.0215	C4H6O5	0.41629	-1.2643	TCA cycle, CAM
		Guanine	C00242	M00959	48	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide degradation
F054	155.07854	L-Ornithine	C00077	M00028 M00844	4	M+Na	155.0791	132.0899	C5H12N2 O2	0.45198	-1.1457	Ornithine biosynthesis arginine biosynthesis
		2-Aminobenzoic acid	C00108	M00023	19	M+NH4	155.0815	137.0477	C7H7NO2			Tryptophan biosynthesis

		p-Aminobenzoic acid	C00568		19	M+NH4	155.0815	137.0477	C7H7NO2			Folate biosynthesis
MG1655	158.95654	3-Mercaptopyruvic acid	C00957		33	M+K	158.9513	119.9881	C3H4O3S	2.2524	1.1714	Cysteine and methionine metabolism
F022	161.08798	D-Alanine	C00993		25	M+H	161.0921	160.0848	C6H12N2O3	4.6502	2.2173	D-amino acid metabolism
F022	162.0701	4-Amino-5-hydroxymethyl-2-methylpyrimidine	C01279		39	M+Na	162.0638	139.0746	C6H9N3O	2.1322	1.0924	Thiamine biosynthesis
MG1655	162.10899	L-Carnitine	C00318		21	M+H	162.1125	161.1052	C7H15NO3	0.44987	-1.1524	
		Carnitine	C00487		25	M+H	162.113	162.113	C7H16NO3			Lysine degradation
F022	164.088	2-Aceto-2-hydroxybutyrate	C06006	M00019 M00570	23	M+NH4	164.0917	146.0579	C6H10O4	2.578	1.3662	Valine, Isoleucine biosynthesis
		2-Dehydropantoate	C00966	M00119	23	M+NH4	164.0917	146.0579	C6H10O4			Pantothenate biosynthesis, Coenzyme A biosynthesis
F022	165.05116	3-Hydroxycinnamic acid	C12621	M00545	21	M+H	165.0546	164.0473	C9H8O3	2.9987	1.5843	Trans-cinamate degradation
		Phenylpyruvic acid	C00166	C00024	21	M+H	165.0546	164.0473	C9H8O3			Phenylalanine biosynthesis
F054	187.10301	Pyridoxine	C00314		25	M+NH4	187.1077	169.0739	C8H11NO3	0.42562	-1.2324	
F022	191.07056	Pyridoxamine	C00534		45	M+Na	191.0791	168.0899	C8H12N2O2	3.4079	1.7689	
ELU39	197.09073	N-Acetylorithine	C00437	M00028	5	M+Na	197.0897	174.1004	C7H14N2O3	3.1626	1.6611	Ornithine metabolism
ELU39	198.08473	Citrulline	C00327	M000844	1	M+Na	198.0849	175.0957	C6H13N3O3	2.4705	1.3048	Arginine biosynthesis
		5-Methylthioribose	C03089		27	M+NH4	198.0795	180.0456	C6H12O4S			Cysteine and methionine metabolism
		Trans-2,3-Dihydroxycinnamate	C12623	M00545	44	M+NH4	198.0761	180.0423	C9H8O4			Trans-cinamate degradation
		4-Hydroxyphenylpyruvic acid	C01179	M00025	44	M+NH4	198.0761	180.0423	C9H8O4			Tyrosine biosynthesis
		3,4-Dihydroxy-L-phenylalanine	C00355		44	M+H	198.0761	197.0688	C9H11NO4			Tyrosine biosynthesis

MG1655	199.0258	5-Dehydro-4-deoxy-D-glucuronate	C04053		23	M+Na	199.0213	176.0321	C6H8O6	0.249	-2.0058	
		(4S)-4,6-Dihydroxy-2,5-dioxohexanoate	C04349		23	M+Na	199.0213	176.0321	C6H8O6			
		Ascorbate	C00072	M00550	23	M+Na	199.0213	176.0321	C6H8O6			Ascorbate degradation
		Ureidosuccinic acid	C00438	M00051	34	M+Na	199.0325	176.0433	C5H8N2O5			de novo pyrimidine biosynthesis
F022	203.04874	L-Rhamnonate	C01934		19	M+Na	203.0526	180.0634	C6H12O6	15.833	3.9848	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		19	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		19	M+Na	203.0526	180.0634	C6H12O6			
		alpha-D-Glucose	C00267	M00001 M000549 M00909	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Fuctose	C00095		19	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar biosynthesis and galactose degradation leloir pathway
		beta-D-glucose	C00221		19	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		19	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		19	M+Na	203.0526	180.0634	C6H12O6			
F022	204.06075	L-Tyrosine	C00082	M00025 M00127	12	M+Na	204.0631	181.0739	C9H11NO3	15.088	3.9153	Tyrosine biosynthesis, thiamine biosynthesis
ELU39	215.09645	3,4-Dihydroxy-L-phenylalanine	C00355		29	M+NH4	215.1026	197.0688	C9H11NO4	2.11	1.0772	Tyrosine biosynthesis
F054	226.94615	2,5-Dichloro-4-oxohex-2-enedioate	C12835		21	M+H	226.9509	225.9436	C6H4Cl2O5	0.23421	-2.0941	
ELU39	233.10867	4-(Glutamylamino)butanoate	C16757	M00136	19	M+H	233.1132	232.1059	C9H16N2O5	2.4439	1.2892	GABA biosynthesis
		N2-Succinyl-L-ornithine	C03415	M00879	19	M+H	233.1132	232.1059	C9H16N2			Arginine

								O5				succinyltransferase pathway
F022	243.09913	Thymidine	C00214		7	M+H	243.0975	242.0903	C10H14N2O5	2.1891	1.1303	Pyrimidine metabolism
F054	244.12412	Porphobilinogen	C00931		21	M+NH4	244.1292	226.0954	C10H14N2O4	0.4843	-1.046	Porphyrin precursor, heme biosynthesis
F022	245.05527	L-Cystathionine	C02291	M00017	6	M+Na	245.0566	222.0674	C7H14N2O4S	6.3849	2.6747	Methionine biosynthesis
		L-Glutamic acid 5-phosphate	C03287	M00015	8	M+NH4	245.0533	227.0195	C5H10NO7P			Proline biosynthesis
ELU39	258.0697	Pantothenic acid	C00864	M00019 M00120	16	M+K	258.0738	219.1107	C9H17NO5	2.16	1.111	Pantothenate biosynthesis, Coenzyme A biosynthesis
		L-Cystine	C00491		47	M+NH4	258.0577	240.0238	C6H12N2O4S2			Cysteine and methionine metabolism
F054	261.05267	3-Deoxy-D-manno-octulosonate	C01187	M00063	21	M+Na	261.0581	238.0689	C8H14O8	0.28609	-1.8054	CMP-KDO biosynthesis
F022	261.05488	3-Deoxy-D-manno-octulosonate	C01187	M00063	12	M+Na	261.0581	238.0689	C8H14O8	2.5994	1.3782	CMP-KDO biosynthesis
F022	261.12267	Cytidine	C00475		13	M+NH4	261.1193	243.0855	C9H13N3O5	9.8906	3.3061	Precursor of cytosine
MG1655	268.14909	Ubiquinone	C00399	M00117	20	M+NH4	268.1543	250.1205	C14H18O4	56.008	5.8076	Oxidative phosphorylation
F054	270.0579	N-Succinyl-L-glutamate	C05931	M00879	2	M+Na	270.0584	247.0692	C9H13NO7	0.10311	-3.2778	Arginine succinyltransferase pathway
ELU39	275.12967	Ubiquinol-1	C00390	M00117	16	M+Na	275.1254	252.1362	14H20O4	2.4115	1.2699	Oxidative phosphorylation
		N2-Succinyl-L-arginine	C03296	M00879	19	M+H	275.135	274.1277	C10H18N4O5			Arginine succinyltransferase pathway
F022	282.03422	D-Galactosamine 6-phosphate	C06377		2	M+Na	282.0349	259.0457	C6H14NO8P	2.2943	1.1981	Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		2	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501		2	M+Na	282.0349	259.0457	C6H14NO			UDP-N-acetyl-D-

								8P				glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352		2	M+Na	282.0349	259.0457	C6H14NO 8P			UDP-N-acetyl-D-glucosamine biosynthesis
F022	306.063	Deoxyguanosine	C00330		10	M+K	306.0599	267.0968	C10H13N 5O4	2.9246	1.5482	Purine metabolism
		Adenosine	C00212	M00958	10	M+K	306.0599	267.0968	C10H13N 5O4			Adenine ribonucleotide degradation, purine metabolism
		D-Glycero-D-manno-heptose 1-phosphate	C07838	M00064	15	M+NH4	306.0585	288.0246	C7H13O1 0P			ADP-L-glycero-D-manno-heptose biosynthesis
		2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate	C04691	M00022	15	M+NH4	306.0585	288.0246	C7H13O1 0P			Shikimate pathway
		2',3'-Cyclic CMP	C02354		47	M+H	306.0486	305.0413	C9H12N3 07P			Pyrimidine metabolism
ELU39	308.13422	N-Succinyl-L,L-2,6-diaminopimelate	C04421	M00016	36	M+NH4	308.1452	290.1114	C11H18N2 07	0.046116	-4.4386	Lysine biosynthesis
ELU39	310.12488	N-acetylneuraminic acid	C00270		37	M+H	310.1133	309.106	C11H19N 09	2.3382	1.2254	Nucleotide sugar metabolism
F054	314.08564	Phosphoribosylformylglycinamide	C04640	M00048	35	M+H	314.0748	313.0675	C8H16N3 08P	2.7652	1.4674	De novo purine biosynthesis
F022	332.0679	N-Acetyl-D-muramic acid	C02713		19	M+K	332.0742	293.1111	C11H19N 08	565.44	9.1432	Precursor for peptidoglycan biosynthesis
		dAMP	C00360		23	M+H	332.0754	331.0682	C10H14N 5O6P			ADP derivative
ELU39	348.06578	3'-AMP	C01367		13	M+H	348.0704	347.0631	C10H14N 5O7P	2.0801	1.0567	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'-monophosphate dGMP	C00362		13	M+H	348.0704	347.0631	C10H14N 5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	13	M+H	348.0704	347.0631	C10H14N 5O7P			Adenine ribonucleotide biosynthesis and degradation

MG1655	354.04634	dAMP	C00360		31	M+Na	354.0574	331.0682	C10H14N 5O6P	4.0476	2.0171	ADP derivative
F054	354.049	dAMP	C00360		24	M+Na	354.0574	331.0682	C10H14N 5O6P	0.44358	-1.1728	ADP derivative
F022	354.04957	dAMP	C00360		22	M+Na	354.0574	331.0682	C10H14N 5O6P	2.2074	1.1423	ADP derivative
F022	370.05309	3'-AMP	C01367		2	M+Na	370.0523	347.0631	C10H14N 5O7P	3.015	1.5922	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'- monophosphate dGMP	C0362		2	M+Na	370.0523	347.0631	C10H14N 5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	2	M+Na	370.0523	347.0631	C10H14N 5O7P			Adenine ribonucleotide biosynthesis and degradation
MG1655	423.09397	Trehalose 6-phosphate	C00689		10	M+H	423.0898	422.0825	C12H23O 14P	0.28785	-1.7966	Other carbohydrate metabolism
		S- Adenosylhomocysteine	C00021		22	M+K	423.0847	384.1216	C14H20N 6O5S			Methionine degradation
MG1655	445.07639	Trehalose 6-phosphate	C00689		10	M+Na	445.0718	422.0825	C12H23O 14P	0.36537	-1.4526	Other carbohydrate metabolism
		dGDP	C00361	M00053	30	M+NH4	445.0632	427.0294	C10H15N 5O10P2			Deoxyribonucleotide biosynthesis
		ADP	C00008	M00053	30	M+NH4	445.0632	427.0294	C10H15N 5O10P2			Deoxyribonucleotide biosynthesis

Table S5.2 Identification table of metabolites significantly affected by the cefotaxime treatment in the 4 *E.coli* strains. Masses were identified using ECMDB and cross referenced with KEGG metabolism database for *E.coli* K-12 MG1655.

Cefotaxime												
Strain	Mass	ID	KEGG	MODULE	ppm	Adduct	Adduct + Compound M/Z	Monoisotopic Mass	Chemical Formula	Fold Change	log2(FC)	Function
MG1655	81.02721	Nitrate	C0244	M00530	28	M+NH4	81.0295	62.9956	HNO3	2.7669	1.4683	Dissimilatory nitrate reduction
F022	98.05503	Trimethylamine N-Oxide	C01104		27	M+Na	98.0576	75.0684	C3H9NO	2.4716	1.3054	Methane metabolism
F022	99.0387	Propylene glycol	C00583		30	M+Na	99.0416	76.0524	C3H8O2	27.171	4.764	
ELU39	122.02266	D-Cysteine	C00793		36	M+H	122.027	121.0197	C3H7NO2S	0.44027	-1.1835	
		L-Cysteine	C00097	M00021	36	M+H	122.027	121.0197	C3H7NO2S			Cysteine biosynthesis
F054	122.02266	D-Cysteine	C00793		36	M+H	122.027	121.0197	C3H7NO2S	0.37352	-1.4207	
		L-Cysteine	C00097	M00021	36	M+H	122.027	121.0197	C3H7NO2S			Cysteine biosynthesis
ELU39	123.0519	Nicotinamide	C00153		28	M+H	123.0553	122.048	C6H6N2O	0.41856	-1.2565	Precursor to NAD
ELU39	135.02364	(R)-Malate	C00497		38	M+H	135.0288	134.0215	C4H6O5	0.44063	-1.1824	
		L-Malic acid	C00149	M00009 M00168	38	M+H	135.0288	134.0215	C4H6O5			TCA cycle, CAM
MG1655	135.02364	(R)-Malate	C00497		38	M+H	135.0288	134.0215	C4H6O5	0.47397	-1.0771	
		L-Malic acid	C00149	M00009 M00168	38	M+H	135.0288	134.0215	C4H6O5			TCA cycle, CAM
F054	140.06606	L-Valine	C00183	M00019	15	M+Na	140.0682	117.079	C5H11NO2	0.4831	-1.0496	Branched chain amino acid metabolism valine-isoleucine metabolism
		Betaine	C00719	M00555	19	M+Na	140.0687	118.0868	C5H12NO2			Serine and threonine biosynthesis
F054	141.0649	Nicotinic acid	C00253	M00115	7	M+NH4	141.0659	123.032	C6H5NO2	0.45893	-1.1236	Precursor to NAD
ELU39	145.12727	Caprylic acid	C06423		34	M+H	145.1223	144.115	C8H16O2	0.48219	-1.0523	Fatty acid biosynthesis
MG1655	152.02783	Pyrroline hydroxycarboxylic acid	C04281		26	M+Na	152.0318	129.0426	C5H7NO3	0.46901	-1.0923	Arginine and proline metabolism
F022	152.04936	(R)-Malate	C00497		39	M+NH4	152.0553	134.0215	C4H6O5	2.1557	1.1081	
		L-Malic acid	C00149	M00009 M00168	39	M+NH4	152.0553	134.0215	C4H6O5			TCA cycle, CAM
		Guanine	C00242	M00959	48	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide

												degradation
F054	152.04936	(R)-Malate	C00497		39	M+NH4	152.0553	134.0215	C4H6O5	0.455	-1.1361	
		L-Malic acid	C00149	M00009 M00168	39	M+NH4	152.0553	134.0215	C4H6O5			TCA cycle, CAM
		Guanine	C00242	M00959	48	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide degradation
MG1655	152.04936	(R)-Malate	C00497		39	M+NH4	152.0553	134.0215	C4H6O5	0.39755	-1.3308	
		L-Malic acid	C00149	M00009 M00168	39	M+NH4	152.0553	134.0215	C4H6O5			TCA cycle, CAM
		Guanine	C00242	M00959	48	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide degradation
ELU39	152.05475	(R)-Malate	C00497		4	M+NH4	152.0553	134.0215	C4H6O5	0.31948	-1.6462	
		L-Malic acid	C00149	M00009 M00168	4	M+NH4	152.0553	134.0215	C4H6O5			TCA cycle, CAM
		Guanine	C00242	M00959	13	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide degradation
F022	152.98198	3-Sulfinylpyruvic acid	C05527		21	M+H	152.9852	151.9779	C3H4O5S	2.1627	1.1128	Cysteine and methionine metabolism
F022	155.0731	L-Ornithine	C00077	M00028 M00844	39	M+Na	155.0791	132.0899	C5H12N2O2	2.1584	1.1099	Ornithine biosynthesis arginine biosynthesis
ELU39	156.03674	Ureidoglycine	C02091		8	M+Na	156.038	133.0487	C3H7N3O3	0.49383	-1.0179	Purine metabolism
		L-Valine	C00183	M00019	35	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism valine-isoleucine metabolism
		Betaine	C00719	M00555	38	M+K	156.0427	118.0868	C5H12NO2			Serine and threonine biosynthesis
MG1655	156.03674	Ureidoglycine	C02091		8	M+Na	156.038	133.0487	C3H7N3O3	0.45925	-1.1226	Purine metabolism
		L-Valine	C00183	M00019	35	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism valine-isoleucine metabolism
		Betaine	C00719	M00555	38	M+K	156.0427	118.0868	C5H12NO2			Serine and threonine biosynthesis
F054	157.04164	(R)-2,3-Dihydroxyisovalerate	C04272	M00019	35	M+Na	157.0471	134.0579	C5H10O4	0.40286	-1.3116	Branched chain amino acid metabolism valine-isoleucine metabolism
		Deoxyribose	C01801		35	M+Na	157.0471	134.0579	C5H10O4			Carbohydrate, pentose phosphate supply

F054	158.02782	L-Homocysteine	C00155	M00017	20	M+Na	158.0246	135.0354	C4H9NO2S	0.49186	-1.0237	Methionine biosynthesis
		L-Allothreonine	C05519		41	M+K	158.0214	119.0582	C4H9NO3			Glycine, serine threonine metabolism
		L-Homoserine	C00263	M00017 M00018	41	M+K	158.0214	119.0582	C4H9NO3			Methionine biosynthesis, threonine biosynthesis
		Acetylphosphate	C00227	M00759	41	M+NH4	158.0213	139.9875	C2H5O5P			Phosphate acetyltransferase-acetate kinase pathway
		L-Threonine	C00188	M00018	41	M+K	158.0214	119.0582	C4H9NO3			Threonine biosynthesis
ELU39	158.95654	3-Mercaptopyruvic acid	C00957		33	M+K	158.9513	119.9881	C3H4O3S	3.1552	1.6578	Cysteine and methionine metabolism
MG1655	158.95654	3-Mercaptopyruvic acid	C00957		33	M+K	158.9513	119.9881	C3H4O3S	2.1357	1.0947	Cysteine and methionine metabolism
ELU39	162.01451	Pimeloyl-[acyl-carrier protein]	C19845	M00123 M00572	10	M+Na	162.0162	139.0269	C6H5NO3	0.48498	-1.044	Biotin biosynthesis, Pimeloyl-ACP biosynthesis, BioC-BioH pathway
F054	162.01451	Pimeloyl-[acyl-carrier protein]	C19845	M00123 M00572	10	M+Na	162.0162	139.0269	C6H5NO3	0.4813	-1.055	Biotin biosynthesis, Pimeloyl-ACP biosynthesis, BioC-BioH pathway
MG1655	162.01451	Pimeloyl-[acyl-carrier protein]	C19845	M00123 M00572	10	M+Na	162.0162	139.0269	C6H5NO3	0.48857	-1.0334	Biotin biosynthesis, Pimeloyl-ACP biosynthesis, BioC-BioH pathway
MG1655	162.10899	L-Carnitine	C00318		21	M+H	162.1125	161.1052	C7H15NO3	0.47731	-1.067	
		Carnitine	C00487		25	M+H	162.113	162.113	C7H16NO3			Lysine degradation
F022	170.03444	D-Glutamic acid	C00217		47	M+Na	170.0424	147.0532	C5H9NO4	2.6711	1.4174	
		O-Acetylserine	C00979	C00021	47	M+Na	170.0424	147.0532	C5H9NO4			Cysteine biosynthesis
		L-Glutamate	C00025	M00015, M00027	47	M+Na	170.0424	147.0532	C5H9NO4			Arginine and proline metabolism, GABA shunt
F022	176.05777	7-Cyano-7-carbaquanine	C15996		6	M+H	176.0567	175.0494	C7H5N5O	2.0323	1.0231	Folate biosynthesis
		2-Maleylacetate	C02222		14	M+NH4	176.0553	158.0215	C6H6O5			Degradation of aromatic compounds
F054	179.04156	L-Histidinal	C01929	M00026	22	M+K	179.0455	140.0824	C6H10N3O	0.49326	-1.0196	Histidine biosynthesis

		Cysteinylglycine	C01419		39	M+H	179.0485	178.0412	C5H10N2O3 S			Glutathione biosynthesis
MG1655	179.04156	L-Histidinal	C01929	M00026	22	M+K	179.0455	140.0824	C6H10N3O	0.49741	-1.0075	Histidine biosynthesis
		Cysteinylglycine	C01419		39	M+H	179.0485	178.0412	C5H10N2O3 S			Glutathione biosynthesis
F054	180.04759	D-4-Hydroxy-2-oxoglutarate	C05946		15	M+NH4	180.0503	162.0164	C5H6O6	0.46187	-1.1144	Arginine and proline Metabolism
		4-Hydroxy-2-oxoglutaric acid	C01127		15	M+NH4	180.0503	162.0164	C5H6O6			Glyoxylate and dicarboxylate metabolism
		L-Histidinol	C00860	M00026	32	M+K	180.0534	141.0902	C6H11N3O			Histidine biosynthesis
F054	188.06555	D-Phenylalanine	C02265		14	M+Na	188.0682	165.079	C9H11NO2	0.4438	-1.172	
		L-Phenylalanine	C00079	M00024	14	M+Na	188.0682	165.079	C9H11NO2			Phenylalanine biosynthesis
F022	189.07858	Tetrahydrodipicolinate	C03972	M00016	44	M+NH4	189.087	171.0532	C7H9NO4	2.0359	1.0257	Lysine biosynthesis
F022	190.00407	Quinolate	C03722	M00115	37	M+Na	190.0111	167.0219	C7H5NO4	2.1427	1.0995	NAD biosynthesis
		Guanine	C00242	M00959	45	M+K	190.0126	151.0494	C5H5N5O			Guanine ribonucleotide degradation
MG1655	199.0258	5-Dehydro-4-deoxy-D-glucuronate	C04053		23	M+Na	199.0213	176.0321	C6H8O6	0.37459	-1.4166	
		(4S)-4,6-Dihydroxy-2,5-dioxohexanoate	C04349		23	M+Na	199.0213	176.0321	C6H8O6			
		Ascorbate	C00072	M00550	23	M+Na	199.0213	176.0321	C6H8O6			Ascorbate degradation
		Ureidosuccinic acid	C00483	M00051	34	M+Na	199.0325	176.0433	C5H8N2O5			De novo pyrimidine biosynthesis
F022	203.04874	L-Rhamnonate	C01934		19	M+Na	203.0526	180.0634	C6H12O6	2.7621	1.4658	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		19	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		19	M+Na	203.0526	180.0634	C6H12O6			
		alpha-D-Glucose	C00267	M00001 M000549 M000909	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Fructose	C00095		19	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar

												biosynthesis and galactose degradation leloir pathway
		beta-D-glucose	C00221		19	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		19	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		19	M+Na	203.0526	180.0634	C6H12O6			
MG1655	203.04874	L-Rhamnonate	C01934		19	M+Na	203.0526	180.0634	C6H12O6	2.3044	1.2044	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		19	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		19	M+Na	203.0526	180.0634	C6H12O6			
		alpha-D-Glucose	C00267	M00001 M000549 M00909	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Fuctose	C00095		19	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar biosynthesis and galactose degradation leloir pathway
		beta-D-glucose	C00221		19	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		19	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		19	M+Na	203.0526	180.0634	C6H12O6			
F022	207.99861	O-Phospho L-serine	C01005	M00020	2	M+Na	207.9981	185.0089	C3H8NO6P	2.0082	1.0059	Serine biosynthesis
F022	210.04333	(2S,4S)-4-Hydroxy-2,3,4,5-tetrahydrodipicolinate	C20258	M00016	29	M+Na	210.0373	187.0481	C7H9NO5	2.1332	1.093	Lysine biosynthesis
F022	211.10117									2.3356	1.2238	
F022	214.00194	2-Amino-3-oxo-4-phosphonoxybutyrate	C07335	M00124	43	M+H	214.0111	213.0038	C4H8NO7P	2.4879	1.3149	Pyridoxal-P biosynthesis

		ate										
		L-Aspartyl-4-phosphate	C03082	M00016 M00017 M00018	43	M+H	214.0111	213.0038	C4H8NO7P			Lysine threonine methionine biosynthesis
ELU39	226.94624	2,5-Dichloro-4-oxohex-2-enedioate	C12835		20	M+H	226.9509	225.9436	C6H4Cl2O5	2.7032	1.4347	
F054	232.0775	N2-Succinyl-L-glutamic acid 5-semialdehyde	C05932	M00879	18	M+H	232.0816	231.0743	C9H13NO6	0.4473	-1.1607	Arginine succinyltransferase pathway
ELU39	233.05536	O-Phospho-4-hydroxy-L-threonine	C06055	M00124	9	M+NH4	233.0533	215.0195	C4H10NO7P	0.49301	-1.0203	Pyridoxal-P biosynthesis
F054	233.05536	O-Phospho-4-hydroxy-L-threonine	C06055	M00124	9	M+NH4	233.0533	215.0195	C4H10NO7P	0.49726	-1.0079	Pyridoxal-P biosynthesis
F054	242.09407	Pantothenic acid	C00864	M00019 M00120	24	M+Na	242.099	219.1107	C9H17NO5	0.49066	-1.0272	Pantothenate biosynthesis, Coenzyme A biosynthesis
F022	243.04403	L-Tryptophan	C00078	M00023	37	M+K	243.053	204.0899	C11H12N2O2	3.729	1.8988	Tryptophan biosynthesis
MG1655	243.04561	L-Tryptophan	C00078	M00023	31	M+K	243.053	204.0899	C11H12N2O2	2.3137	1.2102	Tryptophan biosynthesis
F054	243.04811	L-Tryptophan	C00078	M00023	20	M+K	243.053	204.0899	C11H12N2O2	3.143	1.6521	Tryptophan biosynthesis
ELU39	243.04876	L-Tryptophan	C00078	M00023	18	M+K	243.053	204.0899	C11H12N2O2	2.5429	1.3465	Tryptophan biosynthesis
F022	249.11471	N2-Succinyl-L-glutamic acid 5-semialdehyde	C05932	M00879	26	M+NH4	249.1081	231.0743	C9H13NO6	2.0904	1.0638	Arginine succinyltransferase pathway
ELU39	251.0629	Deoxyuridine	C00526		4	M+Na	251.0638	228.0746	C9H12N2O5	0.43449	-1.2026	Pyrimidine metabolism
		gamma-Glutamylcysteine	C00669		27	M+H	251.0696	250.0623	C8H14N2O5S			Glutathione biosynthesis
MG1655	251.06321	Deoxyuridine	C00526		3	M+Na	251.0638	228.0746	C9H12N2O5	0.44029	-1.1835	Pyrimidine metabolism
		gamma-Glutamylcysteine	C00669		26	M+H	251.0696	250.0623	C8H14N2O5S			Glutathione biosynthesis
F054	251.06333	Deoxyuridine	C00526		2	M+Na	251.0638	228.0746	C9H12N2O5	0.31389	-1.6717	Pyrimidine metabolism
		gamma-Glutamylcysteine	C00669		25	M+H	251.0696	250.0623	C8H14N2O5S			Glutathione biosynthesis
F022	261.04796	3-Deoxy-D-manno-octulosonate	C01187	M00063	39	M+Na	261.0581	238.0689	C8H14O8	0.45913	-1.123	CMP-KDO biosynthesis
		Alpha-D-glucose 6-phosphate	C00668	M00001 M00004	42	M+H	261.037	260.0297	C6H13O9P			Glycolysis, pentose phosphate pathway
		D-Allose 6-phosphate	C02962		42	M+H	261.037	260.0297	C6H13O9P			Sugar metabolism

		phosphate										
		D-Myo-inositol (1)-monophosphate	C01177		42	M+H	261.037	260.0297	C6H13O9P			Inisitol phosphate metabolism
		D-Myo-inositol 4-phosphate	C03546		42	M+H	261.037	260.0297	C6H13O9P			Inisitol phosphate metabolism
F022	265.07602	Thymidine	C00214		13	M+Na	265.0795	242.0903	C10H14N2O5	2.0458	1.0327	Pyrimidine metabolism
F022	266.06812	Cytidine	C00475		25	M+Na	266.0747	243.0855	C9H13N3O5	2.1572	1.1092	Precursor of cytosine
F054	268.10123	Deoxyguanosine	C00330		10	M+H	268.104	267.0968	C10H13N5O4	0.43378	-1.205	Purine metabolism
		Adenosine	C00212	M00958	10	M+H	268.104	267.0968	C10H13N5O4			Adenine ribonucleotide degradation, purine metabolism
		gamma-Glutamylcysteine	C00669		19	M+NH4	268.0962	250.0623	C8H14N2O5S			Glutathione biosynthesis
MG1655	268.14253	Ubiquinone	C00399	M00117	44	M+NH4	268.1543	250.1205	C14H18O4	40.235	5.3304	Oxidative phosphorylation
F022	270.05703	N-Succinyl-L-glutamate	C05931	M00879	5	M+Na	270.0584	247.0692	C9H13NO7	2.0074	1.0053	Arginine succinyltransferase pathway
F054	270.05725	N-Succinyl-L-glutamate	C05931	M00879	4	M+Na	270.0584	247.0692	C9H13NO7	0.38031	-1.3948	Arginine succinyltransferase pathway
F054	273.04743	gamma-Glutamylcysteine	C00669		15	M+Na	273.0516	250.0623	C8H14N2O5S	0.40901	-1.2898	Glutathione biosynthesis
F054	274.13535	N2-succinyl-L-arginine	C03296	M00879	28	M+H	274.1277	273.1204	C10H17N4O5	2.2295	1.1567	Arginine succinyltransferase pathway
F022	282.02832	D-Galactosamine 6-phosphate	C06377		23	M+Na	282.0349	259.0457	C6H14NO8P	2.3247	1.217	Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		23	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501	M00909	23	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	M00909	23	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis

F022	298.09128	1,6-Anhydro-N-acetyl-beta-muramate	C19769		5	M+Na	298.0897	275.1005	C11H17NO7	2.8229	1.4972	Nucleotide sugar biosynthesis
		5'-Methylthioadenosine	C00170		19	M+H	298.0968	297.0896	C11H15N5O3S			Methionine salvage
F022	304.93312	3-phospho-D-glyceroyl phosphate	C00236	M00001 M00002 M00003	35	M+K	304.9224	265.9593	C3H8O10P2	2.4357	1.2843	Glycolysis, Gluconeogenesis
F054	314.0835	Phosphoribosylformylglycineamidine	C04640	M00048	28	M+H	314.0748	313.0675	C8H16N3O8P	4.687	2.2287	De novo purine biosynthesis
F054	332.06703	dAMP	C00360		25	M+H	332.0754	331.0682	C10H14N5O6P	0.027208	-5.1998	ADP derivative
MG1655	332.06769	dAMP	C00360		23	M+H	332.0754	331.0682	C10H14N5O6P	373.07	8.5433	ADP derivative
ELU39	339.0752	AICAR	C04677	M00048	15	M+H	339.07	338.0627	C9H15N4O8P	2.6455	1.4035	De novo purine biosynthesis
F022	354.04832	dAMP	C00360		26	M+Na	354.0574	331.0682	C10H14N5O6P	2.715	1.441	ADP derivative
MG1655	354.04958	dAMP	C00360		22	M+Na	354.0574	331.0682	C10H14N5O6P	2.1291	1.0902	ADP derivative
ELU39	355.07215	5-Amino-6-(5'-phosphoribosylamino)uracil	C01268		20	M+H	355.0649	354.0577	C9H15N4O9P	0.44466	-1.1692	Riboflavin biosynthesis, plants and bacteria
F022	361.06578	AICAR	C04677	M00048	38	M+Na	338.0627	361.052	C9H15N4O8P	2.4847	1.3131	De novo purine biosynthesis
ELU39	361.07054									3.7509	1.9072	
ELU39	365.0533	Xanthylic acid	C00655	C00050	11	M+H	365.0493	364.042	C10H13N4O9P	2.4734	1.3065	Guanine ribonucleotide biosynthesis
F022	370.05291	3'-AMP	C01367		2	M+Na	370.0523	347.0631	C10H14N5O7P	3.2259	1.6897	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'-monophosphate dGMP	C0362		2	M+Na	370.0523	347.0631	C10H14N5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	2	M+Na	370.0523	347.0631	C10H14N5O7P			Adenine ribonucleotide biosynthesis and

												degradation
MG1655	370.06138	3'-AMP	C01367		25	M+Na	370.0523	347.0631	C10H14N5O7P	3.6903	1.8837	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'-monophosphate dGMP	C0362		25	M+Na	370.0523	347.0631	C10H14N5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	25	M+Na	370.0523	347.0631	C10H14N5O7P			Adenine ribonucleotide biosynthesis and degradation
ELU39	370.06139	3'-AMP	C01367		25	M+Na	370.0523	347.0631	C10H14N5O7P	5.4155	2.4371	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'-monophosphate dGMP	C0362		25	M+Na	370.0523	347.0631	C10H14N5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	25	M+Na	370.0523	347.0631	C10H14N5O7P			Adenine ribonucleotide biosynthesis and degradation
F054	370.0618	3'-AMP	C01367		26	M+Na	370.0523	347.0631	C10H14N5O7P	3.1206	1.6418	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		2'-Deoxyguanosine 5'-monophosphate dGMP	C0362		26	M+Na	370.0523	347.0631	C10H14N5O7P			Purine metabolism
		AMP	C00020	M00049 M00958	26	M+Na	370.0523	347.0631	C10H14N5O7P			Adenine ribonucleotide biosynthesis and degradation
F022	380.11062	S-Lactoylglutathione	C03451		4	M+H	380.1122	379.1049	C13H21N3O8S	2.1522	1.1058	Pyruvate metabolism
F054	380.11601	S-Lactoylglutathione	C03451		10	M+H	380.1122	379.1049	C13H21N3O8S	0.37298	-1.4228	Pyruvate metabolism
F054	423.09392	Trehalose 6-phosphate	C00689		10	M+H	423.0898	422.0825	C12H23O14P	0.29864	-1.7435	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		22	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation

MG1655	423.09494	Trehalose 6-phosphate	C00689		12	M+H	423.0898	422.0825	C12H23O14P	0.39329	-1.3463	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		24	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation
ELU39	423.09581	Trehalose 6-phosphate	C00689		14	M+H	423.0898	422.0825	C12H23O14P	0.45199	-1.1456	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		26	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation
ELU39	428.03494	dGDP	C00361	M00053	4	M+H	428.0367	427.0294	C10H15N5O10P2	0.47244	-1.0818	Deoxyribonucleotide biosynthesis
		ADP	C00008	M00053	4	M+H	428.0367	427.0294	C10H15N5O10P2			Deoxyribonucleotide biosynthesis
		Adenosine phosphosulfate	C00224	M00176	18	M+H	428.0272	427.0199	C10H14N5O10PS			Assimilatory sulfate reduction
F022	442.14173	Folate	C00504		12	M+H	442.147	441.1397	C19H19N7O6	2.2108	1.1445	Folate metabolism
MG1655	445.07722	Trehalose 6-phosphate	C00689		12	M+Na	423.0898	422.0825	C12H23O14P	0.49373	-1.0182	Other carbohydrate metabolism
ELU39	484.11616	Tetrahydrofolic acid	C00101	M00621	37	M+K	484.1341	445.171	C19H23N7O6	0.44373	-1.1722	Tetrahydrofolate biosynthesis, glycine cleavage system

Table S5.3 Identification table of metabolites significantly affected by the ciprofloxacin treatment in the 4 *E.coli* strains. Masses were identified using ECMDB and cross referenced with KEGG metabolism database for *E.coli* K-12 MG1655.

Ciprofloxacin												
Strain	Mass	ID	KEGG	MODULE	ppm	Adduct	Adduct + Compound M/Z	Monoisotopic Mass	Chemical Formula	Fold Change	log2(FC)	Function
MG1655	81.02721	Nitrate	C0244	M00530	28	M+NH4	81.0295	62.9956	HNO3	2.422	1.2762	Dissimilatory nitrate reduction
F054	97.02252	D-Lactaldehyde	C00937		36	M+Na	97.026	74.0368	C3H6O2	2.3258	1.2177	Purine metabolism
		(S)-Lactaldehyde	C00424		36	M+Na	97.026	74.0368	C3H6O2			Pyruvate metabolism
		Propanoate	C00163		36	M+Na	97.026	74.0368	C3H6O2			
F054	98.05503	Trimethylamine N-Oxide	C01104		27	M+Na	98.0576	75.0684	C3H9NO	0.38252	-1.3864	Methane metabolism
F022	99.03870	Propylene glycol	C00583		30	M+Na	99.0416	76.0524	C3H8O2	41.95	5.3906	Propanoate metabolism
F054	101.00322	Dimethyl sulfoxide	C11143		1	M+Na	101.0032	78.0139	C2H6OS	2.4843	1.3128	Sulfur metabolism
MG1655	110.02734	Hypotaurine	C00519		3	M+H	110.027	109.0197	C2H7NO2S	2.5364	1.3428	Hypotaurine metabolism
ELU39	112.00041	Oxamate	C01444		1	M+Na	112.0005	89.0113	C2H3NO3	2.5124	1.3291	DEM
F022	112.00041	Oxamate	C01444		1	M+Na	112.0005	89.0113	C2H3NO3	0.14032	-2.8332	DEM
ELU39	112.04662	Cytosine	C00380		35	M+H	112.0505	111.0433	C4H5N3O	3.9367	1.977	Pyrimidine metabolism
F022	112.04662	Cytosine	C00380		35	M+H	112.0505	111.0433	C4H5N3O	4.1877	2.0662	Pyrimidine metabolism
F054	112.04662	Cytosine	C00380		35	M+H	112.0505	111.0433	C4H5N3O	3.7353	1.9012	pyrimidine metabolism
MG1655	112.04662	Cytosine	C00380		35	M+H	112.0505	111.0433	C4H5N3O	9.6471	3.2701	Pyrimidine metabolism
F022	113.01904	D-Lactic acid	C00256		17	M+Na	113.0209	90.0317	C3H6O3	2.166	1.1151	Pyruvate metabolism
		3-Hydroxypropanoate	C01013	M00939	17	M+Na	113.0209	90.0317	C3H6O3			Pyrimidine degradation
		L-Lactic acid	C00186		17	M+Na	113.0209	90.0317	C3H6O3			Pyruvate metabolism
		Glycerone	C00184		17	M+Na	113.0209	90.0317	C3H6O3			Glycerolipid metabolism
MG1655	116.06499	L-Proline	C00148	M00015 M00970	48	M+H	116.0706	115.0633	C5H9NO2	0.25605	-1.9655	Proline biosynthesis and degradation
ELU39	122.02266	D-Cysteine	C00793		36	M+H	122.027	121.0197	C3H7NO2S	0.37521	-1.4142	
		L-Cysteine	C00097	M00021	36	M+H	122.027	121.0197	C3H7NO2S			Cysteine biosynthesis
ELU39	124.03682	Nicotinic acid	C00253	M00115	20	M+H	124.0393	123.032	C6H5NO2	0.24075	-2.0544	Precursor to NAD
F054	134.02625	Cytosine	C00380		46	M+Na	134.0325	111.0433	C4H5N3O	3.9408	1.9785	Pyrimidine metabolism
MG1655	134.02625	Cytosine	C00380		46	M+Na	134.0325	111.0433	C4H5N3O	4.8452	2.2766	Pyrimidine metabolism
MG1655	134.02625	Cytosine	C00380		46	M+Na	134.0325	111.0433	C4H5N3O	2.8788	1.5255	Pyrimidine metabolism

ELU39	134.02878	Cytosine	C00380		28	M+Na	134.0325	111.0433	C4H5N3O	2.6177	1.3883	Pyrimidine metabolism
ELU39	147.10596	L-Lysine	C00047	M00016 M00956	47	M+H	147.1128	146.1055	C6H14N2O 2	0.40812	-1.2929	Lysine metabolism
F054	147.10596	L-Lysine	C00047	M00016 M00956	47	M+H	147.1128	146.1055	C6H14N2O 2	0.43228	-1.21	Lysine metabolism
ELU39	152.02783	Pyrroline hydroxycarboxylic acid	C04281		26	M+Na	152.0318	129.0426	C5H7NO3	0.40057	-1.3199	Arginine and proline metabolism
MG1655	152.02783	Pyrroline hydroxycarboxylic acid	C04281		26	M+Na	152.0318	129.0426	C5H7NO3	0.38598	-1.3734	Arginine and proline metabolism
ELU39	152.05475	(R)-Malate	C00497		4	M+NH4	152.0553	134.0215	C4H6O5	0.382	-1.3884	
		L-Malic acid	C00149	M00009 M00168	4	M+NH4	152.0553	134.0215	C4H6O5			TCA cycle, CAM
		Guanine	C00242	M00959	13	M+H	152.0567	151.0494	C5H5N5O			Guanine ribonucleotide degradation
F022	156.03674	Ureidoglycine	C02091		8	M+Na	156.038	133.0487	C3H7N3O3	0.29715	-1.7507	Purine metabolism
		L-Valine	C00183	M00019	35	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism valine- isoleucine metabolism
		Betaine	C00719	M00555	38	M+K	156.0427	118.0868	C5H12NO2			Serine and threonine biosynthesis
F054	156.03674	Ureidoglycine	C02091		8	M+Na	156.038	133.0487	C3H7N3O3	0.32213	-1.6343	Purine metabolism
		L-Valine	C00183	M00019	35	M+K	156.0421	117.079	C5H11NO2			Branched chain amino acid metabolism valine- isoleucine metabolism
		Betaine	C00719	M00555	38	M+K	156.0427	118.0868	C5H12NO2			Serine and threonine biosynthesis
ELU39	156.99240	Succinic acid	C00042	M00009	17	M+K	156.9898	118.0266	C4H6O4	0.35049	-1.5126	TCA cycle
		Phosphoglycolic acid	C00988		18	M+H	156.9897	155.9824	C2H5O6P			Glyoxylate metabolism
F022	157.04164	(R)-2,3-Dihydroxy- isovalerate	C04272	M00019	35	M+Na	157.0471	134.0579	C5H10O4	0.48546	-1.0426	Valine leucine isoleucine biosynthesis
		Deoxyribose	C01801		35	M+Na	157.0471	134.0579	C5H10O4			Pentose phosphate pathway
F054	157.04164	(R)-2,3-Dihydroxy- isovalerate	C04272	M00019	35	M+Na	157.0471	134.0579	C5H10O4	0.33495	-1.578	Valine leucine isoleucine biosynthesis
		Deoxyribose	C01801		35	M+Na	157.0471	134.0579	C5H10O4			Pentose phosphate

												pathway
ELU39	158.95654	3-Mercaptopyruvic acid	C00957	33	M+K	158.9513	119.9881	C3H4O3S	3.3835	1.7585		Cysteine and methionine metabolism
F054	162.10899	L-Carnitine	C00318	21	M+H	162.1125	161.1052	C7H15NO3	0.48448	-1.0455		
		Carnitine	C00487	25	M+H	162.113	162.113	C7H16NO3				Lysine degradation
MG1655	162.10899	L-Carnitine	C00318	21	M+H	162.1125	161.1052	C7H15NO3	0.35419	-1.4974		
		Carnitine	C00487	25	M+H	162.113	162.113	C7H16NO3				Lysine degradation
ELU39	167.07620	3-(3-Hydroxyphenyl)propanoic acid	C11457	35	M+H	167.0703	166.063	C9H10O3	4.5255	2.1781		Phenylalanine metabolism
ELU39	173.03587	D-Xylulose	C00310	36	M+Na	173.042	150.0528	C5H10O5	2.5327	1.3406		Pentose glucuronate interconversions
		L-Ribulose	C00508	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions
		L-Threo-2-pentulose	C00312	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions
		L-Arabinose	C00259	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions, nucleotide sugar metabolism
		D-Ribulose	C00309	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions
		Ribose	C00121	36	M+Na	173.042	150.0528	C5H10O5				Pentose phosphate pathway
		D-Xylose	C00181	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions, nucleotide sugar metabolism
ELU39	173.03587	D-Xylulose	C00310	36	M+Na	173.042	150.0528	C5H10O5	2.4196	1.2748		Pentose glucuronate interconversions
		L-Ribulose	C00508	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions
		L-Threo-2-pentulose	C00312	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions
		L-Arabinose	C00259	36	M+Na	173.042	150.0528	C5H10O5				Pentose glucuronate interconversions, nucleotide sugar

												metabolism
		D-Ribulose	C00309		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions
		Ribose	C00121		36	M+Na	173.042	150.0528	C5H10O5			Pentose phosphate pathway
		D-Xylose	C00181		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions, nucleotide sugar metabolism
F022	173.03587	D-Xylulose	C00310		36	M+Na	173.042	150.0528	C5H10O5	3.4751	1.797	Pentose glucuronate interconversions
		L-Ribulose	C00508		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions
		L-Threo-2-pentulose	C00312		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions
		L-Arabinose	C00259		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions, nucleotide sugar metabolism
		D-Ribulose	C00309		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions
		Ribose	C00121		36	M+Na	173.042	150.0528	C5H10O5			Pentose phosphate pathway
		D-Xylose	C00181		36	M+Na	173.042	150.0528	C5H10O5			Pentose glucuronate interconversions, nucleotide sugar metabolism
ELU39	174.05946	Orotate	C00295	M00051	49	M+NH4	174.0509	156.0171	C5H4N2O4	3.8135	1.9311	De novo pyrimidine biosynthesis
MG1655	175.06296	3-Carboxy-3-hydroxy-isocaproate	C02504	M00432	16	M+H	175.0601	174.0528	C7H10O5	0.049516	-4.3359	Leucine biosynthesis
		2-Isopropyl-3-oxosuccinate	C04236		16	M+H	175.0601	174.0528	C7H10O5			Valine leucine isoleucine biosynthesis
		Shikimate	C00493		16	M+H	175.0601	174.0528	C7H10O5			Shikimate pathway
MG1655	178.05	L-Histidine	C00135		49	M+Na	178.0587	155.0695	C6H9N3O2	2.6523	1.4073	Histidine biosynthesis
F054	187.10301	Pyridoxine	C00314		25	M+NH4	187.1077	169.0739	C8H11NO3	0.46975	-1.09	
F054	188.06256	D-Phenylalanine	C02265		30	M+Na	188.0682	165.079	C9H11NO2	0.48901	-1.0321	

		L-Phenylalanine	C00079	M00024	30	M+Na	188.0682	165.079	C9H11NO2			Phenylalanine biosynthesis
ELU39	191.07056	Pyridoxamine	C00534		45	M+Na	191.0791	168.0899	C8H12N2O2	3.6178	1.8551	
F022	191.07056	Pyridoxamine	C00534		45	M+Na	191.0791	168.0899	C8H12N2O2	3.563	1.8331	
F054	191.07056	Pyridoxamine	C00534		45	M+Na	191.0791	168.0899	C8H12N2O2	2.3541	1.2352	
ELU39	194.93953	Phosphoglycolic acid	C00988		31	M+K	194.9455	155.9824	C2H5O6P	0.43363	-1.2055	Glyoxylate metabolism
MG1655	194.93953	Phosphoglycolic acid	C00988		31	M+K	194.9455	155.9824	C2H5O6P	0.45834	-1.1255	Glyoxylate metabolism
ELU39	197.09073	N-Acetylmethionine	C00437	M00028	5	M+Na	197.0897	174.1004	C7H14N2O3	0.15218	-2.7162	Ornithine metabolism
MG1655	198.02945	7-Cyano-7-carbaguanine	C15996		46	M+Na	198.0386	175.0494	C7H5N5O	0.47332	-1.0791	Folate biosynthesis
ELU39	198.08473	Citrulline	C00327	M000844	1	M+Na	198.0849	175.0957	C6H13N3O3	0.18702	-2.4187	Arginine biosynthesis
		5-Methylthioribose	C03089		27	M+NH4	198.0795	180.0456	C6H12O4S			Cysteine and methionine metabolism
		Trans-2,3-Dihydroxycinnamate	C12623	M00545	44	M+NH4	198.0761	180.0423	C9H8O4			Trans-cinnamate degradation
		4-Hydroxyphenylpyruvic acid	C01179	M00025	44	M+NH4	198.0761	180.0423	C9H8O4			Tyrosine biosynthesis
		3,4-Dihydroxy-L-phenylalanine	C00355		44	M+H	198.0761	197.0688	C9H11NO4			Tyrosine biosynthesis
F022	198.08473	Citrulline	C00327	M000844	1	M+Na	198.0849	175.0957	C6H13N3O3	0.49651	-1.0101	Arginine biosynthesis
		5-Methylthioribose	C03089		27	M+NH4	198.0795	180.0456	C6H12O4S			Cysteine and methionine metabolism
		Trans-2,3-Dihydroxycinnamate	C12623	M00545	44	M+NH4	198.0761	180.0423	C9H8O4			Trans-cinnamate degradation
		4-Hydroxyphenylpyruvic acid	C01179	M00025	44	M+NH4	198.0761	180.0423	C9H8O4			Tyrosine biosynthesis
		3,4-Dihydroxy-L-phenylalanine	C00355		44	M+H	198.0761	197.0688	C9H11NO4			Tyrosine biosynthesis
MG1655	198.08473	Citrulline	C00327	M000844	1	M+Na	198.0849	175.0957	C6H13N3O3	0.23133	-2.112	Arginine biosynthesis

									3			
		5-Methylthioribose	C03089		27	M+NH4	198.0795	180.0456	C6H12O4S			Cysteine and methionine metabolism
		Trans-2,3-Dihydroxycinnamate	C12623	M00545	44	M+NH4	198.0761	180.0423	C9H8O4			Trans-cinamate degradation
		4-Hydroxyphenylpyruvic acid	C01179	M00025	44	M+NH4	198.0761	180.0423	C9H8O4			Tyrosine biosynthesis
		3,4-Dihydroxy-L-phenylalanine	C00355		44	M+H	198.0761	197.0688	C9H11NO4			Tyrosine biosynthesis
MG1655	199.0258	5-Dehydro-4-deoxy-D-glucuronate	C04053		23	M+Na	199.0213	176.0321	C6H8O6	0.19532	-2.3561	
		(4S)-4,6-Dihydroxy-2,5-dioxohexanoate	C04349		23	M+Na	199.0213	176.0321	C6H8O6			
		Ascorbate	C00072	M00550	23	M+Na	199.0213	176.0321	C6H8O6			Ascorbate degradation
		Ureidosuccinic acid	C00438	M00051	34	M+Na	199.0325	176.0433	C5H8N2O5			De novo pyrimidine biosynthesis
F054	200.97037	D-4-Hydroxy-2-oxoglutarate	C05946		46	M+K	200.9796	162.0164	C5H6O6	2.6789	1.4216	Arginine and proline metabolism
		4-Hydroxy-2-oxoglutaric acid	C01127		46	M+K	200.9796	162.0164	C5H6O6			Glyoxylate metabolism
MG1655	202.10439	cis-3-(Carboxy-ethyl)-3,5-cyclo-hexadiene-1,2-diol	C11588		15	M+NH4	202.1074	184.0736	C9H12O4	0.41202	-1.2792	Phenylalanine metabolism
MG1655	203.04874	L-Rhamnonate	C01934		19	M+Na	203.0526	180.0634	C6H12O6	0.37483	-1.4157	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		19	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		19	M+Na	203.0526	180.0634	C6H12O6			
		alpha-D-Glucose	C00267	M00001 M000549 M00909	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Fructose	C00095		19	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar biosynthesis and

												galactose degradation leloir pathway
		beta-D-glucose	C00221		19	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		19	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	19	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		19	M+Na	203.0526	180.0634	C6H12O6			
ELU39	203.05495	L-Rhamnonate	C01934		12	M+Na	203.0526	180.0634	C6H12O6	8.0756	3.0136	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D- glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		12	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		12	M+Na	203.0526	180.0634	C6H12O6			
		alpha-D-Glucose	C00267	M00001 M000549 M00909	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Fructose	C00095		12	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar biosynthesis and galactose degradation leloir pathway
		beta-D-glucose	C00221		12	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		12	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		12	M+Na	203.0526	180.0634	C6H12O6			
MG1655	203.05495	L-Rhamnonate	C01934		12	M+Na	203.0526	180.0634	C6H12O6	17.601	4.1376	Glycolysis. Nucleotide sugar biosynthesis UDP-N-acetyl-D- glucosamine biosynthesis, prokaryotes
		D-Allose	C01487		12	M+Na	203.0526	180.0634	C6H12O6			
		Inositol	C00137		12	M+Na	203.0526	180.0634	C6H12O6			

		alpha-D-Glucose	C00267	M00001 M000549 M00909	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Fructose	C00095		12	M+Na	203.0526	180.0634	C6H12O6			Nucleotide sugar biosynthesis and galactose degradation leloir pathway
		beta-D-glucose	C00221		12	M+Na	203.0526	180.0634	C6H12O6			
		D-Mannose	C00159		12	M+Na	203.0526	180.0634	C6H12O6			
		Alpha-D-Galactose	C00984	M00554 M00632	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Galactose	C00124	M00632	12	M+Na	203.0526	180.0634	C6H12O6			
		D-Glucose	C00031		12	M+Na	203.0526	180.0634	C6H12O6			
F022	205.06905	Galactitol	C01697		4	M+Na	205.0683	182.079	C6H14O6	11.645	3.5417	Galactose metabolism
		Sorbitol	C00794		4	M+Na	205.0683	182.079	C6H14O6			Fructose mannose metabolism
		Mannitol	C00392		4							Fructose mannose metabolism
ELU39	207.04280	Methylisocitric acid	C04593		34	M+H	207.0499	206.0427	C7H10O7	0.41139	-1.2814	Propanoate metabolism
		Methylcitric acid	C02225		34	M+H	207.0499	206.0427	C7H10O7			Propanoate metabolism
		Lipoic acid	C16241		39	M+H	207.0508	206.0435	C8H14O2S 2			Lipoic acid metabolism
MG1655	207.0428	Methylisocitric acid	C04593		34	M+H	207.0499	206.0427	C7H10O7	0.43417	-1.2037	Propanoate metabolism
		Methylcitric acid	C02225		34	M+H	207.0499	206.0427	C7H10O7			Propanoate metabolism
		Lipoic acid	C16241		39	M+H	207.0508	206.0435	C8H14O2S 2			Lipoic acid metabolism
ELU39	211.10117									0.45633	-1.1318	
F054	212.84886									0.41356	-1.2738	
ELU39	214.00095	2-Amino-3-oxo-4-phosphonoxybutyrate	C07335	M00124	47	M+H	214.0111	213.0038	C4H8NO7P	2.4866	1.3141	Pyridoxal-P biosynthesis
		L-Aspartyl-4-phosphate	C03082	M00016 M00017 M00018	47	M+H	214.0111	213.0038	C4H8NO7P			Lysine threonine methionine biosynthesis
ELU39	214.00245	2-Amino-3-oxo-4-phosphonoxybutyrate	C07335	M00124	40	M+H	214.0111	213.0038	C4H8NO7P	3.4922	1.8042	Pyridoxal-P biosynthesis

		L-Aspartyl-4-phosphate	C03082	M00016 M00017 M00018	40	M+H	214.0111	213.0038	C4H8NO7P			Lysine threonine methionine biosynthesis
		7-Cyano-7-carbaguanine	C15996		47	M+K	214.0126	175.0494	C7H5N5O			Folate biosynthesis
MG1655	220.11236	Panθοthenic acid	C00864	M00019 M00120	25	M+H	220.1179	219.1107	C9H17NO5	0.24982	-2.0011	Pantothenate biosynthesis, Coenzyme A biosynthesis
ELU39	220.11247	Panθοthenic acid	C00864	M00019 M00120	25	M+H	220.1179	219.1107	C9H17NO5	0.24338	-2.0387	Pantothenate biosynthesis, Coenzyme A biosynthesis
F022	226.94615	2,5-Dichloro-4-oxohex-2-enedioate	C12835		21	M+H	226.9509	225.9436	C6H4Cl2O5	2.1431	1.0997	
MG1655	231.03858	2-Amino-3-oxo-4-phosphonooxybutyrate	C07335	M00124	4	M+NH4	231.0377	213.0038	C4H8NO7P	2.0392	1.028	Pyridoxal-P biosynthesis
		L-Aspartyl-4-phosphate	C03082	M00016 M00017 M00018	4	M+NH4	231.0377	213.0038	C4H8NO7P			Lysine threonine methionine biosynthesis
F022	231.03876	2-Amino-3-oxo-4-phosphonooxybutyrate	C07335	M00124	5	M+NH4	231.0377	213.0038	C4H8NO7P	2.59	1.373	Pyridoxal-P biosynthesis
		L-Aspartyl-4-phosphate	C03082	M00016 M00017 M00018	5	M+NH4	231.0377	213.0038	C4H8NO7P			Lysine threonine methionine biosynthesis
ELU39	233.10867	4-(Glutamylamino)butanoate	C16757	M00136	19	M+H	233.1132	232.1059	C9H16N2O5	0.15072	-2.73	GABA biosynthesis
		N2-Succinyl-L-ornithine	C03415	M00879	19	M+H	233.1132	232.1059	C9H16N2O5			Arginine succinyltransferase pathway
F022	233.10867	4-(Glutamylamino)butanoate	C16757	M00136	19	M+H	233.1132	232.1059	C9H16N2O5	0.30835	-1.6974	GABA biosynthesis
		N2-Succinyl-L-ornithine	C03415	M00879	19	M+H	233.1132	232.1059	C9H16N2O5			Arginine succinyltransferase pathway
F054	233.10867	4-(Glutamylamino)butanoate	C16757	M00136	19	M+H	233.1132	232.1059	C9H16N2O5	0.26788	-1.9004	GABA biosynthesis

		N2-Succinyl-L-ornithine	C03415	M00879	19	M+H	233.1132	232.1059	C9H16N2O5			Arginine succinyltransferase pathway
MG1655	233.10867	4-(Glutamylamino)butanoate	C16757	M00136	19	M+H	233.1132	232.1059	C9H16N2O5	0.096314	-3.3761	GABA biosynthesis
		N2-Succinyl-L-ornithine	C03415	M00879	19	M+H	233.1132	232.1059	C9H16N2O5			Arginine succinyltransferase pathway
ELU39	238.97545	D-Erythrose 4-phosphate	C00279	M00004	15	M+K	238.9717	200.0086	C4H9O7P	0.48106	-1.0557	Pentose phosphate pathway
ELU39	242.03727	4-Amino-2-methyl-5-phosphomethylpyrimidine	C04556	M00127	30	M+Na	242.0301	219.0409	C6H10N3O4P	0.38553	-1.3751	Thiamine biosynthesis
F054	242.09427	Pantothenic acid	C00864	M00019 M00120	23	M+Na	242.099	219.1107	C9H17NO5	0.30349	-1.7203	Pantothenate biosynthesis, Coenzyme A biosynthesis
F022	242.09576	Pantothenic acid	C00864	M00019 M00120	17	M+Na	242.099	219.1107	C9H17NO5	0.49079	-1.0268	Pantothenate biosynthesis, Coenzyme A biosynthesis
MG1655	242.09613	Pantothenic acid	C00864	M00019 M00120	16	M+Na	242.099	219.1107	C9H17NO5	0.13989	-2.8376	Pantothenate biosynthesis, Coenzyme A biosynthesis
ELU39	242.09695	Pantothenic acid	C00864	M00019 M00120	12	M+Na	242.099	219.1107	C9H17NO5	0.17271	-2.5335	Pantothenate biosynthesis, Coenzyme A biosynthesis
MG1655	243.04241	L-Tryptophan	C00078	M00023	44	M+K	243.053	204.0899	C11H12N2O2	2.9903	1.5803	Tryptophan biosynthesis
MG1655	243.09647	Thymidine	C00214		4	M+H	243.0975	242.0903	C10H14N2O5	0.44753	-1.1599	Pyrimidine metabolism
MG1655	244.08578	Isochorismate	C00885	M00116	17	M+NH4	244.0816	226.0477	C10H10O6	4.6008	2.2019	Menaquinone biosynthesis
		Prephenate	C00254	M00024, M00025	17	M+NH4	244.0816	226.0477	C10H10O6			Tyrosine biosynthesis phenylalanine biosynthesis
		Chorismate	C00251	M00022, M00023, M00024, M00025,	17	M+NH4	244.0816	226.0477	C10H10O6			Central to aromatic amino acid metabolism. Shikimate pathways, tryptophan, tyrosine,

				M00116 M00117								phenylalanine biosynthesis.
		N-Acetylmannosamine	C00645		27	M+Na	244.0792	221.0899	C8H15NO6			Nucleotide sugar biosynthesis
		N-Acetyl-D-glucosamine	C00140		27	M+Na	244.0792	221.0899	C8H15NO6			Nucleotide sugar biosynthesis
		Cytidine	C00475		29	M+H	244.0928	243.0855	C9H13N3O5			Precursor of cytosine
F022	244.08637	Isochorismate	C00885	M00116	20	M+NH4	244.0816	226.0477	C10H10O6	2.6577	1.4102	Menaquinone biosynthesis
		Prephenate	C00254	M00024, M00025	20	M+NH4	244.0816	226.0477	C10H10O6			Tyrosine biosynthesis phenylalanine biosynthesis
		Chorismate	C00251	M00022, M00023, M00024, M00025, M00116 M00117	20	M+NH4	244.0816	226.0477	C10H10O6			Central to aromatic amino acid metabolism. Shikimate pathways, tryptophan, tyrosine, phenylalanine biosynthesis.
		N-Acetylmannosamine	C00645		26	M+Na	244.0792	221.0899	C8H15NO6			Nucleotide sugar biosynthesis
		N-Acetyl-D-glucosamine	C00140		26	M+Na	244.0792	221.0899	C8H15NO6			Nucleotide sugar biosynthesis
		Cytidine	C00475		30	M+H	244.0928	243.0855	C9H13N3O5			Precursor of cytosine
MG1655	244.12243	Porphobilinogen	C00931		28	M+NH4	244.1292	226.0954	C10H14N2O4	3.1315	1.6468	Porphyrin precursor, heme biosynthesis
F022	245.05476	L-Glutamic acid 5-phosphate	C03287	M00015	6	M+NH4	245.0533	227.0195	C5H10NO7P	2.9897	1.58	Proline biosynthesis
		L-Cystathionine	C02291	M00017	8	M+Na	245.0566	222.0674	C7H14N2O4S			Methionine biosynthesis
ELU39	251.06318	Deoxyuridine	C00526		3	M+Na	251.0638	228.0746	C9H12N2O5	0.36025	-1.4729	Pyrimidine metabolism
		gamma-Glutamylcysteine	C00669		26	M+H	251.0696	250.0623	C8H14N2O5S			Glutathione biosynthesis
ELU39	255.06808	gamma-Glutamyl-gamma-	C15700		24	M+K	255.0742	216.111	C9H16N2O4	0.35003	-1.5145	GABA biosynthesis

		butyraldehyde										
F022	255.09020	4-(Glutamylamino)butanoate	C15767	M00136	19	M+Na	255.0951	232.1059	C9H16N2O5	0.40841	-1.2919	GABA biosynthesis
		N2-Succinyl-L-ornithine	C03415	M00879	19	M+Na	255.0951	232.1059	C9H16N2O5			Arginine succinyltransferase pathway
		Nicotinamide riboside	C03150		31	M+H	255.0981	255.0981	C11H15N2O5			Nicotinate metabolism
MG1655	256.09664	3-Deoxy-D-manno-octulosonate	C01187	M00063	24	M+NH4	256.1027	238.0689	C8H14O8	0.47483	-1.0745	CMP-KDO biosynthesis
		7,8-Dihydroneopterin	C04874	M00126	29	M+H	256.104	255.0968	C9H13N5O4			Tetrahydrofolate biosynthesis
		gamma-Glutamyl-L-putrescine	C15699		36	M+K	256.1058	217.1426	C9H19N3O3			GABA biosynthesis
ELU39	256.09774	3-Deoxy-D-manno-octulosonate	C01187	M00063	19	M+NH4	256.1027	238.0689	C8H14O8	0.3554	-1.4925	CMP-KDO biosynthesis
		7,8-Dihydroneopterin	C04874		25	M+H	256.104	255.0968	C9H13N5O4			Tetrahydrofolate biosynthesis
		gamma-Glutamyl-L-putrescine	C15699		31	M+K	256.1058	217.1426	C9H19N3O3			GABA biosynthesis
MG1655	258.06909	Pantothenic acid	C00864	M00019 M00120	18	M+K	258.0738	219.1107	C9H17NO5	0.43038	-1.2163	Pantothenate biosynthesis, Coenzyme A biosynthesis
		L-Cystine	C00491		44	M+NH4	258.0577	240.0238	C6H12N2O4S2			Cysteine and methionine metabolism
ELU39	258.06970	Pantothenic acid	C00864	M00019 M00120	16	M+K	258.0738	219.1107	C9H17NO5	0.4556	-1.1341	Pantothenate biosynthesis, Coenzyme A biosynthesis
		L-Cystine	C00491		47	M+NH4	258.0577	240.0238	C6H12N2O4S2			Cysteine and methionine metabolism
F054	260.04756	D-Galactosamine 6-phosphate	C06377		21	M+H	260.053	259.0457	C6H14NO8P	0.46553	-1.1031	Galactose metabolism
		N-Acetylmannosamine	C00645		21	M+K	260.0531	221.0899	C8H15NO6			
		N-Acetyl-D-glucosamine	C00140		21	M+K	260.0531	221.0899	C8H15NO6			
		alpha-D-Glucosamine	C06156		21	M+H	260.053	259.0457	C6H14NO8			UDP-N-acetyl-D-

		1-phosphate							P			glucosamine biosynthesis
		Glucosamine-1P	C04501	M00909	21	M+H	260.053	259.0457	C6H14NO8 P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	M00909	21	M+H	260.053	259.0457	C6H14NO8 P			UDP-N-acetyl-D-glucosamine biosynthesis
ELU39	260.04817	D-Galactosamine 6-phosphate	C06377		18	M+H	260.053	259.0457	C6H14NO8 P	0.43899	-1.1877	Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		18	M+H	260.053	259.0457	C6H14NO8 P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501	M00909	18	M+H	260.053	259.0457	C6H14NO8 P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	M00909	18	M+H	260.053	259.0457	C6H14NO8 P			UDP-N-acetyl-D-glucosamine biosynthesis
		N-Acetylmannosamine	C00645		19	M+K	260.0531	221.0899	C8H15NO6			
		N-Acetyl-D-glucosamine	C00140		19	M+K	260.0531	221.0899	C8H15NO6			
F054	261.05221	3-Deoxy-D-manno-octulosonate	C01187	M00063	23	M+Na	261.0581	238.0689	C8H14O8	0.40559	-1.3019	CMP-KDO biosynthesis
F022	261.05340	3-Deoxy-D-manno-octulosonate	C01187	M00063	18	M+Na	261.0581	238.0689	C8H14O8	0.30007	-1.7367	CMP-KDO biosynthesis
F022	261.12235	Cytidine	C00475		12	M+NH4	261.1193	243.0855	C9H13N3O 5	0.43245	-1.2094	Precursor of cytosine
MG1655	266.06812	Cytidine	C00475		25	M+Na	266.0747	243.0855	C9H13N3O 5	7.2813	2.8642	Precursor of cytosine
F022	266.06826	Cytidine	C00475		24	M+Na	266.0747	243.0855	C9H13N3O 5	3.4269	1.7769	Precursor of cytosine
F022	266.06913	Cytidine	C00475		21	M+Na	266.0747	243.0855	C9H13N3O 5	4.3931	2.1352	Precursor of cytosine
F054	266.06932	Cytidine	C00475		20	M+Na	266.0747	243.0855	C9H13N3O 5	4.3531	2.122	Precursor of cytosine

F022	267.06411	Pseudouridine	C02067		20	M+Na	267.0588	244.0695	C9H12N2O6	2.1683	1.1166	Pyridoxal-P biosynthesis
		Uridine	C00299		20	M+Na	267.0588	244.0695	C9H12N2O6			Pyrimidine metabolism
		Pyridoxine 5'-phosphate	C00627	M00124	37	M+NH4	267.074	249.0402	C8H12NO6P			Pyridoxal-P biosynthesis
MG1655	267.07108	Pyridoxine 5'-phosphate	C00627	M00124	37	M+NH4	267.074	249.0402	C8H12NO6P	2.9295	1.5507	Pyridoxal-P biosynthesis
		Biotin	C00120	M00123	24	M+Na	267.0774	244.0882	C10H16N2O3S			Biotin biosynthesis
ELU39	268.10123	Deoxyguanosine	C00330		10	M+H	268.104	267.0968	C10H13N5O4	0.48254	-1.0513	Purine metabolism
		Adenosine	C00212	M00958	10	M+H	268.104	267.0968	C10H13N5O4			Adenine ribonucleotide degradation, purine metabolism
		gamma-Glutamylcysteine	C00669		19	M+NH4	268.0962	250.0623	C8H14N2O5S			Glutathione biosynthesis
MG1655	268.10123	Deoxyguanosine	C00330		10	M+H	268.104	267.0968	C10H13N5O4	0.44203	-1.1778	Purine metabolism
		Adenosine	C00212	M00958	10	M+H	268.104	267.0968	C10H13N5O4			Adenine ribonucleotide degradation, purine metabolism
		gamma-Glutamylcysteine	C00669		19	M+NH4	268.0962	250.0623	C8H14N2O5S			Glutathione biosynthesis
MG1655	268.14788	Ubiquinone	C00399	M00117	24	M+NH4	268.1543	250.1205	C14H18O4	126.29	6.9806	Oxidative phosphorylation
F054	270.05583	N-Succinyl-L-glutamate	C05931	M00879	10	M+Na	270.0584	247.0692	C9H13NO7	0.15059	-2.7313	Arginine succinyltransferase pathway
F022	270.05789	N-Succinyl-L-glutamate	C05931	M00879	2	M+Na	270.0584	247.0692	C9H13NO7	3.5914	1.8445	Arginine succinyltransferase pathway
F022	274.09410	Deoxyadenosine	C00559		11	M+Na	274.0911	251.1018	C10H13N5O3	7.7967	2.9629	Purine metabolism
ELU39	274.09525	Deoxyadenosine	C00559		15	M+Na	274.0911	251.1018	C10H13N5O3	4.4405	2.1507	Purine metabolism
ELU39	275.12967	Ubiquinol-1	C00390	M00117	16	M+Na	275.1254	252.1362	14H20O4	0.39219	-1.3504	Oxidative

												phosphorylation
		N2-Succinyl-L-arginine	C03296	M00879	19	M+H	275.135	274.1277	C10H18N4O5			Arginine succinyltransferase pathway
F022	275.12967	Ubiquinol-1	C00390	M00117	16	M+Na	275.1254	252.1362	14H20O4	3.5261	1.8181	Oxidative phosphorylation
		N2-Succinyl-L-arginine	C03296	M00879	19	M+H	275.135	274.1277	C10H18N4O5			Arginine succinyltransferase pathway
MG1655	275.12967	Ubiquinol-1	C00390	M00117	16	M+Na	275.1254	252.1362	14H20O4	0.086225	-3.5358	Oxidative phosphorylation
		N2-Succinyl-L-arginine	C03296	M00879	19	M+H	275.135	274.1277	C10H18N4O5			Arginine succinyltransferase pathway
F022	277.07352	D-Galactosamine 6-phosphate	C06377		22	M+NH4	277.0795	259.0457	C6H14NO8P	0.25373	-1.9787	Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		22	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501		22	M+NH4	277.0795	259.0457	C6H14NO8P			
		Glucosamine 6-phosphate	C00352		22	M+NH4	277.0795	259.0457	C6H14NO8P			
		Nicotinamide riboside	C03150		24	M+Na	277.08	255.0981	C11H15N2O5			
ELU39	277.07424	D-Galactosamine 6-phosphate	C06377		19	M+NH4	277.0795	259.0457	C6H14NO8P	0.30109	-1.7317	Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		19	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501		19	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352		19	M+NH4	277.0795	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Nicotinamide riboside	C03150		21	M+Na	277.08	255.0981	C11H15N2			

								O5				
ELU39	282.03374	D-Galactosamine 6-phosphate	C06377	4	M+Na	282.0349	259.0457	C6H14NO8P	0.41491	-1.2691	Galactose metabolism	
		alpha-D-Glucosamine 1-phosphate	C06156	4	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine-1P	C04501	4	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine 6-phosphate	C00352	4	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
MG1655	282.04398	Cytidine	C00475	17	M+K	282.0487	243.0855	C9H13N3O5	7.8545	2.9735	Precursor of cytosine	
		D-Galactosamine 6-phosphate	C06377	32	M+Na	282.0349	259.0457	C6H14NO8P			Galactose metabolism	
		alpha-D-Glucosamine 1-phosphate	C06156	32	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine-1P	C04501	32	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine 6-phosphate	C00352	32	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
F022	282.04537	Cytidine	C00475	12	M+K	282.0487	243.0855	C9H13N3O5	2.9404	1.556	Precursor of cytosine	
		D-Galactosamine 6-phosphate	C06377	37	M+Na	282.0349	259.0457	C6H14NO8P			Galactose metabolism	
		alpha-D-Glucosamine 1-phosphate	C06156	37	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine-1P	C04501	37	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis	
		Glucosamine 6-phosphate	C00352	37	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine	

												biosynthesis
F054	282.0456	Cytidine	C00475		11	M+K	282.0487	243.0855	C9H13N3O5	3.5104	1.8116	Precursor of cytosine
		D-Galactosamine 6-phosphate	C06377		38	M+Na	282.0349	259.0457	C6H14NO8P			Galactose metabolism
		alpha-D-Glucosamine 1-phosphate	C06156		38	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine-1P	C04501	M00909	38	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
		Glucosamine 6-phosphate	C00352	M00909	38	M+Na	282.0349	259.0457	C6H14NO8P			UDP-N-acetyl-D-glucosamine biosynthesis
F022	283.04688	Biotin	C00120	M00123	16	M+K	283.0513	244.0882	C10H16N2O3S	2.6511	1.4066	Biotin biosynthesis
MG1655	283.0478	Biotin	C00120	M00123	12	M+K	283.0513	244.0882	C10H16N2O3S	4.8496	2.2779	Biotin biosynthesis
ELU39	283.04863	Biotin	C00120	M00123	10	M+K	283.0513	244.0882	C10H16N2O3S	2.4458	1.2903	Biotin biosynthesis
ELU39	285.03283	Sorbitol 6-phosphate	C01096		6	M+Na	285.0346	262.0454	C6H15O9P	2.1176	1.0824	Fructose mannose metabolism
		Galactitol 1-phosphate	C06311		6	M+Na	285.0346	262.0454	C6H15O9P			Galactose metabolism
		2-[(2R,5Z)-2-Carboxy-4-methylthiazol-5(2H)-ylidene]ethyl phosphate	C20246	M00127	8	M+NH4	285.0305	266.9966	C7H10NO6PS			Thiamine biosynthesis
ELU39	287.24039	Retinol	C00473		12	M+H	287.2369	286.2297	C20H30O	3.0787	1.6223	
F054	290.0835	Deoxyguanosine	C00330		9	M+Na	290.086	267.0968	C10H13N5O4	2.3329	1.2221	Purine metabolism
		Adenosine	C00212	M00958	9	M+Na	290.086	267.0968	C10H13N5O4			Adenine ribonucleotide degradation, purine metabolism
		N-Succinyl-2-amino-6-ketopimelate	C04462	M00016	12	M+H	290.087	289.0798	C11H15NO8			Lysine metabolism

F054	291.07189	Inosine	C00294	C00958	7	M+Na	291.07	268.0808	C10H12N4O5	3.0891	1.6272	Adenine ribonucleotide degradation, AMP => Urate
		2(alpha-D-Mannosyl)-D-glycerate	C11544		11	M+Na	291.0687	268.0794	C9H16O9			Other carbohydrate metabolism
MG1655	292.19351	4-Hydroxy-3-polyprenylbenzoate	C04858	M00117	10	M+NH4	292.1907	274.1569	C17H22O3	0.29761	-1.7485	Ubiquinone biosynthesis
MG1655	294.05436	6-Phosphogluconic acid	C00345	M00004	14	M+NH4	294.0585	276.0246	C6H13O10P	0.44697	-1.1617	Pentose phosphate pathway
		7,8-Dihydroneopterin	C04874	M00126	19	M+K	294.0599	255.0968	C9H13N5O4			Tetrahydrofolate biosynthesis
ELU39	296.02692	2-C-Methyl-D-erythritol-2,4-cyclodiphosphate	C11453	M00096	5	M+NH4	296.0284	275.98	C5H10O9P2	0.46943	-1.091	C5 isoprenoid biosynthesis, non-mevalonate pathway
MG1655	297.11148	N2-Succinyl-L-arginine	C03296	M00879	18	M+Na	297.1169	274.1277	C10H18N4O5	0.31368	-1.6726	Arginine succinyltransferase pathway
ELU39	298.09456	5'-Methylthioadenosine	C00170		8	M+H	298.0968	297.0896	C11H15N5O3S	2.1401	1.0977	Methionine salvage
		1,6-Anhydro-N-acetyl-beta-muramate	C19769		16	M+Na	298.0897	275.1005	C11H17NO7			Nucleotide sugar biosynthesis
F054	306.0669	Deoxyguanosine	C00330		10	M+K	306.0599	267.0968	C10H13N5O4	2.0087	1.0063	Purine metabolism
		Adenosine	C00212	M00958	10	M+K	306.0599	267.0968	C10H13N5O4			Adenine ribonucleotide degradation, purine metabolism
		D-Glycero-D-mannoheptose 1-phosphate	C07838	M00064	15	M+NH4	306.0585	288.0246	C7H13O10P			ADP-L-glycero-D-manno-heptose biosynthesis
		2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate	C04691	M00022	15	M+NH4	306.0585	288.0246	C7H13O10P			Shikimate pathway
MG1655	308.13193	N-Succinyl-L,L-2,6-diaminopimelate	C04421	M00016	43	M+NH4	308.1452	290.1114	C11H18N2O7	0.42316	-1.2407	Lysine biosynthesis
ELU39	308.13545	N-Succinyl-L,L-2,6-diaminopimelate	C04421	M00016	32	M+NH4	308.1452	290.1114	C11H18N2O7	0.46571	-1.1025	Lysine biosynthesis

MG1655	309.03751	5'-Phosphoribosylglycinamide	C03838	M00048	27	M+Na	309.0458	286.0566	C7H15N2O8P	0.31888	-1.6489	De novo purine biosynthesis
		dUMP	C00365	M00938	35	M+H	309.0482	308.041	C9H13N2O8P			Pyrimidine biosynthesis
MG1655	310.12643	N-acetylneuraminate	C00270		42	M+H	310.1133	309.106	C11H19NO9	2.3554	1.236	Nucleotide sugar metabolism
F022	310.12735	N-acetylneuraminate	C00270		45	M+H	310.1133	309.106	C11H19NO9	5.3758	2.4265	Nucleotide sugar metabolism
ELU39	310.12870	N-acetylneuraminate	C00270		50	M+H	310.1133	309.106	C11H19NO9	3.4383	1.7817	Nucleotide sugar metabolism
F054	314.08428	Phosphoribosylformylglycineamide	C04640	M00048	30	M+H	314.0748	313.0675	C8H16N3O8P	4.8928	2.2906	De novo purine biosynthesis
ELU39	331.00252	alpha-D-Ribose 1-methylphosphonate 5-phosphate	C20423		21	M+Na	330.9954	308.0062	C6H14O10P2	0.42935	-1.2198	Energy intermediates
F022	332.06653	N-Acetyl-D-muramoate	C02713		23	M+K	332.0742	293.1111	C11H19NO8	423.59	8.7265	Precursor for peptidoglycan biosynthesis
		dAMP	C00360		27	M+H	332.0754	331.0682	C10H14N5O6P			ADP derivative
F054	332.06662	N-Acetyl-D-muramoate	C02713		23	M+K	332.0742	293.1111	C11H19NO8	3.4326	1.7793	Precursor for peptidoglycan biosynthesis
		dAMP	C00360		27	M+H	332.0754	331.0682	C10H14N5O6P			ADP derivative
MG1655	332.06715	N-Acetyl-D-muramoate	C02713		21	M+K	332.0742	293.1111	C11H19NO8	841.6	9.717	Precursor for peptidoglycan biosynthesis
		dAMP	C00360		25	M+H	332.0754	331.0682	C10H14N5O6P			ADP derivative
ELU39	332.06765	N-Acetyl-D-muramoate	C02713		20	M+K	332.0742	293.1111	C11H19NO8	628.07	9.2948	Precursor for peptidoglycan biosynthesis
		dAMP	C00360		23	M+H	332.0754	331.0682	C10H14N5O6P			ADP derivative
F022	347.07336	Adenosine 2',3'-cyclic	C02353		37	M+NH4	347.0863	329.0525	C10H12N5	0.49262	-1.0215	Precursor to adenosine,

		phosphate							O6P			derivative of 2'3'-Cyclic AMP
		Cyclic AMP	C00575		37	M+NH4	347.0863	329.0525	C10H12N5O6P			Biofilm Formation
F054	348.06611	N-acetylneuraminate	C00270		9	M+H	310.1133	309.106	C11H19NO9	2.0074	1.0053	Precursor to adenosine, derivative of 2'3'-Cyclic AMP
		AMP	C00020		12	M+H	348.0704	347.0631	C10H14N5O7P			Purine metabolism
		3'-AMP	C01367	M00049 M00958	12	M+H	348.0704	347.0631	C10H14N5O7P			Adenine ribonucleotide biosynthesis and degradation
F054	354.04809	dAMP	C00360		26	M+Na	354.0574	331.0682	C10H14N5O6P	5.2314	2.3872	ADP derivative
MG1655	354.04857	dAMP	C00360		25	M+Na	354.0574	331.0682	C10H14N5O6P	83.557	6.3847	ADP derivative
F022	354.05001	dAMP	C00360		21	M+Na	354.0574	331.0682	C10H14N5O6P	12.156	3.6036	ADP derivative
MG1655	354.05018	dAMP	C00360		20	M+Na	354.0574	331.0682	C10H14N5O6P	6.601	2.7227	ADP derivative
ELU39	354.05025	dAMP	C00360		20	M+Na	354.0574	331.0682	C10H14N5O6P	5.1729	2.371	ADP derivative
F054	355.12234	S-(Hydroxymethyl)glutathione	C14180		16	M+NH4	337.0944	355.1282	C11H19N3O7S	4.7877	2.2593	Methane metabolism
ELU39	355.12236	S-(Hydroxymethyl)glutathione	C14180		16	M+NH4	337.0944	355.1282	C11H19N3O7S	14.059	3.8135	Methane metabolism
ELU39	359.10149	N1-(5-Phospho-a-D-ribose)-5,6-dimethylbenzimidazole	C04778		3	M+H	359.1003	358.093	C14H19N2O7P	0.37619	-1.4105	
		Pantetheine 4'-phosphate	C01134	M00120	6	M+H	359.1036	358.0964	C11H23N2O7PS			Coenzyme A biosynthesis
F022	377.10660	Pantetheine 4'-phosphate	C01134	M00120	39	M+Na	377.0918	358.0964	C11H23N2O7PS	2.2189	1.1499	Coenzyme A biosynthesis
MG1655	377.14328	Riboflavin	C00255	M00125	6	M+H	377.1456	376.1383	C17H20N4	0.45113	-1.1484	Riboflavin biosynthesis

									O6			
		S-Adenosylmethionine	C01137		16	M+Na	377.1372	355.1552	C14H23N6O3S			Methionine salvage
F054	380.11626	S-Lactoylglutathione	C03451		11	M+H	380.1122	379.1049	C13H21N3O8S	0.40164	-1.316	Pyruvate metabolism
F022	393.08125	Pantetheine 4'-phosphate	C01134	M00120	40	M+K	393.0657	358.0964	C11H23N2O7PS	2.0558	1.0397	Coenzyme A biosynthesis
F054	423.09437	Trehalose 6-phosphate	C00689		11	M+H	423.0898	422.0825	C12H23O14P	0.42387	-1.2383	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		23	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation
MG1655	423.09528	Trehalose 6-phosphate	C00689		13	M+H	423.0898	422.0825	C12H23O14P	0.26025	-1.942	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		25	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation
ELU39	423.09606	Trehalose 6-phosphate	C00689		15	M+H	423.0898	422.0825	C12H23O14P	0.22658	-2.1419	Other carbohydrate metabolism
		S-Adenosylhomocysteine	C00021		27	M+K	423.0847	384.1216	C14H20N6O5S			Methionine degradation
F054	442.14369	Folate	C00504		7	M+H	442.147	441.1397	C19H19N7O6	0.33397	-1.5822	Folate metabolism
MG1655	445.07635	Trehalose 6-phosphate	C00689		10	M+Na	423.0898	422.0825	C12H23O14P	0.31285	-1.6764	Other carbohydrate metabolism
		dGDP	C00361	M00053	29	M+NH4	445.0632	427.0294	C10H15N5O10P2			Deoxyribonucleotide biosynthesis
		ADP	C00008	M00053	20	M+NH4	445.0632	427.0294	C10H15N5O10P2			Deoxyribonucleotide biosynthesis
ELU39	445.07657	Trehalose 6-phosphate	C00689		11	M+Na	423.0898	422.0825	C12H23O14P	0.35723	-1.4851	Other carbohydrate metabolism
		dGDP	C00361	M00053	30	M+NH4	445.0632	427.0294	C10H15N5O10P2			Deoxyribonucleotide biosynthesis
		ADP	C00008	M00053	30	M+NH4	445.0632	427.0294	C10H15N5O10P2			Deoxyribonucleotide biosynthesis

MG1655	483.99196	CTP	C00063	M00052	0	M+H	483.9918	482.9845	C9H16N3O 14P3	3.2484	1.6997	Pyrimidine ribonucleotide biosynthesis
--------	-----------	-----	--------	--------	---	-----	----------	----------	------------------	--------	--------	--

S5 – Mass Spectrometry Parameters

Parameters for C:\MassLynx\Heather.PRO\ACQUDB\esi_pos_sens_3min_600.EXP
Created by 4.2 SCN983

Lock Spray Configuration:

Tuning on Analyte

Temperature Correction:

Temperature Correction Disabled

Instrument Configuration:

Lteff	1800.0
Veff	7199.60
Resolution	10000
Min Points in Peak	2
Acquisition Device	WatersADC
Acquisition Algorithm	ADC Mode
ADC Trigger Threshold (V)	1.00
ADC Input Offset (V)	-1.50
Average Single Ion Intensity	19
ADC Amplitude Threshold	3
ADC Centroid Threshold	-1
ADC Ion Area Threshold	3
ADC Ion Area Offset	10
ADC Pushes Per IMS Increment	1
EDC Delay Coefficient	1.4100
EDC Delay Offset	0.4000

Experimental Instrument Parameters

Instrument Parameter Filename

C:\MassLynx\Heather.PRO\ACQUDB\esi060423.IPR (MODIFIED)

Polarity ES+

Capillary (kV)	2.4300
Source Temperature (°C)	100
Sampling Cone	9.0000
Source Offset	28.0000
Source Gas Flow (mL/min)	0.00
Desolvation Temperature (°C)	280
Cone Gas Flow (L/Hr)	0.0
Desolvation Gas Flow (L/Hr)	731.0
Nebuliser Gas Flow (Bar)	7.0
LM Resolution	4.4
HM Resolution	15.0
Aperture 1	0.0
Pre-filter	2.0
Ion Energy	0.2
Manual Trap Collision Energy	FALSE
Trap Collision Energy	4.0
Manual Transfer Collision Energy	FALSE
Transfer Collision Energy	2.0
Manual Gas Control	FALSE
Trap Gas Flow (mL/min)	2.00
HeliumCellGasFlow	180.00
IMS Gas Flow (mL/min)	90.00
Detector	3300
DetectorCache	0
Sample Infusion Flow Rate (µL/min)	5
Sample Flow State	LC
Sample Fill Volume (µL)	250
Sample Reservoir	Wash
LockSpray Infusion Flow Rate (µL/min)	10
LockSpray Flow State	Infusion
LockSpray Reservoir	B
LockSpray Capillary (kV)	3.0
Use Manual LockSpray Collision Energy	FALSE

Collision Energy	4.0
Acceleration1	70.0
Acceleration2	200.0
Aperture2	40.0
Transport1	70.0
Transport2	70.0
Steering	-0.20
Tube Lens	45
Pusher	1900.0
Pusher Offset	-0.28
Puller	1370.0
Pusher Cycle Time (μ s)	Automatic
Pusher Width (μ s)	Automatic
Collector	60
Collector Pulse	10.0
Stopper	10
Stopper Pulse	20.0
Entrance	62
Static Offset	180
Puller Offset	0.00
Reflectron Grid (kV)	1.471
Flight Tube (kV)	10.00
Reflectron (kV)	3.780
Use Manual Trap DC	FALSE
Trap DC Entrance	1.0
Trap DC Bias	2.0
Trap DC	-2.0
Trap DC Exit	0.0
Use Manual IMS DC	FALSE
IMS DC Entrance	-20.0
Helium Cell DC	1.0
Helium Exit	-20.0
IMSBias	2.0

IMS DC Exit	20.0	
USe Manual Transfer DC		FALSE
Transfer DC Entrance	5.0	
Transfer DC Exit	15.0	
Trap Manual Control		OFF
Trap Wave Velocity (m/s)	300	
Trap Wave Height (V)	0.5	
IMS Manual Control		OFF
IMS Wave Velocity (m/s)	300	
IMS Wave Height (V)	0.0	
Transfer Manual Control		OFF
Transfer Wave Velocity (m/s)	247	
Transfer Wave Height (V)	0.2	
Step Wave 1 In Manual Control		OFF
Enable Reverse Operation		OFF
Step Wave 1 In Velocity (m/s)	300.0	
Step Wave 1 In Height	15.0	
Step Wave 1 Out Manual Control		OFF
Step Wave 1 Out Velocity (m/s)	300.0	
Step Wave 1 Out Height	15.0	
Step Wave 2 Manual Control		OFF
Step Wave 2 Velocity (m/s)	300.0	
Step Wave 2 Height	1.0	
Use Manual Step Wave DC		OFF
Step Wave TransferOffset	25.0	
Step Wave DiffAperture1	3.0	
Step Wave DiffAperture2	-0.0	
Use Automatic RF Settings		TRUE
StepWave1RFOffset	300.0	
StepWave2RFOffset	350.0	
Target Enhancement Enabled		FALSE
Target Enhancement Mode		EDC
Target Enhancement Mass		785.0

Target Enhancement Trap Height (V)	4.0
Target Enhancement Extract Height (V)	15.0
Mobility Trapping Manual Release Enabled	FALSE
Mobility Trapping Release Time (μ s)	500
Mobility Trap Height (V)	15.0
Mobility Extract Height (V)	0.0
Trag Gate LUT table enabled	FALSE
TriWave Trap Gate LookUp Table	
Using Drift Time Trimming	FALSE
Drift Time Bins	0
Using Mobility Delay after Trap Release	TRUE
IMS Wave Delay (μ s)	1000
Variable Wave Height Enabled	FALSE
Wave Height Ramp Type	Linear
Wave Height Start (V)	10.0
Wave Height End (V)	40.0
Wave Height Using Full IMS	TRUE
Wave Height Ramp (%)	100.0
Wave Height Look Up Table	
Variable Wave Velocity Enabled	FALSE
Wave Velocity Ramp Type	Linear
Wave Velocity Start (m/s)	1000.0
Wave Velocity End (m/s)	300.0
Wave Velocity Using Full IMS	TRUE
Wave Velocity Ramp (%)	100.0
Wave Velocity Look Up Table	
Backing	2.75e0
Source	6.85e-3
Sample Plate	1.24e3
Trap	8.83e-3
Helium Cell	1.00e-4
IMS	9.90e-5
Transfer	7.85e-3

TOF	4.52e-7
IMSRFOffset	300
IMSMobilityRFOffset	250
TrapRFOffset	300
Use Automatic RF Settings	TRUE
AutoStepWave1RFOffset	300
AutoStepWave2RFOffset	300
TransferRFOffset	350
MS Profile Type	Auto P
MSProfileMass1	100
MSProfileDwellTime1	20
MSProfileRampTime1	20
MSProfileMass2	300
MSProfileDwellTime2	20
MSProfileRampTime2	40
MSProfileMass3	500
PusherInterval	39.000000
PusherOffset	0.250000
LockMassValidSigma	5

Acquisition mass range

Start mass	50.000
End mass	600.000

Calibration mass range

Start mass	0.000
End mass	0.000

Experiment Reference Compound Name: N/A

Function Parameters - Function 1 - TOF MS FUNCTION

Scan Time (sec)	1.000
Interscan Time (sec)	0.014
Start Mass	50.0

End Mass	600.0
Start Time (mins)	0.00
End Time (mins)	3.00
Data Format	Continuum
Analyser	Sensitivity Mode
ADC Sample Frequency (GHz)	3.0
ADC Pusher Frequency (μ s)	39.0
ADC Pusher Width (μ s)	1.00
Use Tune Page Cone Voltage	YES
Using Auto Trap Collision Energy (eV)	4.000000
Using Auto Transfer Collision Energy (eV)	2.000000
Sensitivity	Normal
Dynamic Range	Normal
Save Collapsed Retention Time Data	No
Use Rule File Filtering	No
FragmentationMode	CID
Calibration	Dynamic 2