



Libraries and Learning Services

University of Auckland Research Repository, ResearchSpace

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

Watching Bacteria Adapt

Investigating How Pathogens Evolve to Cause Disease

Hannah Marie Read

*A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy
in Molecular Medicine, The University of Auckland, 2016.*

Abstract

Experimental evolution has provided us with many insights into varied evolutionary processes, however there has been a lack of studies seeking to follow natural infection and transmission of a pathogen through its natural host. I have followed the adaptation of the murine enteropathogen *Citrobacter rodentium* as it infects laboratory mice. Ten independent lineages were started, with mice from five of the lineages receiving a low dose of the quinolone nalidixic acid in their drinking water to alter their normal microbiome, and mice from the remaining five lineages receiving untreated drinking water. Mice were orally inoculated with the bioluminescent *C. rodentium* derivative ICC180 and individually housed animals allowed to infect naïve animals through tightly controlled mouse-to-mouse exposure, a process which was repeated weekly over a period of five months, by which time at least 20 transmissions had occurred. The *in vivo*-adapted *C. rodentium* were isolated from the mice and compared and competed with the ancestral strain. Multiple phenotypic and genotypic differences were observed, including the evolution of a ‘hypertransmissible’ isolate which preferentially transmits to naïve animals when in competition with the wildtype strain, an isolate which gained the ability to form aggregates in rich media, and a hypermutable isolate. In addition to the *in vivo* evolution experiment, a complementary *in vitro* evolution experiment was performed, with *C. rodentium* passaging through laboratory media using the traditionally established method of serial transfer. These *in vitro*-adapted isolates adapted to a restricted media with 1% glucose supplementation, and half of the isolates exhibited a trade-off of reduced growth on rich media accompanying improved growth on the restricted media. Such a trade-off was not observed in the *in vivo*-adapted *C. rodentium*. In conclusion, a five month period of experimental evolution was sufficient for observable adaptation to occur, with separate evolutionary trajectories for *C. rodentium* adapting *in vivo* versus *in vitro*. Populations have been stored for future analyses, and the bank of samples will be an important resource for evolutionary biologists. The work detailed in this thesis extends existing flask-based experimental evolution with an *in vivo* model following the entire infectious process.

Acknowledgements

For their support and encouragement, I would like to recognise the following people

(in no particular order)

for contributions to my sanity:

My supervisor, boss, mentor, and friend: **Siouxsie Wiles**,

who inspired me to become a scientist

and whose infectious enthusiasm spurred action and innovation.

The ever-encouraging members of the Wiles lab group,

in particular **Benedict Uy** and **Jimmy Dalton**.

I do not wish to even imagine how different this journey may have been
without their support and late-night pep talks.

The friends and family who have supported me,

with a special mention to my Dad, **Clive Read**,

for encouraging me to ignore everyone else, to do a PhD,
and to not quit.

Table of Contents

List of Figures.....	vii
List of Tables	ix
Chapter 1: Literature Review.....	1
1.1 Introduction	1
1.2 Experimental Evolution.....	2
1.2.1 Experimental evolution to date: <i>in vitro</i>	3
1.2.2 Experimental evolution to date: <i>in vivo</i>	11
1.2.3 Pathogen Evolution	13
1.2.4 The gap in the literature	13
1.3 <i>Citrobacter rodentium</i>	15
1.3.1 Diarrhoeal Illness	15
1.3.2 The History of <i>Citrobacter rodentium</i>	15
1.3.3 Transmissible Murine Colonic Hyperplasia.....	18
1.3.4 <i>C. rodentium</i> virulence factors	19
1.3.5 <i>C. rodentium</i> is still in the process of adapting to laboratory mice	23
1.4 Thesis Aims.....	25
Chapter 2: Materials and Methods.....	26
2.1 Introduction	26
2.2 Materials.....	27
2.2.1 The pRhomo plasmid	32
2.2.2 Antibiotics.....	34
2.3 Methods.....	35
2.3.1 Bacterial strains and culture conditions.....	35
2.3.2 Mice	35
2.3.3 Infection of mice	35
2.3.4 <i>In vivo</i> biophotonic imaging.....	36
2.3.5 <i>In vivo</i> Evolution Experiment	36
2.3.6 <i>In vivo</i> competition experiments	37
2.3.7 Evolution of <i>C. rodentium</i> ICC180 to defined laboratory media	37
2.3.8 Attachment of <i>Citrobacter rodentium</i> to mouse fibroblast cells	37
2.3.9 Evaluation of desiccation resistance.....	38
2.3.10 Bacterial staining.....	38
2.3.11 Assessment of spontaneous mutation rate	39
2.3.12 Assessment of recombination rate using pRhomo	39
2.3.13 <i>In vitro</i> growth curves.	39
2.3.14 Genome sequencing	40
2.3.15 Statistical analyses	40

Chapter 3: Comparison of <i>C. rodentium</i> strains ICC169 and ICC180.....	41
3.1 Introduction	41
3.2 The <i>in vitro</i> and <i>in vivo</i> effects of constitutive light expression on a bioluminescent strain of the mouse enteropathogen <i>Citrobacter rodentium</i>	43
3.2.1 Introduction.....	43
3.2.2 Materials and methods	45
3.2.3 Results.....	50
3.3.4 Discussion	59
3.3.5 Supplementary Tables.....	62
3.3.6 Supplementary Figures.....	70
3.3 Conclusions	73
Chapter 4: Experimental <i>in vivo</i> Evolution of <i>C. rodentium</i>	74
4.1 Introduction	74
4.2 Methods.....	76
4.3 Results	77
4.3.1 Impact of nalidixic acid treatment on the mouse gut microbiota	77
4.3.2 Bacterial shedding during the <i>in vivo</i> evolution experiment and the occurrence of transmission failure events ..	79
4.3.3 Comparison of <i>C. rodentium</i> ICC180 levels shed by mice during the <i>in vivo</i> evolution experiment	82
4.3.4 Biophotonic imaging of <i>C. rodentium</i> ICC180 from within infected mice.....	86
4.3.5 No detectable change in the severity of disease	93
4.4 Discussion	94
4.4.1 High rates of successful transmission throughout the <i>in vivo</i> evolution experiment.....	94
4.4.2 No loss of bioluminescence was observed following <i>in vivo</i> evolution.....	94
4.4.3 Presence of nalidixic acid associated with reduction in the number of <i>C. rodentium</i> shed from animals.....	95
4.4.5 Conclusions	96
Chapter 5: <i>In vivo</i> assessment of <i>in vivo</i> -adapted <i>C. rodentium</i>	97
5.1 Introduction	97
5.2 Results	98
5.2.1 <i>In vivo</i> infection dynamics of the <i>in vivo</i> -adapted <i>C. rodentium</i> strains after oral gavage of mice	98
5.2.2 The <i>in vivo</i> -adapted <i>C. rodentium</i> strains are differentially shed from animals co-infected with ICC169	100
5.2.3 Differences in bacterial burden as measured by bioluminescence readings in co-infected animals	106
5.2.4 The <i>in vivo</i> -adapted strains show changes in transmissibility compared to the ancestral ICC180 strain	107
5.3 Discussion	109
5.3.1 Adaptation to mice results in changes in the numbers of <i>C. rodentium</i> shed in single infections in 3/10 lineages	109
5.3.2 Shedding and transmissibility of <i>in vivo</i> -adapted <i>C. rodentium</i> when in competition with ICC169 are not necessarily linked.....	110
5.3.3 Treatment of mice with nalidixic acid does not appear to influence the trajectory of <i>in vivo</i> adaptation as measured by transmissibility and bacterial shedding	111

5.3.4 Shedding and <i>in vivo</i> bacterial burden, as assessed by bioluminescence, are not necessarily correlated.....	112
5.3.5 Conclusions.....	115
Chapter 6: <i>In vitro</i> assessment of <i>in vivo</i> -adapted <i>C. rodentium</i>	116
6.1 Introduction	116
6.2 Results	117
6.2.1 Emergence of an aggregating <i>C. rodentium</i> strain following <i>in vivo</i> adaptation	117
6.2.2 Aggregating <i>C. rodentium</i> N5 _{P20} does not disperse in the presence of cellulase.....	120
6.2.3 No obvious microscopic variation between evolved <i>C. rodentium</i> and the ancestor	122
6.2.4 No change in tolerance to desiccation in <i>in vivo</i> -adapted <i>C. rodentium</i>	124
6.2.5 No change in the ability of <i>in vivo</i> -adapted <i>C. rodentium</i> strains to attach to mouse fibroblast cells	125
6.2.6 Emergence of a hypermutable strain of <i>C. rodentium</i> following <i>in vivo</i> adaptation	127
6.2.7 No detectable change in recombination rate for <i>in vivo</i> -adapted <i>C. rodentium</i>	129
6.3 Discussion	130
6.3.1 <i>In vivo</i> -adapted <i>C. rodentium</i> show no changes in desiccation tolerance or attachment to mouse fibroblast cells	130
6.3.2 Changes in growth in rich and restricted media	131
6.3.3 One strain has evidence of increased mutation rate, no evidence of recombination rate change	132
6.3.4 Conclusions.....	133
Chapter 7: Genomic changes in <i>in vivo</i> -adapted <i>C. rodentium</i>	134
7.1 Introduction	134
7.2 Results	135
7.2.1 Single nucleotide polymorphisms and other genetic changes	136
7.2.2 Genes with multiple genetic changes present.....	142
7.2.3 Intragenic SNPs constitute the majority of SNPs and are found throughout the genome	145
7.2.4 Genetic changes in <i>C. rodentium</i> adapted to the mouse environment.....	147
7.2.5 Genetic changes in <i>C. rodentium</i> adapted to the mouse environment in the presence of nalidixic acid	155
7.2.6 Categorisation of the genes mutated in the <i>in vivo</i> -adapted <i>C. rodentium</i> strains.....	162
7.2.7 Little evidence of foreign DNA uptake in <i>in vivo</i> -adapted <i>C. rodentium</i>	164
7.2.8 Minimal changes to plasmids present in <i>C. rodentium</i>	165
7.3 Discussion	166
7.3.1 Genes known to be important for <i>C. rodentium</i> virulence are conserved	166
7.3.2 Mutations present in genes encoding for hypothetical proteins	167
7.3.3 The Type 6 Secretion System <i>cts2S</i> gene is mutated in the majority of “N” strains	167
7.3.4 Genes involved in ethanolamine metabolism are frequently altered.....	167
7.3.5 Similar genetic changes present in strains from distinct lineages	168
7.3.6 The plasmids present in <i>C. rodentium</i> appear to be stable	168
7.3.7 Conclusions.....	169

Chapter 8: Experimental <i>in vitro</i> Evolution of <i>C. rodentium</i>	170
8.1 Introduction	170
8.2 Methods.....	171
8.3 Results	172
8.3.1 Reproducible loss of bioluminescence following evolution of ICC180 in laboratory media.....	172
8.3.2 <i>In vitro</i> -adapted <i>C. rodentium</i> has improved growth in DM media	172
8.3.3 <i>In vitro</i> -adapted <i>C. rodentium</i> has decreased growth in rich media.....	174
8.4 Discussion	176
8.4.1 Light production appears to confer no benefit <i>in vitro</i>	176
8.4.2 Evidence of a growth trade-off for <i>in vitro</i> -adapted <i>C. rodentium</i>	177
8.4.3 Conclusions.....	177
Chapter 9: Discussion and Future Directions.....	178
9.1 The differences between <i>in vivo</i> and <i>in vitro</i> evolution	178
9.1.1 Loss of bioluminescence was observed following adaptation to laboratory media, and not following adaptation to a mouse	178
9.1.2 Evidence of growth trade-offs in <i>in vitro</i> -adapted <i>C. rodentium</i> absent in <i>in vivo</i> -adapted <i>C. rodentium</i>	179
9.2 The effect of nalidixic acid treatment on the evolution of <i>C. rodentium</i>	180
9.3 Phenotypes and genotypes observed for each of the <i>in vivo</i> -adapted lineages	181
9.3.1 W1 _{P20} has improved transmissibility and a history of reduced light from the rectum	181
9.3.2 W2 _{P20} has improved transmissibility and a history of reduced light from the rectum	182
9.3.3 W3 _{P20} performs the same <i>in vivo</i> as the ancestor and has the greatest number of genetic changes	183
9.3.4 W4 _{P20} has improved transmissibility and no notable unique intragenic SNPs	184
9.3.5 W5 _{P20} is shed at higher levels and has increased growth in DM media.....	185
9.3.6 N1 _{P20} has improved transmissibility and improved growth in DM media.....	186
9.3.7 N2 _{P20} has improved transmissibility while being shed at lower numbers.....	187
9.3.8 N3 _{P20} has decreased shedding and an increased mutation rate	188
9.3.9 N4 _{P20} transmits preferentially to ICC169 <i>in vivo</i> and has increased growth in DM media	189
9.3.10 N5 _{P20} has improved transmissibility and aggregates in LB media.....	190
9.4 Future directions	192
9.5 Conclusions.....	194
Appendices	195
References.....	200

List of Figures

Figure 1.1. Number of published experimental evolution papers is increasing.	2
Figure 1.2. <i>C. rodentium</i> attachment to enterocytes.	20
Figure 2.1. The pRhomo plasmid before and after recombination	33
Figure 4.1. Schematic of the <i>Citrobacter rodentium</i> <i>in vivo</i> evolution experiment.	76
Figure 4.2. <i>Citrobacter rodentium</i> successfully transmits in controlled mouse-to-mouse infections over a period of 5 months (overleaf).....	79
Figure 4.3. <i>Citrobacter rodentium</i> is shed at relatively stable levels throughout the <i>in vivo</i> experimental evolution experiment (overleaf).....	82
Figure 4.4. Administration of low-dose nalidixic acid in the drinking water does not change the overall amount or peak levels of <i>Citrobacter rodentium</i> shed over the 5-month evolution experiment.	85
Figure 4.5. Representative images showing detectable bioluminescent <i>Citrobacter rodentium</i> from within anaesthetised mice and the regions of interest measured.	86
Figure 4.6. Amount of light detected from each region of interest over the course of the evolution experiment does not change in the majority of lineages (overleaf).....	87
Figure 4.7. AUC values for light detected from the abdomen and rectum of mice vary between lineages.	90
Figure 4.8. The relationship between light detected from infected mice and the amount of <i>Citrobacter rodentium</i> recovered from stools varies between different evolutionary lineages (overleaf)	91
Figure 4.9. No change in disease severity as measured by change in mouse weight.....	93
Figure 5.1. The <i>in vivo</i> -adapted <i>C. rodentium</i> strains W1 and W5 are shed at higher numbers compared with mice infected with the ancestral ICC180 strain.....	99
Figure 5.2. Schematic of the co-infection experiments.....	101
Figure 5.3. Competition of the <i>in vivo</i> -adapted <i>C. rodentium</i> strains evolved in the absence of antibiotics with ICC169 reveals fitness advantages in some lineages (overleaf).....	102
Figure 5.4. Competition of the <i>in vivo</i> -adapted <i>C. rodentium</i> strains evolved in the presence of nalidixic acid with ICC169 reveals fitness advantages in some lineages (overleaf).....	102
Figure 5.5. Competition indices reveal varying fitness of <i>in vivo</i> -adapted <i>C. rodentium</i> when competed head-to-head with ICC169 in mice (overleaf).	102
Figure 5.6. Higher levels of lights detected from mice infected with two <i>in vitro</i> -adapted strains when competed with ICC169.....	106
Figure 6.1. Some lineages of <i>in vivo</i> -adapted <i>Citrobacter rodentium</i> show improved growth in minimal media (overleaf).	117
Figure 6.2. Emergence of aggregates in the N5 _{P20} lineage of <i>in vivo</i> -adapted <i>Citrobacter rodentium</i> in rich media (overleaf).	117
Figure 6.3. N5 _{P20} aggregates fluoresce under UV light when stained with calcofluor white.....	121
Figure 6.4. <i>In vivo</i> -adapted <i>Citrobacter rodentium</i> show no morphological changes under light microscopy.	123
Figure 6.5. <i>In vivo</i> -adapted <i>Citrobacter rodentium</i> show no change in tolerance to desiccation.	124
Figure 6.6. No change in ability of <i>in vivo</i> -adapted <i>Citrobacter rodentium</i> strains to attach to mouse L929 fibroblast cells.	126
Figure 6.7. One lineage of <i>in vivo</i> -adapted <i>Citrobacter rodentium</i> has an increased mutation rate to two separate antibiotics.	128
Figure 6.8. <i>in vivo</i> -adapted <i>Citrobacter rodentium</i> shows no apparent change in recombination rate.....	129
Figure 7.1. PROVEAN scores for all SNPs and unique SNPs (overleaf).....	145

Figure 7.2. Proportions SNPs in genes with estimated functions.	163
Figure 8.1. Schematic of the <i>Citrobacter rodentium</i> <i>in vitro</i> evolution experiment.	171
Figure 8.2. <i>In vitro</i> -adapted <i>Citrobacter rodentium</i> shows improved growth in minimal media.	173
Figure 8.3. <i>In vitro</i> -adapted <i>Citrobacter rodentium</i> shows reduced growth in rich media.	175

List of Tables

Table 1.1. Benefits of using mice and <i>C. rodentium</i> as a host-pathogen model	14
Table 1.2. <i>C. rodentium</i> outbreaks.....	17
Table 1.3. Mouse knockouts which can or cannot clear <i>C. rodentium</i> infection	18
Table 1.4. <i>C. rodentium</i> genes important for virulence and their putative functions	22
Table 2.1. Bacterial Strains.....	27
Table 2.2. Animals.....	28
Table 2.3. Eukaryotic cell lines.....	28
Table 2.4. Media and Solution Recipes	29
Table 2.5. Chemicals and Media.....	30
Table 2.6. Stains	30
Table 2.7. Anaesthetics.....	31
Table 2.8. Enzymes.....	31
Table 2.9. DNA isolation	32
Table 2.10. Sample Storage	32
Table 2.11. Antibiotics.....	34
Table 3.1. Characteristics of the <i>Citrobacter rodentium</i> strains used in this study.....	42
Table 4.1. OTUs with a significant difference in relative abundance after 24 hours treatment with nalidixic acid	78
Table 4.2. Slopes of AUC of CFU shed in each transmission chain.....	84
Table 4.3. Linear regression of total flux for abdominal and rectal regions of interest	89
Table 5.1. Changes in shedding at the peak of infection and varying transmission improvements in <i>in vivo</i> -adapted <i>C. rodentium</i> when competed with ICC169.....	108
Table 5.2. Summary of observed changes of <i>in vivo</i> -adapted <i>C. rodentium</i> for <i>in vivo</i> assays	114
Table 6.1. Summary of observed changes of <i>in vivo</i> -adapted <i>C. rodentium</i> for <i>in vitro</i> assays	133
Table 7.1. Intergenic and intragenic genetic changes present at $\geq 25\%$ frequency.....	137
Table. 7.2. List of genes with multiple genetic events.....	144
Table 7.3. Percentage of intragenic and unique intragenic SNPs.	145
Table 7.4. Mutations in W1 _{P20} genome.....	147
Table 7.5. Mutations in W2 _{P20} genome.....	149
Table 7.6. Mutations in W3 _{P20} genome.....	150
Table 7.7. Mutations in W4 _{P20} genome.....	152
Table 7.8. Mutations in W5 _{P20} genome.....	154
Table 7.9. Mutations in N1 _{P20} genome.....	155
Table 7.10. Mutations in N2 _{P20} genome.....	157
Table 7.11. Mutations in N3 _{P20} genome.....	158
Table 7.12. Mutations in N4 _{P20} genome.....	159
Table 7.13. Mutations in N5 _{P20} genome.....	161
Table. 7.15. Genetic changes to pCROD2 in <i>in vivo</i> -adapted <i>C. rodentium</i>	165
Table 8.1. Summary of observed changes of <i>in vitro</i> -adapted <i>C. rodentium</i>	176

Co-Authorship Form

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 3

Nature of contribution by PhD candidate	Experimental work, writing, design
Extent of contribution by PhD candidate (%)	80

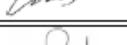
CO-AUTHORS

Name	Nature of Contribution
Grant Mills	Galleria mellonella infections
Sarah Johnson	Technical assistance with mouse infections
Peter Tsai	Bioinformatics assistance and location of insertion
Dr. James Dalton	Confirmation of SNPs by PCR
Dr. Lars Barquist	Statistical analysis of BIOLOG Phenotypic Microarray plates
Dr. Cris Print	Bioinformatics support

Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Grant Mills		08/04/2016
Sarah Johnson		06/04/2016
Peter Tsai		15/04/2016
Dr. James Dalton		08/04/2016
Dr. Lars Barquist		05/04/2016
Dr. Cris Print		15/04/2016

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 3

Nature of contribution by PhD candidate	Experimental work, writing, design
Extent of contribution by PhD candidate (%)	80

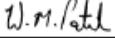
CO-AUTHORS

Name	Nature of Contribution
Dr. Wayne Patrick	Biochemical effects discussion and protein modelling
Dr. Siouxsie Wiles	Writing, editing, experimental design

Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Dr. Wayne Patrick		10/04/2016
Dr. Siouxsie Wiles		28/07/2016

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 4

Nature of contribution by PhD candidate	Experimental work, writing, design
Extent of contribution by PhD candidate (%)	90

CO-AUTHORS

Name	Nature of Contribution
Elahe Kia	16S sequence extraction and sequencing analysis
Associate Professor Mike Taylor	16S sequence analysis

Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Elahe Kia		30/05/2016
Associate Professor Mike Taylor		30/05/2016

Chapter 1: Literature Review

1.1 Introduction

Since the first human glimpsed bacteria through the lens of a microscope¹, our study of microorganisms has led us to a greater understanding and appreciation for the bacteria around us. Microorganisms have a huge impact on humanity, playing a profound and varied role in our lives, from providing us with nutrients² to causing disease and suffering; our commensals may even influence the ultimate length of our lives³.

Our close relationship with our microbiome has important implications for our health and wellbeing from the day we are born, with babies delivered via caesarean section colonised by different types of bacteria and having a greater risk of developing allergies compared with those delivered via vaginal birth⁴, to how we live our lives, with an altered gut microbiota associated with type I diabetes⁵. It has recently been shown that in mice with a genetic propensity for developing obesity, changing their microbiota can make them resistant to developing metabolic syndromes⁶. Between individuals there is a great amount of diversity between our microbiota⁷, and this diversity has the potential to distinguish us and our microbiome from other individuals⁷, with interpersonal differences dwarfing intrapersonal differences⁸. While some bacteria are harmful and cause disease, others play an important role in our defence against these pathogens. For example, *Clostridium scindens* produces secondary bile acids that prevent its relative *C. difficile* from infecting its host⁹.

However, not all of our encounters with bacteria are mutually beneficial. Our relationship with pathogenic bacteria has been long, with *Mycobacterium tuberculosis*, the causative agent of the infection commonly known as tuberculosis (TB), found in human bones 9,000 years old¹⁰. Other illnesses are more recent, and new diseases have emerged. The causative agent of the Plague, *Yersinia pestis*, evolved as recently as 1,500 years ago¹¹. In 2011, a strain of Enteropathogenic *E. coli* (EAEC) which had acquired the Shiga toxin gene via horizontal gene transfer became highly virulent and caused severe diarrhoea in over 3,000 people¹²⁻¹⁴. Evolution is clearly an ongoing process.

This research's unique contribution to the scientific field is the development, execution, and analysis of a novel 5-month *in vivo* evolution experiment, following the evolution of a bacterium still in a state of adaptation to its natural murine host. I have investigated the resulting isolates from this complex evolution model and compared them to the same bacteria adapted to laboratory media which is more traditionally used in evolution experiments. By this we were uniquely placed to determine initial changes as a result of bacterial adaptation to a new and complex environment.

1.2 Experimental Evolution

Historically, studies of evolution were restricted to descriptive studies, with a major drawback being that such studies lacked the rigour of the scientific method found in other disciplines¹⁵. Experimental evolution changed the playing field, allowing hypotheses of evolutionary mechanisms to be tried and tested, and evolution processes to be explored and manipulated in real time¹⁵. Bacteria are constantly changing and adapting to their environment, with important implications for us as human beings. This takes form as the development of antibiotic resistance, the emergence of new diseases, and changes in disease symptoms and severity. Bacteria are an attractive model for evolution experiments, due to their large population sizes and relatively fast generation times¹⁶. The ability to store bacteria in an unchanging state at low temperatures uniquely allows direct comparisons between ancestor and descendants, allowing more rigorous and objective measures of relative fitness and change over the evolution period. The experimental evolution field is growing, aided by advances in technology enabling improved understanding of molecular mechanisms and genetic backgrounds of adaptations, with the number of published articles on the rise (Fig 1.1). This section will offer a brief summary of some of the key evolution experiments to date and the lessons learned, and will describe the ‘hole’ in current research that the work described in this thesis aims to fill.

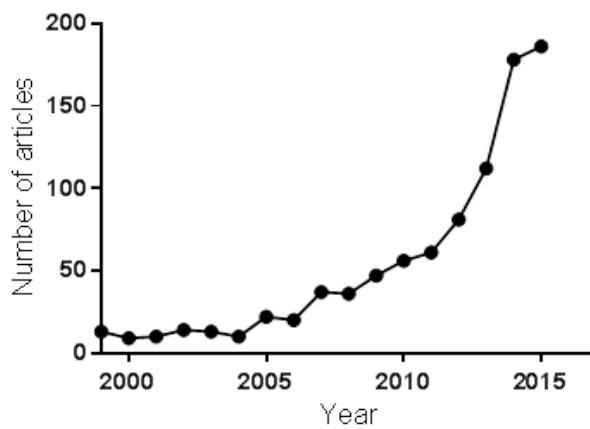


Figure 1.1. Number of published experimental evolution papers is increasing. This graph depicts the number of articles listed on the database PubMed with the words “experimental evolution” in either their title or abstract between the years 2000 and 2015, inclusive.

1.2.1 Experimental evolution to date: *in vitro*

Historically, evolution experiments have been performed in the laboratory in flasks, by transferring samples of microorganisms in a defined laboratory media at set time points or via continuous culture methods. Among the earliest of these such experiments include those performed by William Dallinger, who adapted flagellates to increasing temperatures; over several years he increased the temperature that the flagellates were exposed to, and found ‘adapted’ flagellates which could survive at the increased temperature of 70°C but could no longer thrive at the original temperature of 16°C¹⁷. The importance of these experiments was, unfortunately, underappreciated, and the experimental evolution field remained dormant until the later part of the 20th century. Rather than attempt to detail all *in vitro* evolution experiments published, I will highlight the broader concepts established by a variety of well-known experiments, providing a taster so as to appreciate the breadth of studies and some of the important principles of evolution that have been gained from this work.

1.2.1.1 Trade-offs: benefits are frequently accompanied by a price

In 1988, Lenski and colleagues began what is to date the longest evolution experiment ever performed, dubbed the “Long-term Experimental Evolution” (LTEE). This experiment, which is still ongoing, has now reached over 60,000 estimated generations. Briefly, twelve independent lineages were started: six from an *E. coli* B strain (*Ara*⁻) and six from a spontaneous mutant strain which is able to utilise L-arabinose (*Ara*⁺)¹⁷. The populations were each grown in a defined laboratory media containing glucose as the carbon source, and subcultured daily. Lenski and colleagues observed that the mean fitness increase was approximately 37% over the first 2,000 generations¹⁸, which was subsequently shown to be a function of a decreased lag phase and increased growth rate¹⁹. The authors noted that the strains grew better in the presence of glucose, but not maltose, a glucose disaccharide, suggesting that the increased growth rate is due to improvements in glucose transport²⁰. Glucose and maltose are utilised similarly after transport into the bacterial cell and maltose is broken down into two glucose molecules, therefore differences in growth on these two carbon sources indicates that the uptake mechanisms, which differ between the two sugars, had changed following adaptation to glucose²⁰. Lenski’s evolved lineages were also found to have a tendency to be more fit when grown on lactose, but less fit on melibiose, indicating that adaptation to glucose is accompanied by improvement in lactose utilisation but that this comes at a “trade-off” cost of decreased fitness on melibiose²¹. All twelve lineages lost the ability to metabolise D-ribose, with the loss detected as early as generation 500 and become fixed in all lineages by generation 2,000²². Restoring D-ribose catabolic function by restoring a functional *rbs* operon to two of the mutants resulted in decreased fitness, and strains which had artificially had the *rbs* operon impaired resulted in improved fitness, indicating that removing

D-ribose metabolism is beneficial in the environment that the bacteria are evolving in, perhaps indicating a high cost of functional *rbs* and its lack of utility in the media used²².

A further trade-off was observed after 20,000 generations: at this point, the average fitness of the populations had improved by approximately 70%, relative to the ancestor²³. The authors noted that the evolving *E. coli*, which had been consistently grown at 37°C, had begun to show evidence of specialising in growth at that temperature, and had a decreased growth rate at lower (20°C) and higher (42°C) temperatures²⁴. After going back through the fossil record, they discovered that this narrowing of optimal temperature could be seen as early as generation 2,000²⁴. They also noted that the evolved populations had become more sensitive to freeze-thaw cycles than the ancestor, with an average mortality rate of 53.7% per daily freeze-thaw cycle compared to the initial 34.4%²⁵; interestingly, the populations with the higher competitive fitness also had the higher mortality during freeze-thaw cycles ($r = 0.5882$)²⁵.

The LTEE focuses on bacteria grown in shaking broth with a mixed population; in contrast, important evolution experiments have been performed by Rainey and colleagues investigating populations grown in static conditions with varied environments within a single flask (referred to as “microcosms”). They grew *Pseudomonas fluorescens* in flasks with no shaking, and found that after 3 days growth there were three distinct colony phenotypes when the bacteria were grown on solid media: the ancestral “smooth” colonies (SM), and two new phenotypes: the wrinkly spreader (WS) and fuzzy spreader (FS) varieties²⁶. These WS cells form biofilms at the air-liquid interface in the flask²⁶, which stay at the surface of the broth due to their hydrophobicity, instead of buoyancy. Once the biofilm grows beyond a certain size, the biofilm collapses to the bottom of the flask, and so the WS phenotype is under selective pressure to be present at lower frequencies. Both the ancestral SM phenotype and WS phenotype populations can stably co-exist, as the fitness of each of these types is negative frequency dependent²⁷; that is, it is disadvantageous for each of the types to be present at high numbers.

A trade-off was observed for specialising in growth at the air-liquid interface; the WS cells had a reduced growth on many carbon sources; however, further selection and specialisation at the air-liquid interface reduced the observable trade-off, which the authors propose to be a result of compensatory adaptations²⁷. It is important to note that although a reduced fitness cost was observed, as shown by regaining the ability to grow on certain carbon sources, this cost was not removed, and the WS did not regain the ancestral phenotype of complete growth on the carbon sources. Similarly, while the WS phenotype is stable in the static microcosm, when grown on solid media the phenotype is unstable and quickly reverts to the ancestral SM phenotype, indicating that adaptation to the static microcosm comes with reduced fitness in the solid media environment²⁸.

Evidence of trade-offs have been seen in a wide variety of evolution experiments, and importantly there is the suggestion that adaptation to one environment can alter the virulence of a pathogen. Ellis and Cooper adapted *Burkholderia cenocepacia* to an onion-based liquid medium for 1,000 generations, which resulted in adaptation to the onion medium and a mean fitness improvement of 78% compared to the ancestor²⁹. However, by the 500th generation many of the *B. cenocepacia* lineages were unable to kill nematodes²⁹. This loss of virulence was due to adaptation to the onion, and not simply due to adaptation to an environment with no requirement for particular virulence genes, as *B. cenocepacia* evolved on the same media (M9 buffer) but with galactose, instead of macerated onion, did not have any observable reduction in virulence²⁹.

The examples listed in this section are related in that they support the idea that every adaptation is accompanied by a price, or a trade-off. Whether it is reduced metabolism on different substrates in response to improved utilisation of a single substrate²¹, decreased growth at temperatures different to the temperature that the microbes have adapted to²⁴, growth and cooperation in a spatially differentiated environment (as found in the static microcosms) leading to reduceds fitness in solid media²⁸, to loss of virulence of a pathogen when it adapts to an alternative host environment²⁹, the single message is clear: adaptation to one condition occurs with a reduced ability to thrive in another.

1.2.1.2 Microbes readily adapt to stressful conditions

In another study dubbed the Evolutionary Adaptation to Temperature study, Lenski and colleagues grew populations of the same *E. coli* strain at various temperatures. Briefly, an *E. coli* isolate which had adapted to 37°C for 2,000 generations, was used to start six replicate lineages in each of the experimental temperature groups: 32°C, 37°C, 42°C, and a 32°C /42°C ‘thermal generalist’ group exposed to 32°C and 42°C in an alternating cycle³⁰. Later, the experiment was expanded to include a group at 20°C³¹. The groups were grown and subcultured in their designated temperatures for 2,000 generations³⁰. The rate of adaptation to the initial “ancestral” environment, 37°C, was the slowest; and out of the lineages adapting to new environments, the highest temperature had the faster rate of adaptation, followed by the lowest temperature, and then the fluctuating environment³⁰. Each of the groups, after adaptation to their particular temperature for 2,000 generations, was then grown at 20°C for 2,000 generations to determine if the history of a population would influence the adaptation to the new environment³¹. The authors found that all groups adapted to the new 20°C environment, with the history of the populations having no observed effect on adaptation to the new environment³¹. Lenski and colleagues investigated extending the range of temperatures which their *E. coli* populations could survive, and discovered “thermotolerant” mutants (mutants which could survive a lethal temperature of 44°C) in populations from the 42°C and 32°C groups; the thermotolerant mutants did not have

improved growth at 42°C, indicating that survival at 44°C did not necessarily require or was linked to improved fitness at the high temperature environment 42°C^{32,33}.

Zhou and colleagues published a study wherein they investigated what conditions bacteria can adapt to, and what limits they can be pushed to, by studying tolerance to high salt conditions (halodurancy) in the anaerobe *Desulfovibrio vulgaris*³⁴. The authors observed that it required only 100 generations for *D. vulgaris* to adapt to media with a salt-stress, with an improved salt tolerance following a 2-fold improvement in growth rate. Combining adaptive evolution with genomic studies, they determined that only a few beneficial mutations were required for this adaptation³⁴ which, combined with the small number of generations required, indicates that halodurancy is relatively easy to evolve. Interestingly, within the first 100 generations, pre-existing genetic variations were selected for with a rapid improvement in halodurancy, with a much more gradual improvement following this initial burst. This suggests that the variation which resulted in halodurancy was already present in the population, and so adaptation to some stressors might be extremely rapid due to the inherent variation present in a population³⁴. Zhou's study, and the work of others, serves to highlight the power of experimental evolution when combined with genome sequencing techniques to match phenotype and genotype, pin-pointing gene function and processes to outcome and effect.

Both these studies highlight the remarkable ability of microorganisms to adapt to stressful environments, a lesson that resonates with the original work by William Dallinger¹⁷. Indeed, the true limits of microbes and what they cannot eventually overcome is yet to be determined.

1.2.1.3 The importance of genetic variation

As suggested by the study by Zhou and colleagues on *D. vulgaris*, selection works on existing genetic variation, and therefore the variability, mutability, and generation time of a population would play a role in determining the speed and material with which adaptation can occur. In 2015, Weller and Wu investigated the role of bacterial generation time on the gain of genetic variation and found that decreased generation times, as found in bacteria which form spores, is associated with decreased spontaneous genetic mutations compared with bacteria with faster generation times, as found in bacteria which do not form spores³⁵. Evolution experiments have also identified the emergence of “mutators”: strains with higher mutation rates than the ancestor. After 10,000 generations of the LTEE, three lineages, two from the Ara⁻ group and one from the Ara⁺ group, developed higher spontaneous resistance rates to the antibiotic nalidixic acid compared to the ancestral mutation rate³⁶. After 20,000 generations, another ‘mutator’ strain had emerged, bringing the total mutator lineages to four out of twelve²³.

At generation 7,000 of the Lenski LTEE, the role of conjugation was investigated by adding plasmid-carrying Hfr *E. coli* K12 cells³⁷. These donor cells had a variety of phenotypic differences, such as resistance to antibiotics, the ability to grow in the absence of certain amino acids, or resistance to bacteriophages, each of which could be assessed. The evolved *E. coli* lineages continued to be evolved for a further 1000 generations, with Hfr *E. coli* from freezer stocks added to the media every fifth day (every 33 generations). Recombination was first detected at a median time of generation 600, with variation between the different lineages. Many lineages had entirely lost the ancestral phenotype by generation 1000, while one still maintained an 80% proportion of ancestral phenotype. By this, the authors propose that conjugation is able to provide greater leaps in adaptation than by mutation alone³⁷, and showed that evolution in a more complex environment has the potential to have greater changes in fitness than observed in the LTEE.

These studies listed offer evidence as to the importance of genetic variability, either in the form of mutation rate³⁶, horizontal gene transfer³⁷, or simply in the generation time of the organism and the speed on which the beneficial genetic changes which accrue can be seen as observable increased phenotypic fitness³⁵.

1.2.1.4 The Cit⁺ lineage and the speciation debate

After 31,500 generations of the Lenski LTEE, a single lineage of *E. coli* developed the ability to utilise citrate (Cit⁺)³⁸, which was present in the media and which the ancestral strain is unable to use in aerobic conditions. All of the *E. coli* capable of metabolising citrate possess a mutation in *citT*, enabling the citrate transporter to be expressed under aerobic conditions; however the phenotype observed is only weak metabolism of citrate³⁹. In combination with another mutation, which increases C₄-dicarboxylate transporter expression, citrate metabolism is enhanced resulting in strong utilisation of citrate³⁹. These two mutations alone, however, are insufficient to fully recreate the strong citrate utilisation phenotype, indicating a function for other mutations present³⁹. This gain of ability to utilise citrate in an aerobic environment was proposed to be contingent on the history of the population; that is, that the change to Cit⁺ would not appear independently of the history of that particular lineage, and this was tentatively suggested to be initial evidence of speciation^{38,40,41}.

This finding and associated interpretations have been challenged by a subsequent study: in 2016, Hofwegen and colleagues described the evolution of a Cit⁺ *E. coli* strain in as few as 12 generations, with improved citrate utilisation taking only 100 generations^{42,43}. The difference between the two experiments is that in the Hofwegen study, the flasks of cells were grown for one week before diluting and serially passaging, whereas in the Lenski experiment the populations were subcultured every 24 hours, therefore reducing the amount of time for growth and selection on the alternative carbon source,

citrate⁴³. In media equivalent to the media used in the LTEE, Cit⁺ mutants of the same *E. coli* B strain could be isolated after 63 days⁴²; a substantially shorter period of evolution than 31,500 generations (15 years). Significantly, the changes required to gain the Cit⁺ phenotype in the Hofwegen study were similar to the changes required in the LTEE⁴². As the associated changes and phenotype switch could be reliably achieved under a much shorter time frame and without any particular “history,” Hofwegen and colleagues assert that the evidence for a speciation event is scarce⁴².

1.2.1.5 The importance of cooperation

Microbial evolution experiments are commonly performed with populations, rather than individuals; these experiments have offered some insight as to the benefits of cooperation and the eventual negative cost to cheating. Within the Rainey *P. fluorescens* experiments, “cheaters” have been observed within WS communities. The researchers initiated microcosms with WS cells and followed them over 10 days; at day 5, “ancestral-like” SM cells were present, and when these cells were competed with the ancestral SM cells there was no difference in competitive fitness⁴⁴. These “cheater” cells successfully avoided the cost of cooperation in the WS community, that is, by over production of acetylated cellulose, and yet are still held in the biofilm, gaining the benefits of being in the biofilm without contributing. If the WS community grows with any cheaters, then the group is ultimately doomed due to the success of the biofilm. Without cheaters, the increasing weight of the biofilm causes it to eventually collapse at around day 8. However in the presence of cheaters, the biofilm collapses after only 4 days. At day 3, 24% of cells in the biofilm are cheats⁴⁴. As they do not contribute to the strength of the biofilm, but do contribute to the weight, the biofilm collapses prematurely, therefore encouraging cooperation and discouraging cheats.

1.2.1.6 Changes in the environment influence the rate of coevolution

P. fluorescens, which had grown in the static conditions for 7 days, and then in static conditions for another two successive 7-day periods, maintained a high diversity of populations; in contrast, *P. fluorescens* which was moved to a shaking environment for two successive 7-day periods following growth in the static condition underwent an overall loss in diversity of bacterial types²⁶, indicating that mixing of the community was associated with a reduction in bacterial diversity. Similarly, mixing of the populations influences the rate of coevolution of *P. fluorescens* and a parasitic bacteriophage, with coevolution in a shaking environment leading to both more resistant *P. fluorescens* as well as more infective bacteriophage⁴⁵. Briefly, microcosms with *P. fluorescens* and the bacteriophage were under two conditions: static (no shaking), or shaking (with 1 minute of shaking every half hour). Culture was transferred to fresh medium every 48 hours for 16 transfers (120 bacterial generations). The rate of

change in bacteriophage infectivity was twice as fast in the mixed (shaking) populations⁴⁵. The bacteria from the mixed community were also more broadly resistant than the unmixed bacteria, being more resistant to both mixed and unmixed phage⁴⁵. The phage from the mixed community had evolved to be more infective than unmixed phage, and the bacteria from the unmixed community were less resistant to mixed than unmixed phage⁴⁵.

1.2.1.7 Adaptation to new niches and the importance of diverse communities

While the initial evolution experiments may start with a known, clonal population of the ancestor, following adaptation to their new environment a wide variety of phenotypes are often observed. In 2014, Callahan, Fukami, and Fisher published a paper investigating the evolution of *P. fluorescens* to different niches in laboratory media, by allowing the bacteria to adapt to the laboratory environment and then investigating the formation of any niches within the communities^{46,47}. They found that niches appeared within 100 generations and became fixed after approximately 400 generations^{46,47}.

Poltak and Cooper encouraged the adaptation of *B. cenocepacia* biofilms for 1,500 generations in a model which involved daily cycles of the stages of biofilm development: colonisation, assemblage of the biofilm, and dispersal⁴⁸. Briefly, media was inoculated with a 7 mm plastic bead which was colonised with a *B. cenocepacia* biofilm⁴⁸. At each 24 hour time point, a newly colonised bead was transferred to a new tube⁴⁸. Successfully passaged *B. cenocepacia* must have successfully colonised a bead, dispersed from the bead, and then successfully colonised a new bead. A number of phenotypes were detected: studded, ruffled spreader, and wrinkly, with names referring to their appearance on solid media⁴⁸. All phenotypes had appeared by generation 300, and all had relatively greater fitness, biofilm production, and planktonic growth rate compared to the ancestral type⁴⁸. The community of different phenotypes provided benefits to all members, with mixed communities performing better than expected, which was thought to be due to cross-feeding of different types with secondary metabolites from other phenotypes and due to the phenotypes occupying separate spatial niches⁴⁸. Confocal scanning microscopy was employed to determine the architecture of the biofilms produced, and showed that the wrinkly phenotype colonises the bead first, reaching a maximum between 4-6 hours exposure to the bead; next, the ruffled spreader phenotype colonises, reaching a maximum at 18 hours, and then finally the studded phenotype attaches to the ruffled spreader and wrinkly phenotypes, steadily growing until reaching 85% of the biofilm mass by 24 hours exposure to the bead⁴⁸, indicating that the biofilm thrives with a diverse community of distinct phenotypes.

Korona and colleagues evolved the soil bacterium *Ralstonia* species strain TFD41 for 1,000 generations in two different minimal media environments: liquid media, and solid media⁴⁹. The media

contained minimal salts with 2,4-D as the sole carbon source⁴⁹. All evolved strains were more fit than the ancestral strain in the minimal media, with the strains evolved to the solid media having higher fitness than those evolved to the liquid media⁵⁰. They found that lineages adapted to the liquid media were more similar, while those evolved to the solid media diverged more from each other⁴⁹. When the evolved *Ralstonia* was compared in a rich media, containing tryptone, peptone, yeast extract, and acetic acid, each of the lineages had similar fitness as the ancestor in the liquid or solid condition they had evolved to, whereas when the lineages were put in an environment opposite to their evolved condition, in the rich media the evolved lineages were less fit than the ancestor⁵⁰. Further analysis of the strains evolved to the different conditions for 1,000 generations by Riley and colleagues revealed a plethora of phenotypic differences, including the loss of capsule formation and a change in cell length: 8/12 liquid-evolved strains were longer, and 2/12 liquid-evolved strains shorter than the ancestor; 3/6 solid-evolved strains were longer, and 1/6 solid-evolved strains were shorter than the ancestor⁵¹. All of the evolved strains had an increased sensitivity to bile salts, and the majority exhibited improved adhesion to a sand matrix, compared with the ancestral strain⁵¹.

1.2.2 Experimental evolution to date: *in vivo*

The observations of adaptation to an *in vitro* environment do not necessarily reflect what occurs following adaptation to an *in vivo* environment; MacLean and Vogwill have discussed the disparity between trends of antibiotic resistance in the absence of antibiotic pressure in clinical observations and *in vitro* evolution experiments, highlighting a limitation of the design of evolution experiments not truly mimicking real-life scenarios⁵². Since the turn of the millennia, there has been growing interest in taking experimental evolution out of the flask and into an animal host. While none of these studies have the breadth or longevity of the Lenski LTEE, these first steps offer important insight into the evolutionary processes which occur in a more complex environment.

In 2001, Giraud and colleagues published a study which involved inoculating germ-free laboratory mice with the non-pathogenic *E. coli* MG1655 strain or *E. coli* with a mutation in the *mutS* gene conferring higher mutation rates⁵³. Mice were followed for 400 days and the fitness of the *E. coli* strains determined at various time points. The strain with a higher mutation rate was shown to have an advantage over MG1655 as shown by higher bacterial numbers in the animals over the first 42 days, however this advantage was transient and MG1655 reached the same fitness level as the mutator strain after approximately 400 days⁵³.

In 2004, Nilsson and colleagues published an *in vivo* evolution study investigating *Salmonella typhimurium* adaptation to laboratory mice⁵⁴. They investigated two strains of *S. typhimurium*: a wildtype and a strain with a mutation in the *mutS* gene which confers a high mutation rate. *S. typhimurium* was allowed to adapt to BALB/c mice for 8-10 passages, with mice artificially infected via intraperitoneal injection at each passage, and with each infection lasting 3-4 days. Following adaptation, the strains were each competed against a separate kanamycin resistant strain to enable distinction between the recovered bacteria. The results of their study indicated that *S. typhimurium* is able to rapidly increase its fitness to the mouse environment, however in the high mutation rate strain this specialisation comes with a trade-off cost of reduced metabolic capability⁵⁴. The wildtype strains, in contrast, adapted to the mouse environment without any subsequent loss of fitness in laboratory tests⁵⁴.

In 2010, Lee and colleagues published a study detailing the adaptation of non-pathogenic *E. coli* K-12 to mouse colonies (each containing three breeding pairs) for 1,031 days⁵⁵. Germ-free mouse colonies were inoculated with either *E. coli* with a type 1 pilus constitutively expressed or *E. coli* lacking a type 1 pilus. Within 79 days, the *E. coli* originally expressing the type 1 pilus had no evidence of type 1 pilus production, as determined by immunoassays. The authors observed that the size of cultured *E. coli* colonies from these mouse colonies diversified over the course of the

experiment, and that *E. coli* was present in the animals at an increased density than from animals at the beginning of the experiment⁵⁵.

In 2013, Lüttich and colleagues serially passaged the opportunistic pathogen *Candida albicans* SC5314 through laboratory mice using a systemic infection model, whereby BALB/c mice were infected intravenously with *C. albicans*; for subsequent passages mice were artificially inoculated with *C. albicans* isolated from the kidneys of infected animals⁵⁶. Over the period of eight passages, very few phenotypic differences were observed and the authors concluded that *C. albicans* is already well adapted to the mouse kidney⁵⁶.

In 2014, McCarthy and colleagues colonised germ-free piglets with human- and pig-associated strains of *Staphylococcus aureus*, following genetic and shedding changes over a period of 16 days⁵⁷. Horizontal transfer of genetic information from the pig-associated to the human-associated *S. aureus* was detected as early as 4 hours post-colonisation, occurring at significantly higher levels *in vivo* than predicted from a complementary *in vitro* study⁵⁷.

In 2015, Mikonranta and colleagues explored the evolution of the opportunistic pathogen *Serratia marcescens* db11 to an outside environment and to a host environment (fruit flies)⁵⁸. *S. marcescens* was adapted to either an outside environment, or to fruit flies via contaminated food. Subsequent passages of fruit flies followed isolation of *S. marcescens* from the fruit flies, growth in rich media, and then infection again via contaminated food. Ten passages were performed and the isolated bacteria then assessed for any changes. The authors found that *S. marcescens* that had adapted to infection of the fruit flies had reduced virulence, as assessed by death rate of infected fruit flies, compared with the ancestor and the bacteria adapted to the outside environment. This study, while showing a natural infection of fruit flies with *S. marcescens*, included enrichment of the bacteria in laboratory media between every passage, and therefore does not fully encapsulate the complete infectious cycle in its natural conditions.

These *in vivo* studies, while immensely valuable and beginning to highlight the differences between *in vitro* and *in vivo* evolution experiments, do not follow the relationship between a natural pathogen to its natural host, and do not allow the pathogen to infect subsequent hosts via natural transmission processes. The route of infection impacts subsequent adaptation⁵⁹, and therefore following pathogens through natural infection processes is required for assessments to reflect real-world processes. In this way, the studies described are still an artificial glimpse to a process which is currently being performed on a massive scale in our day to day lives.

1.2.3 Pathogen Evolution

While scientists have spent decades studying carefully controlled evolution experiments, pathogenic bacteria have been performing their own ‘experiment’ out in the field. The causative agent of Plague, *Yersinia pestis*, evolved from the enteric pathogen *Y. pseudotuberculosis* as recently as 1,500 years ago¹¹. *Vibrio cholerae*, the bacterium responsible for cholera, can become hyperinfectious after growth and passage through a human host⁶⁰. In 2011, there was an outbreak of *E. coli* food poisoning in Europe with over 3,000 cases of disease, which was subsequently found to be due to an Enteropathogenic *E. coli* (EAEC) which had acquired the Shiga toxin gene via horizontal gene transfer and become highly virulent^{12–14}. Pathogens are under no obligation to wait for the rate of scientific research to reach their rate of adaptation.

1.2.4 The gap in the literature

As stated previously, there is very little published literature about experimental evolution of pathogens to their hosts, and very little which follows evolution of the pathogen for an extended period of time. A lot of the lessons that we have learned are taken from long-term *in vitro* evolution experiments, which while providing a wealth of information, may not apply to a more complex environment as found in a host organism.

There are a great number of questions still to be answered; whether adaptation to a host results in increased or decreased virulence has no conclusive answer. Some evidence shows that it is in the best interests of bacteria to allow their host to survive as long as possible, with commensals hailed as the most ‘adapted’ of bacteria. Indeed, Price and colleagues found that chronic carriage of the noncommensal bacterium *Burkholderia pseudomallei* in one patient lead to reduced virulence of the *B. pseudomallei* isolated – indicating that adaptation to the host resulted in a more commensal-like state⁶¹. In the case of *Neisseria meningitidis*, colonisation of the nasopharynx of healthy individuals allows transmission and spread of bacteria, while the ‘switch’ to a pathogenic state and invasion of the bloodstream and central nervous system, makes it difficult for these populations to transmit to new hosts, and therefore the more invasive disease state is considered not to be in the best interests of *N. meningitidis*^{62–64}. However, in the case of other bacteria, such as *M. tuberculosis*, there is evidence that virulence is increasing over time^{65,66}. There is a need to understand what microbes will do in a given environment before they happen – to understand what factors can influence potential outcomes of adaptation.

We are already influencing the bacteria in our environment and perhaps driving adaptation to places we cannot necessarily predict. One example is the widespread use of antibiotics, not only in human and veterinary medicine, but also in farming practices. While our understanding of selective pressure allows

that antibiotic exposure may lead to the selection of resistant mutations, it does not make predictions regarding pathogenicity, infectivity, or other attributes. The relationship between antibiotic resistance and virulence is complicated, and depends much on the organism and antibiotic resistance mechanism. Clinical *S. aureus* isolates that are rifampicin-resistant are less virulent than rifampicin-sensitive strains⁶⁷ while strains of *Salmonella typhimurium* resistant to streptomycin, rifampicin, or nalidixic acid are unable to cause disease in mice but can quickly gain extra mutations which restore virulence while remaining resistant to the antibiotics⁶⁸. In contrast, the efflux pumps in *Vibrio cholerae* play an important role in antibiotic resistance as well as increasing virulence in mice⁶⁹. *S. aureus* clinical isolates which are resistant to the antibiotic vancomycin are more virulent in the insect *Galleria mellonella* model⁷⁰. *Salmonella enterica* lacking efflux pumps are also both more sensitive to antibiotics and have reduced virulence⁷¹. Likewise, in *Klebsiella pneumonia*, efflux pumps make the bacteria both resistant to antibiotics and better able to cause pneumonia in mice⁷².

The research presented in this thesis explores both of these questions: the adaptation of a pathogen to its host, and the influence antibiotic exposure may have on the evolution of the pathogen. In choosing a host-pathogen model, a number of considerations are important. These include the ability of the pathogen and host to be monitored and manipulated, the ease and cost of housing and caring for the host, the hazards associated with using the pathogen and any potential for creation of more dangerous human pathogens. I have chosen the host-pathogen model of *Citrobacter rodentium* and laboratory mice for these evolution experiments, which is uniquely placed to be both suitable for studying host-pathogen interactions as well as being a relatively recent pathogen with the potential for further adaption (Table 1.1).

Table 1.1. Benefits of using mice and *C. rodentium* as a host-pathogen model

Easily manipulated host and pathogen	Laboratory mice have a long history of being used in research, and practices regarding infection of mice with <i>C. rodentium</i> have been established.
Complete infectious cycle	<i>C. rodentium</i> rapidly spreads between mice in the same cage, allowing natural transmission of the disease to occur. This also allows the complete infectious cycle to be followed with no artificial emphasis on any aspect.
No danger to humans	<i>C. rodentium</i> is a mouse pathogen, and there are no reports of this bacterium infecting humans.

1.3 *Citrobacter rodentium*

1.3.1 Diarrhoeal Illness

The global burden of food borne disease is significant, and comparable to other major infectious diseases such as tuberculosis, malaria, and HIV/AIDS⁷³. In developing countries, diarrhoeal illness is the cause of 21% of all deaths in children under 5, which amounts to 2.5 million deaths per year⁷⁴. Enteropathogenic *E. coli* (EPEC) is a significant cause of diarrhoea in children in developing countries⁷⁵, and was estimated to cause around 24 million cases of diarrhoeal illness worldwide in 2010, with these infections resulting in 37,000 deaths⁷³. Enterohaemorrhagic *E. coli* (EHEC) is more of a problem in developed countries, with around 73,000 cases reported annually in the USA⁷⁶. Unfortunately, EPEC and EHEC are difficult to model in the laboratory as these pathogens do not cause a comparable disease in laboratory mice⁷⁷. This changed when, in the middle of the last century, several laboratory animal houses over the world experienced outbreaks of a diarrhoeal disease never seen before.

1.3.2 The History of *Citrobacter rodentium*

In 1964, two outbreaks of severe diarrhoeal disease with high mortality occurred in laboratory mouse colonies in two separate countries: in the Argonne National Laboratory (ANL) in the United States of America⁷⁸, and in the National Institute of Health in Japan⁷⁹. The disease was infectious, with sick animals co-housed with uninfected animals leading to the spread of disease⁷⁸. Gram-negative non-motile rods were isolated from infected animals, and the pathogen was dubbed atypical *Citrobacter freundii* ANL in the USA and atypical *Escherichia coli* Ex30, or murine pathogenic *E. coli* (MPEC) in Japan. The disease itself was called catarrhal enterocolitis and infectious megaenteron, respectively, both reflecting the disease symptoms of thickening, or hyperplasia, of the colons of the infected animals. The *C. freundii* ANL was considered atypical as the biochemical assays employed showed that the bacteria did not utilise citrate as the sole carbon source, and did not lead to production of hydrogen sulphide on Kligler iron agar, which is indicative of *C. freundii*⁷⁹.

Another outbreak in Fort Deitrick in the USA occurred in 1968, with infected mice suffering rectal prolapse as well as the characteristic colonic hyperplasia, but with no diarrhoea⁸⁰. Atypical *C. freundii*, again, was isolated from the infected animals and this isolate was dubbed *C. freundii* Ediger⁸⁰. A fourth disease outbreak in laboratory mice, this time accompanied with diarrhoea, occurred in the University of Tokyo, Japan, in 1971, again with the characteristic colonic hyperplasia⁸¹. In 1972, yet another outbreak of disease with loose stools and hyperplasia of the colon occurred in the USA, with another atypical *C. freundii* isolated, this time dubbed *C. freundii* biotype 4280⁸². Each of these

geographically separate outbreaks in mouse colonies had very similar symptoms: while there was slight variation in whether there was presence of rectal prolapse and the degree of diarrhoea, hyperplasia of the colon was common to all⁸³.

Experiments by Barthold and colleagues showed that F3444 inbred rats and Syrian hamsters were unable to be infected with the atypical *C. freundii* isolate⁸⁴, and therefore the pathogen in question was traditionally thought to only infect and cause disease in mice. However, when the latest outbreak occurred in Spain, 1999, the known host range of the bacteria was diversified to include Mongolian gerbils⁸⁵. Bacteria isolated from the infected gerbils had slightly differed biochemical characteristics, with no urease production and an inability to utilise L-arabinose⁸⁵.

Prior to 1993, only three species of *Citrobacter* were officially recognised: *C. freundii*, *C. koseri*, and *C. amalonaticus*⁸⁶. The *C. freundii* isolates associated with the colonic hyperplasia disease in the USA did not match any of the biochemical characteristics of any known *Citrobacter* species, and so were named atypical *C. freundii*. Later, DNA-DNA hybridisation techniques were used to reclassify the *Citrobacter* taxon, with eleven *Citrobacter* ‘genomospecies’ formed which were defined by DNA relatedness⁸⁶. Genomospecies 9 included all of the atypical *C. freundii* isolates from the USA. In 1995, genomospecies 9 and *C. freundii* biotype 4280 were compared and found to be genetically and biochemically related, and combined in the form of a new species: *Citrobacter rodentium*⁸³. Finally, in 2000 Luperchio and colleagues found that the strains isolated from the outbreaks in Japan were indistinguishable from those from the outbreaks in the USA⁸⁷, indicating that each of the outbreaks, occurring in different parts of the world in isolated animal units, were caused by the same pathogen. It is important to note that for each of these outbreaks, the ultimate source of infection was never identified and *C. rodentium*’s presence outside of animal laboratories has not been fully explored, and therefore the exact origins and the full host range of *C. rodentium* remains unknown.

Table 1.2. *C. rodentium* outbreaks
Adapted from Mundy *et al.*⁸⁸

Bacteria isolated	Location	Year	Species	Reference
<i>C. freundii</i> ANL	Argonne National Laboratory, USA	1964	Mouse	Brennan <i>et al.</i> (1965)
Murine pathogenic <i>E. coli</i>	National Institute of Health, Japan	1964	Mouse	Muto <i>et al.</i> (1969)
<i>C. freundii</i> Ediger	Fort Detrick, USA	1968	Mouse	Ediger <i>et al.</i> (1974)
Murine pathogenic <i>E. coli</i>	University of Tokyo, Japan	1971	Mouse	Itoh <i>et al.</i> (1978)
<i>C. freundii</i> 4280	USA	1972	Mouse	Barthold <i>et al.</i> (1976)
<i>C. rodentium</i>	Section of Laboratory Animal Resources, Spain	1999	Mongolian gerbils	Puente-Redondo <i>et al.</i> (1999)

1.3.3 Transmissible Murine Colonic Hyperplasia

The disease caused by *C. rodentium* infection is known as transmissible murine colonic hyperplasia⁸⁴, and its severity is largely determined by the genetic background of the infected animals. Infection in normal adult mice is often subclinical^{89,90}, whereas infection of suckling mice and some transgenic mouse lines results in severe clinical disease and high mortality^{87,90} (Table 1.3). In mice susceptible to clinical disease, this manifests itself as classical signs of sickness, including listlessness, ruffled coats, and a hunched posture⁹⁰. The colons of infected animals are often thickened and shortened⁹¹. Animals which do die from *C. rodentium* infection do so from dehydration and fatal fluid loss following severe diarrhoea⁹².

Following artificial infection of healthy animals with *C. rodentium*, the bacteria colonises the caecal patch, a lymphoid structure in the caecum^{93,94}. The bacteria attach to the intestinal cells and then undergo changes in gene expression that enables colonisation of the lower colon and rectum⁹⁵. The peak of infection occurs at approximately a week post exposure; at this point, *C. rodentium* will be present at between 10^9 and 10^{10} colony forming units (CFU) per gram of colon, and will have replaced 90% of the aerobic bacteria in the colon^{78,84,93,94,96}. As epithelial cells are shed from the intestines, the infection begins to clear, with complete clearance of the pathogen occurring approximately 2-3 weeks post infection^{93,94}. Mice which clear the infection gain a lasting immunity to any further *C. rodentium* infection⁹⁷.

When it was observed that *C. rodentium* colonises the mucosa of the large intestines via characteristic attaching and effacing (A/E) lesions⁹⁸⁻¹⁰⁰ (Fig. 1.2), similar to EPEC and EHEC in humans, *C. rodentium* became of interest to the EHEC and EPEC research communities.

Table 1.3. Mouse knockouts which can or cannot clear *C. rodentium* infection
Adapted from Mundy *et al.*⁸⁸

Mouse strains which cannot clear infection	Mouse strains which clear infection
RAG1 knockouts (no T and B cells)	IgA-deficient
μMT knockouts (no B cells)	IgM-deficient
CD4 knockouts	IgG3-deficient
TcRβ knockouts (no αβ T cells)	TcRδ knockouts (no γδ T cells)
	CD8 knockouts

1.3.4 *C. rodentium* virulence factors

C. rodentium, much like EHEC and EPEC, cause attaching-effacing (A/E) lesions in the intestines, which are characterised by localised destruction (or effacement) of the brush border microvilli, close attachment of the bacterium to the apical host cell membrane, and formation of a pedestal-like structure on the host cell¹⁰¹. When viewed via transmission electron microscopy, a gap of only 10 nm can be seen between the enterocyte and the intimately attached bacterium⁸⁹. The pedestal-like structure on which the bacterium is perched can protrude as far as 10 µm beyond the normal surface of the enterocyte⁸⁹. Formation of these A/E lesions is a multi-step process (Fig. 1.2): first, the bacterium undergoes initial attachment to the enterocyte; second, the translocated intimin receptor (Tir) is moved to the enterocyte via a type III secretion system (T3SS); third, Tir on the enterocyte surface intimately attaches to intimin on the bacterium, and the cytoskeleton of the enterocyte dissolves; and finally, filamentous actin is polymerised to form the characteristic pedestal protruding from the enterocyte. Formation of A/E lesions is critical for successful infection of the animal, as *C. rodentium* strains with dysfunctional Tir¹⁰⁰, Intimin⁹⁸, or T3SS¹⁰² are unable to colonise or cause disease.

Genes required for A/E lesion formation are found on a pathogenicity island called the locus of enterocyte effacement, or LEE¹⁰³. The LEE encodes for a T3SS, a complex that spans both bacterial membranes and is used by many Gram negative bacteria to put virulence factors directly into host cells¹⁰⁴. The LEE5 operon encodes the outer membrane adhesin Intimin (*eae*)¹⁰⁵, the translocated intimin receptor Tir (*tir*)¹⁰⁶, and CesT (*cesT*), which is the chaperone of Tir^{107,108}. The LEE4 operon encodes for the structural needle protein EscF (*escF*)¹⁰⁹, the translocator proteins EspA (*espA*) and EspD (*espD*), and the effector proteins EspB (*espB*)^{110–112} and EspF (*espF*)¹¹³.

The Ler (LEE-encoded regulator) directly activates operons LEE2, LEE3, LEE4, and the *tir* gene^{114–117}. GrlA (global regulator of LEE activator) and GrlR (global regulator of LEE repressor) appear to function indirectly, possibly through modulating Ler^{118,119}. A fourth regulatory gene is found in the LEE3 operon, which is named *mpc*¹²⁰. Mpc is involved in the regulation of LEE expression, and removing the gene adversely impacts *C. rodentium*'s ability to form pedestals and therefore reduces virulence¹¹⁸. The *cfc* gene cluster encodes for a type IV pilus, unique to *C. rodentium*¹²¹. This pilus is essential for colonisation of the mouse intestine, as *C. rodentium* lacking *cfc* cannot be cultured from stools from day 3 post-infection¹²¹. The cluster of genes includes twelve tandem genes: *cfcA*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*, *J*, *P*, and *V*¹²¹.

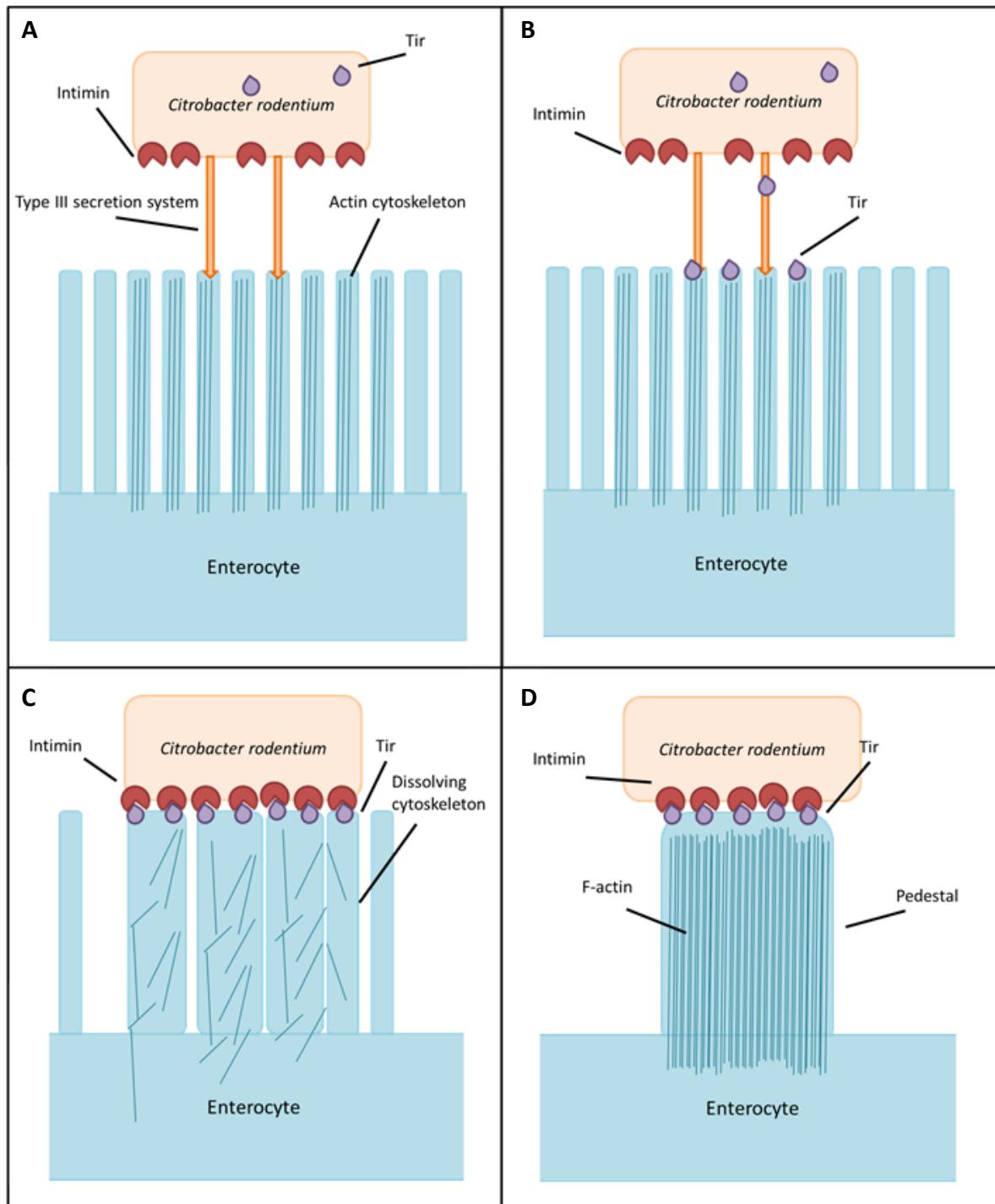


Figure 1.2. *C. rodentium* attachment to enterocytes.

(A) Initial attachment of the bacterium to an intestinal cell. (B) Tir (translocated intimin receptor) is translocated to the host cell via the type III secretion system. (C) The bacterium is intimately attached to the host cell, with intimin binding to tir. The host cytoskeleton dissolves, leading to effacement of microvilli. (D) Filamentous actin polymerisation and characteristic pedestal formation. Adapted from Wales *et al*⁸⁹.

C. rodentium lacking *escF* is incapable of establishing an infection of C3H/Hej mice, and can only be recovered from stools for the first 24 hours post infection¹²². Intimin and EspB are essential for pathogenesis of EPEC in human volunteer studies^{123,124}, and for virulence of *C. rodentium* in mice^{98,125}. Tir is also essential for successful *C. rodentium* infection¹⁰⁰. Mutations in *orf4*, *espD*, and *escD* of *C. rodentium* result in reduced virulence in mice¹²¹. EspI is dependent on the T3SS for infection in C3H/Hej mice¹²². *C. rodentium* lacking *espI* are recovered from infected mice 3-4 logs lower than the wildtype, and take longer to clear from C3H/Hej mice (day 33 post infection)¹²².

EspF, EspG, EspH, and Map are not essential for A/E lesion formation *in vitro*^{113,126,127}. EspF is involved in the disruption of tight junctions and in the induction of apoptosis^{128,129}, and *C. rodentium* lacking *espF* have a later peak of infection (day 12 post-infection compared with day 7 post-infection), slightly reduced hyperplasia, and take longer to clear from C3H/Hej mice¹²². EspG has an unknown function, and *C. rodentium* lacking *espG* show no colonisation defect in C3H/Hej mice^{122,126}. EspH is thought to downregulate filopodium formation and enhance pedestal formation¹²⁷. *C. rodentium* lacking *espH*, when infected into C3H/Hej mice, were recovered from stools the same as wildtype *C. rodentium*, indicating no colonisation defect¹²². Map is thought to target host mitochondria and interfere with membrane potential¹³⁰, and also to be involved in cytoskeletal rearrangements resulting in formation of Cdc42-dependent filopodia¹³¹. *C. rodentium* lacking *map* were recovered from infected mice 3-4 logs lower than the wildtype, as well as having a later peak of infection (day 12 post-infection compared with day 7 post-infection) and a longer clearance time (day 40 post-infection compared with day 14 post-infection) from the C3H/Hej mice¹²².

Recently, evidence of horizontal gene transfer has been shown for *C. rodentium* with the RegA regulon (a member of the AraC/XylS superfamily), which has also been horizontally transferred to *Escherichia* spp as well as *C. rodentium*¹³². In *C. rodentium*, RegA regulates the LEE pathogenicity island via *grlA*¹³³. Bicarbonate ions, which are found in the mouse gut, are required for RegA to influence gene expression^{95,133}, and therefore play an important role in allowing *C. rodentium* to determine its location in a host and respond appropriately.

Table 1.4. *C. rodentium* genes important for virulence and their putative functions

Gene	Function
<i>grlR</i>	Global regulator of LEE repressor
<i>grlA</i>	Global regulator of LEE activator
<i>ler</i>	Ler transcriptional regulator
<i>escR</i>	Structural protein
<i>escS</i>	Structural protein
<i>escT</i>	Structural protein
<i>escU</i>	Structural protein
<i>escC</i>	Structural protein
<i>sepD</i>	Secretion switching protein
<i>escJ</i>	Structural protein
<i>espZ</i>	Effector protein
<i>mpc</i>	Regulator
<i>escV</i>	Structural protein
<i>escN</i>	Translocator
<i>sepQ</i>	Structural protein
<i>espH</i>	Effector protein
<i>map</i>	Cytoskeletal rearrangement; effector protein
<i>tir</i>	Translocated intimin receptor
<i>eae</i>	Outer membrane adhesion; intimin
<i>escD</i>	Translocator protein
<i>sepL</i>	Secretion switching protein
<i>espA</i>	Translocator protein
<i>espD</i>	Translocator protein
<i>espB</i>	Effector protein
<i>escF</i>	Structural needle protein
<i>espF</i>	Effector protein; disruption of tight junctions and induces apoptosis
<i>espG</i>	Unknown function; effector protein
<i>espH</i>	Downregulation of filopodium and enhance pedestal formation
<i>espI</i>	Secreted protein
<i>cfc</i>	Colonization factor <i>Citrobacter</i> ; CFC type IV pilus

1.3.5 *C. rodentium* is still in the process of adapting to laboratory mice

As stated before, *C. rodentium* is a recently discovered pathogen, and therefore may have adapted to laboratory mice or gained the ability to cause disease in mice relatively recently. Analysis of *C. rodentium* isolates have revealed that the organism is still undergoing rapid adaptation, with a large degree of genomic flux characterised by rearrangements of chromosomal sequences and expansion of insertion sequence elements¹³⁴, indicating genetic flexibility and the potential for further change. Further flexibility is enabled by transcriptional slippage, which is a mechanism used by bacteria to produce two different proteins from the same gene. In *C. rodentium*, two differently sized TssM1 proteins (an inner membrane protein involved in the type VI secretion system) can be produced: one 88-kDa version which is truncated due to a stop codon¹³⁵, and one with a slippage event enabling the full-length protein 130-kDa to be produced¹³⁶. The production of the full-length version occurs at a 1/3 ratio to the truncated form¹³⁶. A similar mechanism is used by *Yersinia pseudotuberculosis* – albeit in the opposite manner: slippage in *C. rodentium* leads to a full-length form, while slippage in *Y. pseudotuberculosis tssM3* leads to a truncated version¹³⁶.

Petty and colleagues infected mice with *C. rodentium* and let the animals infect new animals for a total of ten passages, via natural transmission through coprophagia and grooming. *C. rodentium* which was isolated from infected animals at the tenth passage had gained resistance to the virulent phage φCR1, which is known to target lipopolysaccharide (LPS) as a receptor, indicating a change to the LPS of the *C. rodentium* isolate^{134,137}. This finding strongly suggests that *C. rodentium* is still actively changing its genome and has the potential to rapidly evolve further.

Previous experiments by Wiles and colleagues led to the discovery of a ‘hyperinfectious’ *C. rodentium* state following natural transmission of the pathogen from one mouse to another; *C. rodentium* in this hyperinfectious state is 1000-fold more infectious than the bacterium grown in rich laboratory media¹³⁸. This phenotype is transient and the bacterium returns to normal infectivity after a single overnight passage in laboratory media¹³⁸. As well as increasing the infectivity of *C. rodentium*, natural transmission also alters the initial site of infection, with the hyperinfectious *C. rodentium* bypassing colonisation of the caecal patch and directly colonising the colon¹³⁸.

Efforts have been undertaken to understand the mechanism of the hyperinfectious phenotype. Interestingly, the hyperinfectious *C. rodentium* is more sensitive to acid stress, instead of gaining tolerance to acid, as would be traditionally anticipated¹³⁹. Recently, RNASeq has been used to investigate the genes expressed by hyperinfectious *C. rodentium* during infection¹⁴⁰. The genes *escR*, *escS*, *escU*, *escC*, *sepB*, *escJ*, *espZ*, *mpc*, *escV*, *espH*, *cesF*, *map*, *tir*, *cesT*, *escD*, *sepL*, *espA*, *espD*, *espB*, *cesD2*, *escF* and *espF* were all upregulated in the hyperinfectious state but not in a media made from mouse faeces¹⁴⁰. Other known virulence genes essential for colonisation, such as *lifA*, and a newly

described gut colonisation fimbria (*gcf*)¹⁴¹ were not induced in the hyperinfectious state, indicating that some genes known to be important for colonisation in artificial infections are perhaps not as important for natural infections¹⁴⁰.

Taken together, these published results further emphasise the differences between artificial infections and infections which have come about via natural means, such as natural grooming and coprophagia behaviour, and that artificial infections may not fully mimic the infection dynamics that occur in nature. The ability of *C. rodentium* to rapidly and easily naturally infect its host makes it an ideal candidate for investigating natural transmission of a pathogen. That it is still in the process of adapting to its murine host means that it is uniquely positioned to be used to investigate the process of adaptation and evolution of pathogens. These properties of *C. rodentium* make it aptly situated to be used as a model of bacterial adaptation to a host, to determine the range of effects on its adaptation, and to measure changes with respects to its ancestor.

1.4 Thesis Aims

The work presented in this doctoral thesis aims to investigate the genotypic and phenotypic changes that occur following a prolonged period of adaptation and transmission of the enteropathogen *C. rodentium* through its natural host, laboratory mice. I have chosen C57BL/6 mice as the model host, as in this strain of mice a measurable and symptomatic disease accompanies *C. rodentium* infection, but rarely leads to mortality of the animals. Within the *in vivo* evolution experiment I have compared two treatment groups: one group of mice which received a low dose of the quinolone antibiotic nalidixic acid (which the *C. rodentium* strain used is resistant to) in order to alter the normal microflora, and another group which received no antibiotic treatment. In this way, I aim to investigate the question of whether antibiotic treatment will influence the course of evolution of *C. rodentium*. I will evaluate the null hypothesis:

- 1) Low-dose nalidixic acid treatment will have no impact on the normal microflora or the *in vivo* adaptation of *C. rodentium*

Should the data reject this null hypothesis, I will then move to evaluate two alternate hypotheses:

- 2) Low-dose nalidixic acid treatment will alter normal microflora, reducing the amount of genetic information available for recombination purposes and thus limiting adaptation; or
- 3) Low-dose nalidixic acid treatment will alter normal microflora, resulting in vacant space for *C. rodentium* and thus remove the barriers for colonisation, enabling efficient infection and faster adaptation

I have also simultaneously followed the *in vitro* evolution of *C. rodentium*, passaging the bacterium through laboratory media in a similar fashion to the well-known LTEE, using the traditionally established method of serially passaging bacteria daily to fresh batches of defined laboratory media. I hypothesise that the artificial *in vitro* model and *in vivo* model, as a result of the two distinctly separate environments, will lead to observably different phenotypes.

During these studies I have performed genotypic and phenotypic assessments of the ‘evolved’ strains to identify any evidence of adaptation to the mouse environment and to establish whether a 5-month period is sufficient for measurable adaptation to occur. This work is a novel undertaking, and seeks to extend current flask-based evolution experiments to a more complex and dynamic *in vivo* environment, allowing evaluation of whether lessons learned from *in vitro* studies translate to *in vivo* experimental models. In addition, the bank of stored samples generated over the course of the evolution experiments provides an extensive “fossil record”, allowing future in-depth analysis of the ‘evolved’ lineages and revivification of the bacteria and continuation of the evolution process. This bank will be a unique resource for future evolutionary biologists.

Chapter 2: Materials and Methods

2.1 Introduction

This chapter includes the catalogue number and supplier information for the major solutions and reagents which were used in the experiments. The methods for each of the experiments are divided by type; first, growth of bacteria and care of animals, then the experiments performed using mice, and lastly the experiments performed *in vitro*.

2.2 Materials

Table 2.1. Bacterial Strains

Species	Strain	Description	Reference
<i>Citrobacter rodentium</i>	ICC169	A spontaneous nalidixic acid resistant mutant of <i>C. rodentium</i>	Wiles and colleagues, 2005 ¹⁴²
	ICC180	Kanamycin resistant and bioluminescent derivative of <i>Citrobacter rodentium</i> ICC169	Wiles and colleagues, 2004 ⁹³
	W1 _{P20}	<i>In vivo</i> -adapted isolate which has passaged through 20 C57BL/6 mice and has <i>C. rodentium</i> ICC180 as its ancestor	This study
	W2 _{P20}	As W1	This study
	W3 _{P20}	As W1	This study
	W4 _{P20}	As W1	This study
	W5 _{P20}	As W1	This study
	N1 _{P20}	<i>In vivo</i> -adapted isolate which has passaged through 20 nalidixic acid treated C57BL/6 mice and has <i>C. rodentium</i> ICC180 as its ancestor	This study
	N2 _{P20}	As N1	This study
	N3 _{P20}	As N1	This study
	N4 _{P20}	As N1	This study
	N5 _{P20}	As N1	This study
	L1 _{G1200}	Laboratory-adapted isolate which has passaged through Modified Davis & Mingoli media with 1% glucose for five months and has <i>C. rodentium</i> ICC180 as its ancestor	This study
	L2 _{G1200}	As L1	This study
	L3 _{G1200}	As L1	This study
	L4 _{G1200}	As L1	This study
	L5 _{G1200}	As L1	This study
	L6 _{G1200}	As L1	This study
<i>Escherichia coli</i>	UPEC 536	Clinical uropathogenic <i>E. coli</i> strain originally isolated from a pyelonephritis infection	Knapp and colleagues 1986 ¹⁴³

Table 2.2. Animals

Item	Catalogue #	Supplier
<i>Galleria mellonella</i> Wax moth larvae	7100	http://www.biosuppliers.com
<i>Mus musculus</i> C57BL/6Elite	N/A	Vernon Jansen Unit, University of Auckland

Table 2.3. Eukaryotic cell lines

Item	Catalogue #	Supplier
L-929 murine fibroblast cells	N/A	Cryosite Ltd

Table 2.4. Media and Solution Recipes

Media/Solution	Composition	
	Constituent	Amount per litre
Super Optimal Broth (SOB)	Tryptone	20g
	Yeast extract	5g
	Sodium chloride	0.584g
	Potassium chloride	0.186g
	MgSO ₄	2.4g
	Dipotassium hydrogen phosphate anhydrous	10.5g
Modified Davis & Mingioli Media	Potassium dihydrogen phosphate	4.5g
	Sodium citrate dihydrate	5g
	Ammonium sulphate	1g
	Magnesium sulphate heptahydrate	24.65mg
L929 mouse fibroblast growth media	Newborn calf serum	10% v/v
	DMEM	90% v/v
TAE buffer	Sodium EDTA	3.75g
	Tris	48.4g
	Sodium acetate	6.8g
	Glacial Acetic Acid	16.9ml

Table 2.5. Chemicals and Media

Item	Catalogue #	Supplier
LB broth (Lennox)	240230	Becton, Dickinson & Co.
LB Agar (Lennox)	240110	Becton, Dickinson & Co.
SOC Medium	15544-034	Invitrogen
Dipotassium hydrogen phosphate anhydrous	60354	Sigma-Aldrich
Potassium dihydrogen phosphate	PO02620100	Sharlau
Sodium citrate dehydrate	W302600	Sigma-Aldrich
Ammonium sulphate	A6387	Sigma-Aldrich
Magnesium sulphate heptahydrate	M1880-500G	Sigma-Aldrich
Dulbecco's Modified Eagle Medium (DMEM)	11995-065	Gibco
Newborn Calf Serum	161010-159	Gibco
Roswell Park Memorial Institute (RPMI) Medium 1640	11874-093	Gibco
Tryptone	211705	Becton, Dickinson & Co.
Yeast extract	288620	Becton, Dickinson & Co.
Sodium chloride	S9888	Sigma-Aldrich
Potassium chloride	P9541	Sigma-Aldrich
Magnesium sulphate	M7506	Sigma-Aldrich
Phosphate buffered saline tablets	P4417	Sigma-Aldrich

Table 2.6. Stains

Item	Catalogue #	Supplier
FLUKA Calcofluor white	18909	Sigma-Aldrich
Crystal violet	57119	Searle Scientific Services
Iodine	285642Y	BDH Laboratory Supplies
Safranin	343122N	BDH Laboratory Supplies

Table 2.7. Anaesthetics

Item	Catalogue #	Supplier
Isoflurane	N/A	Lunan Better Pharmaceutical Vernon Jansen Unit, University of Auckland
Ketamine/xylazine	N/A	Vernon Jansen Unit, University of Auckland

Table 2.8. Enzymes

Item	Catalogue #	Supplier
Cellulase	C1184-5KU	Sigma-Aldrich

Table 2.9. DNA isolation

Item	Catalogue #	Supplier
1 Kb Plus DNA ladder	10787018	Thermo Fisher
SYBR Safe DNA Gel Stain	S33102	Thermo Fisher
Agarose	A9539	Sigma-Aldrich
PowerSoil® DNA Isolation Kit	12888	Qiagen
Genomic-tip 20/G	10223	Qiagen
Genomic DNA Buffer Set	19060	Qiagen
Proteinase K	19133	Qiagen

Table 2.10. Sample Storage

Item	Catalogue #	Supplier
Microvette 500 Z-Gel	20.1344	Sarstedt
Nunc® CryoVials®	V7884	Sigma-Aldrich

2.2.1 The pRhomo plasmid

The pRhomo plasmid, generously donated by Rodríguez-Beltrán and colleagues¹⁴⁴, is a 6,169 bp length plasmid used to determine bacterial homologous recombination rates. It contains an ampicillin resistance gene (*bla*), and a gentamicin resistance gene (*aacC1*) flanked by two non-functional parts of a tetracycline resistance gene (*tetA* 5' and *tetA* 3') which share a 627 bp region of 100% homology. A recombination event is required for the gentamicin resistance gene to be removed and the tetracycline resistance *tetA* gene to be whole, leading to a 4,500 bp plasmid. Therefore, without recombination the plasmid confers ampicillin and gentamicin resistance; and with recombination the plasmid confers ampicillin and tetracycline resistance (Fig. 2.1).

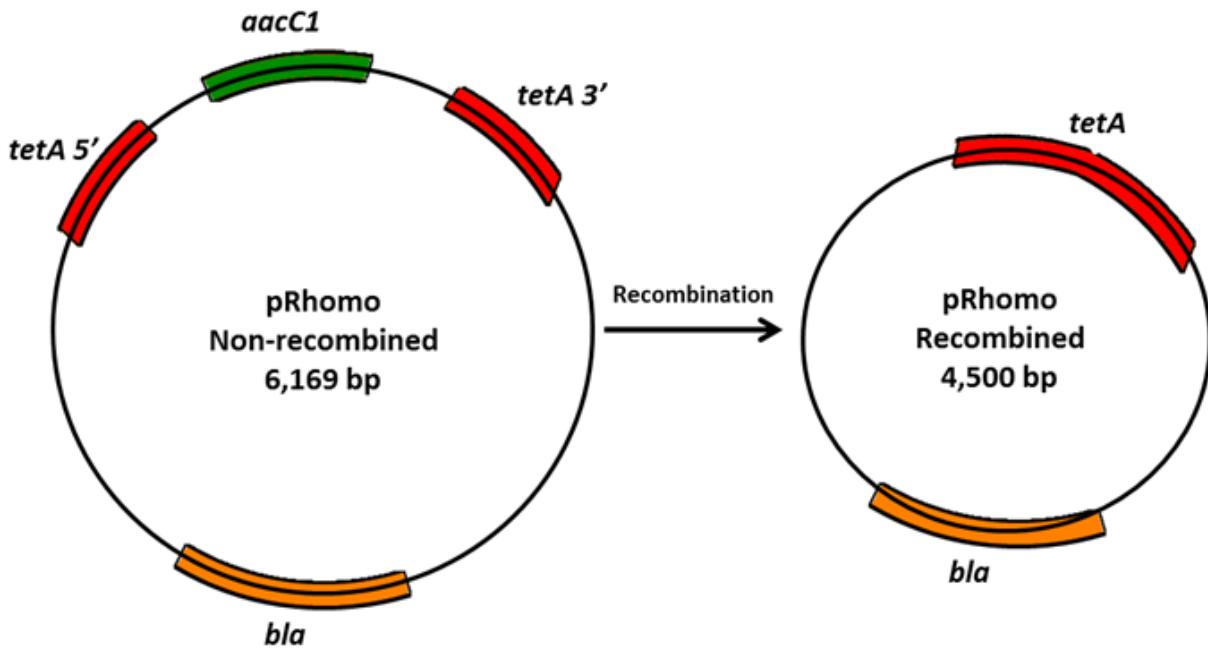


Figure 2.1. The pRhomo plasmid before and after recombination. The non-recombined version of the pRhomo plasmid contains the gentamicin resistance gene (*aacC1*) flanked by two parts of the tetracycline resistance gene (*tetA 5'* and *tetA 3'*). The two parts have an overlapping homologous region of 627 bp. After a recombination event, the two parts are restored as a functional tetracycline resistance gene (*tetA*), and the gentamicin resistance gene is removed. Figure adapted from Supplementary Figure 1, Rodríguez-Beltrán (2015)¹⁴⁴.

2.2.2 Antibiotics

For the purposes of bacterial selection, ampicillin was used at $100 \mu\text{g ml}^{-1}$. Nalidixic acid and kanamycin were both used at $50 \mu\text{g ml}^{-1}$. Gentamicin and tetracycline were both used at $10 \mu\text{g ml}^{-1}$. Streptomycin was used at $20 \mu\text{g ml}^{-1}$ and rifampicin was used at $80 \mu\text{g ml}^{-1}$. For antibiotic treatment of mice, Nalidixic acid was used at $10 \mu\text{g ml}^{-1}$ total concentration in the drinking water. Each antibiotic was resuspended in sterile Milli-Q water to make 1000x stock solutions.

Table 2.11. Antibiotics

Item	Catalogue #	Supplier
Ampicillin sodium salt	A9518	Sigma-Aldrich
Nalidixic acid sodium salt	N4382	Sigma-Aldrich
Gentamicin sulphate salt	G1264	Sigma-Aldrich
Kanamycin	60615	Sigma-Aldrich
Tetracycline	A1685,0025	AppliChem
Streptomycin	S9137	Sigma-Aldrich
Rifampicin	R8883	Sigma-Aldrich

2.3 Methods

2.3.1 Bacterial strains and culture conditions

Bacteria were revived and grown from frozen stocks stored at -80°C in order to prevent adaptation of *C. rodentium* over multiple laboratory subcultures. Bacteria were grown with shaking at 200 revolutions per minute (RPM) at 37°C in LB-Lennox media (Fort Richard Laboratories Ltd., Auckland, New Zealand) or defined minimal media. Antibiotics (kanamycin [50 µg ml⁻¹], nalidixic acid [50 µg ml⁻¹]) were only added to the media if they were required for selection, unless otherwise stated.

2.3.2 Mice

Female 6-8 week old C57BL/6Elite mice were provided by the Vernon Jansen Unit (University of Auckland) and came from specific-pathogen free (SPF) stocks. All animals were housed in individually HEPA-filtered cages (IVC) with sterile bedding and free access to sterilised food and water. Autoclaved cardboard toilet rolls were added to the cages for recreational purposes and to encourage close proximity of infected and uninfected mice. Mice which received antibiotic treatment were given nalidixic acid in their drinking water to a total concentration of 10 µg ml⁻¹. Experiments were performed in accordance with the New Zealand Animal Welfare Act (1999) and institutional guidelines provided by the University of Auckland Animal Ethics Committee, which reviewed and approved these experiments under applications R1003 and R1496.

2.3.3 Infection of mice

Bacteria, grown overnight in LB-Lennox broth, were spun at 4.5k RPM for 5 minutes, and resuspended in a tenth of the volume of sterile PBS, producing a 10x concentrated inoculum. Animals were orally inoculated using a gavage needle with 200 µl of bacteria (approximately 10⁸ CFU) and biophotonic imaging used to determine correct delivery of bacteria to the stomach. The number of viable bacteria used as an inoculum was determined by retrospective plating onto LB-Lennox Agar. Stool samples were recovered aseptically at various time points after inoculation, and the number of viable bacteria per gram of stool was determined after homogenisation at 0.1 g ml⁻¹ in PBS and plating onto selective LB-Lennox Agar containing either 50 µg ml⁻¹ kanamycin (for selection of ICC180 or adapted *C. rodentium*) or 50 µg ml⁻¹ nalidixic acid (for selection of ICC169).

2.3.4 *In vivo* biophotonic imaging

Biophotonic imaging was used to noninvasively measure the bioluminescent signal emitted by *C. rodentium* ICC180 from anaesthetised mice to provide information regarding the localisation of the bacterium. Prior to being imaged, the abdominal area of each mouse was shaved using a Vidal Sasoon handheld facial hair trimmer to minimise any potential signal impedance by melanin within pigmented skin and fur. Bioluminescence (given as photons second⁻¹ cm⁻² sr⁻¹) was measured after gaseous anaesthesia with isoflurane using the IVIS® Kinetic camera system (Perkin Elmer). A photograph (reference image) was taken under low illumination before quantification of photons emitted from ICC180 at a binning of four over 1 minute using the Living Image software (Perkin Elmer). The sample shelf was set to position D (field of view, 12.5 cm). For anatomic localisation, a pseudocolour image representing light intensity (blue, least intense to red, most intense) was generated using the Living Image software and superimposed over the gray-scale reference image. Bioluminescence in specific regions of individual mice also was quantified using the region of interest tool in the Living Image software program (given as photons second⁻¹) and used to calculate Area Under Curve (AUC) values for each individual animal.

2.3.5 *In vivo* Evolution Experiment

Two groups (n=6) of mice were housed in IVCs in an animal facility and fed autoclaved food and water *ad libitum*. One group of mice received 10 µg ml⁻¹ nalidixic acid in their drinking water (antibiotic treated animals), and all subsequent animals in this group also received the antibiotic nalidixic acid. Mice were anaesthetised with isoflurane and infected with ICC180 as described in section 2.3.3. Stool samples were recovered aseptically at various time points after inoculation, and the number of viable bacteria per gram of stool was determined after homogenisation at 0.1 g ml⁻¹ in PBS and plating onto LB-Lennox Agar containing 50 µg ml⁻¹ kanamycin. After 7 days, the ten animals (five from each group) with the highest bacterial burden were separated into 10 different “lineages” in separate cages, with no mingling between the lineages occurring henceforth. Each of these animals was dubbed passage (P) 0. An uninfected naïve animal was housed with each of these infected animals and allowed to be infected via natural grooming and coprophagy practices – this animal was dubbed P1. 7 days later (at the peak of infection), a new uninfected animal was housed with P1 and allowed to be infected again via natural processes – this new animal was then dubbed P2. The process was repeated over a period of 5 months, until the final passage number P22 was reached. On days of co-mingling, when the new uninfected animal was housed with the latest passage number, the mouse prior to the latest was culled under isoflurane and ketamine/xylazine anaesthetic, with blood removed via cardiac puncture, and the colon and rectum removed via dissection. The serum from the blood was then isolated using

Microvette 500 Z-Gel (Sarstedt) and stored at -20°C for future analysis, and colons snap frozen in a dry ice and ethanol bath before storage at -80°C.

2.3.6 *In vivo* competition experiments

Groups of six C57BL/6 mice were housed in IVCs in an animal facility and fed autoclaved food and water *ad libitum*. The mice were orally inoculated using gavage needles with 200 µl of a 1:1 mix of ICC169 and the competing *C. rodentium* strain (containing approximately 10⁸ CFU of bacteria) and biophotonic imaging used to determine correct delivery of bacteria to the stomach. The number of viable bacteria used as an inoculum was determined by retrospective plating onto LB-Lennox Agar containing either nalidixic acid or kanamycin. Stool samples were recovered aseptically at various time points after inoculation, and the number of viable bacteria per gram of stool was determined after homogenisation at 0.1 g ml⁻¹ in PBS and plating onto LB-Lennox Agar containing the appropriate antibiotics. The number and ratio of colonies growing on each antibiotic was used to calculate AUC values.

2.3.7 Evolution of *C. rodentium* ICC180 to defined laboratory media

Six 50 ml conical polypropylene tubes were inoculated with 15 ml defined minimal media (supplemented with 1% glucose) with 50 µl (approximately 10⁸ CFU) of ICC180. The tubes were incubated at 37°C with shaking at 200 RPM for 24 hours, allowing approximately 8 generations to occur. 50 µl was subcultured into fresh media every 24 hours, and samples were stored at -80°C with 25% v/v glycerol every 7 days (approximately every 50 generations). The experiment was continued until day 140 (the same number of days as the *in vivo* mouse evolution), at which point approximately 1,200 generations had occurred.

2.3.8 Attachment of *Citrobacter rodentium* to mouse fibroblast cells

Mouse fibroblasts (L-929 cells, obtained from the American Type Culture Collection [ATCC] through Cryosite Ltd.), were grown in DMEM + 10% v/v foetal calf serum, in a staggered layout within a 96 well plate, till confluent (approximately 10⁵ fibroblast cells per well). Media only controls were also present. Bacteria grown overnight in LB-Lennox broth were grown in DMEM + 10% v/v foetal calf serum overnight, and then diluted in fresh media and 100 µl (10⁶ bacterial cells) added to wells containing cells or the media only control wells. The plates were incubated without shaking for 1 hour. The bacteria were gently washed out of the wells with 200 µl of sterile PBS three times, and then the plates were imaged using the VICTOR X Light Plate reader (Perkin Elmer). The relative light units

(RLU) measured from wells with no fibroblast cells were subtracted from the wells with fibroblast cells, to determine the relative number of bacteria which remained attached to the fibroblasts. Experiments were repeated with new biological repeats of bacteria and fibroblasts on different days on 3 separate occasions.

2.3.9 Evaluation of desiccation resistance

C. rodentium was grown overnight from freezer stocks in LB-Lennox broth. Bacteria were washed in sterile PBS to remove excess nutrients, and 20 µl of culture (approximately 10^7 cells) added to sufficient 1.7 ml tubes to allow a tube to be revived at each time point. The tubes were left open to dry at room temperature for the duration of the experiment. Every 24 hours, 40 µL of sterile PBS was added to a tube and vortexed for approximately 20 seconds to ensure resuspension of the dried bacteria, then plated onto selective media. Experiments were repeated with 4 independent biological repeats.

2.3.10 Bacterial staining

The Gram Stain was performed on heat fixed bacterial samples, as described previously¹⁴⁵. Briefly, heat fixed samples were submerged in Crystal Violet for 60 seconds, then gently rinsed with water, then stained with Iodine for 60 seconds and rinsed again. Acetone was used as a decolourising agent, and added to the stained samples for approximately 10 seconds, or until the stain on densely stained regions had ceased lifting from the slide. The samples were then rinsed, and submerged in the counter-stain Safranin for 60 seconds. Samples were rinsed again, gently blotted, and then viewed using light microscopy. Photographs of stained samples were taken using the Leica AirLab application (Leica Microsystems).

Calcofluor white is a fluorochrome which binds to β-1,3 and β-1,4 polysaccharides, such as cellulose and chitin^{146,147}. Uropathogenic *Escherichia coli* (UPEC) strain 536 grown in Roswell Park Memorial Institute (RPMI) media was used as a positive control¹⁴⁷. Aggregates were heat fixed on sterile microscope slides, and 20 µl of calcofluor white was added to the samples. The stained sample was covered with a glass coverslip and the slides observed using light microscopy and fluorescence determined using UV-2A light.

2.3.11 Assessment of spontaneous mutation rate

C. rodentium was grown overnight from freezer stocks in LB-Lennox broth, and then 10 µl subcultured into 1.5 ml of LB-Lennox broth and grown overnight. 50 µl of the culture was plated onto LB agar plates containing 80 µg ml⁻¹ rifampicin, or a 1:10 diluted culture of 50 µl plated on LB agar containing 20 µg ml⁻¹ streptomycin. Cultures were also diluted onto LB agar with no selection to retrospectively determine the dose on each plate, in order to determine the proportion of resistant colonies. The percentage of resistant colonies was used as a proxy for mutation rate¹⁴⁸. Experiments were repeated on different days with independent cultures at least nine times.

2.3.12 Assessment of recombination rate using pRhomo

C. rodentium was grown overnight from freezer stocks in LB-Lennox broth, and then made chemically competent by the previously published Hanahan protocol^{149,150}. Competent *C. rodentium* was then transformed with the pRhomo plasmid¹⁴⁴ via heat-shock at 42°C for 60 seconds, and recovered with SOC media for 1 hour, then plated onto LB agar plates containing ampicillin (100 µg ml⁻¹) and gentamicin (10 µg ml⁻¹), and ampicillin (100 µg ml⁻¹) and tetracycline (10 µg ml⁻¹), and grown overnight at 37°C. The ampicillin and tetracycline containing plate is included to ensure that no already recombined plasmid is transformed into the bacteria. Five separate transformants were selected and inoculated in 1 ml of LB-Lennox broth with ampicillin, and grown overnight. Serial dilutions of the overnight culture was plated onto LB agar containing ampicillin and gentamicin, and ampicillin and tetracycline, and the ratio of gentamicin resistant:tetracycline resistant colonies assessed to determine the recombination rate. Experiments were repeated on four separate occasions with freshly made batches of competent bacteria.

2.3.13 In vitro growth curves.

10 ml of either LB-Lennox or defined minimal media was inoculated with 20 µl of a culture grown overnight in LB-Lennox broth. Samples were removed at regular intervals to measure bioluminescence, using a VICTOR X Light Plate reader (Perkin Elmer), and viable counts, by plating onto LB-Lennox Agar (Fort Richard Laboratories Ltd., Auckland, New Zealand). Overnight cultures were plated to retrospectively determine the initial inocula. Experiments were performed on five separate occasions and results used to calculate Area Under Curve values for each strain.

2.3.14 Genome sequencing

Genomic DNA was prepared from bacteria grown overnight in LB-Lennox broth using Genomic-tips 20/G (Qiagen). Isolated DNA was tested for purity and concentration using a NanoDrop™ (Thermo Scientific) and run on a 1% TAE agarose gel at 80V for 2 hours. Whole genome sequencing was performed using the Illumina MiSeq platform by NZGL (New Zealand). Data was analysed using a virtual machine running Bio-Linux 7¹⁵¹. Reads were mapped to the reference strain *C. rodentium* ICC168 (Genbank accession number FN543502.1) and analysed using BreSeq version 0.24rc6, which identified predicted mutations that were statistically valid. To determine any changes to plasmids, reads were also mapped to each of the known *C. rodentium* plasmids: pCROD1 (Genbank accession number FN543503.1), pCROD2 (Genbank accession number FN543504.1), pCROD3 (Genbank accession number FN543505.1), and pCRP3 (Genbank accession number NC_003114). MIRA 4 was used to assemble unmapped reads into contigs to investigate the presence of new genetic information¹⁵².

For investigation of the microbial communities present in the stool, mouse stools were snap frozen in a dry ice and ethanol bath, and then extracted for 16S sequencing using a MoBio PowerSoil® kit (Qiagen).

2.3.15 Statistical analyses

Data handling and statistical analyses were determined following a flowchart “The BioStat Decision Tool”, developed by Dr. Siouxsie Wiles and Dr. Anne Bishop (<http://flexiblelearning.auckland.ac.nz/biostat-tree/>). All significance levels cut-offs were arbitrarily set to 0.05, unless specified otherwise. Data was analysed using GraphPad Prism 6. Data was tested for normality using D’agostino-Pearson test; data which failed normality was analysed using a non-parametric test, data which passed normality was analysed using a parametric test. Two-tailed tests were used to test the hypothesis that the evolved *C. rodentium* would have either an increased or a decreased measure compared with the ancestral ICC180. When comparing multiple experimental groups, Dunn’s post hoc multiple comparison test was applied.

Chapter 3: Comparison of *C. rodentium* strains ICC169 and ICC180

3.1 Introduction

The *in vivo* evolution experiment was performed using a bioluminescently-tagged *C. rodentium* strain ICC180, which enables monitoring of bacterial burden and location within each infected mouse without necessitating euthanasia. All *in vivo*-adapted *C. rodentium* isolates, however, would be unable to be directly competed with the ancestral ICC180 strain, as it is inherently impossible to distinguish between the ancestral ICC180 strain and the *in vivo*-adapted *C. rodentium* strains, both being bioluminescently-tagged and resistant to kanamycin and nalidixic acid (Table 3.1). Addition of further selective markers to distinguish between the ancestor and the *in vivo*-adapted strains would carry the risk of impacting on the fitness of the changed strain, therefore necessitating assessing the neutrality of the change. Similarly, any molecular manipulation would require further subculturing steps in laboratory media, providing an opportunity for adaptation to the laboratory environment.

Therefore, an alternative strategy was devised: the *in vivo*-adapted strains and their ancestor would each be competed with *C. rodentium* ICC169, the non-bioluminescent, kanamycin sensitive, nalidixic acid resistant parent of ICC180⁹³ (Table 3.1). ICC180 had been shown previously to be equally as fit as ICC169 at infecting mice⁹³, allowing confidence that any differences between the strains would be negligible. ICC180 and the wild type ICC169 are distinguishable by differences in antibiotic resistance, as well as by the presence or absence of bioluminescent colonies. The *in vivo*-adapted strains and the ICC180 ancestral strain could therefore be indirectly competed by comparing ICC169 with ICC180 and ICC169 with the *in vivo*-adapted strains.

To facilitate this, ICC169 and ICC180 were fully characterised and competed to provide a solid benchmark for future comparisons between the ‘evolved’ strains. This comprehensive comparison of ICC169 and ICC180 was subsequently published in PeerJ (2016). This paper is reproduced here, reformatted and abridged for consistency purposes, with permission from the co-authors and PeerJ¹⁵³. Please note that there is a minor error in Figure 5 of the article: the x (bottom) axis on Fig. 5C and Fig. 5E should read “Time Post-infection (**Days**)”, not “Time Post-infection (**Hours**)”. At the time of submission this error is in the process of being rectified.

Table 3.1. Characteristics of the *Citrobacter rodentium* strains used in this study

Strain	Characteristics*	Reference
ICC180	Nal ^R , Km ^R , lux+	Wiles, 2004 ⁹³
<i>In vivo</i> -adapted <i>C. rodentium</i>	Nal ^R , Km ^R , lux+	This study
ICC169	Nal ^R	Wiles, 2005 ¹⁴²

*Key: Nal^R, resistant to nalidixic acid; Kan^R, resistant to kanamycin; lux+, bioluminescent.

3.2 The *in vitro* and *in vivo* effects of constitutive light expression on a bioluminescent strain of the mouse enteropathogen *Citrobacter rodentium*

Read H,^{1,2} Mills G,^{1,2} Johnson S,^{1,2} Tsai P,³ Dalton J,^{1,2,4} Barquist L,⁵ Print C,^{2,3,4} Patrick WM,^{4,6} Wiles S.^{1,2,4}

¹Bioluminescent Superbugs Lab, ²Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand; ³Bioinformatics Institute, School of Biological Sciences, University of Auckland, New Zealand; ⁴Maurice Wilkins Centre for Molecular Biodiscovery, New Zealand; ⁵Institute for Molecular Infection Biology, University of Würzburg, Würzburg, Germany; ⁶Department of Biochemistry, University of Otago, Dunedin, New Zealand.

3.2.1 Introduction

Bioluminescence is the by-product of a chemical reaction which has evolved in a wide variety of creatures for different purposes. This ‘living light’ allows fireflies like *Photinus pyralis* to find a mate¹⁵⁴, larvae like the New Zealand glow worm *Arachnocampa luminosa* to lure prey¹⁵⁵, and the bacterium *Aliivibrio fischeri* (formally *Vibrio fischeri*) to camouflage its nocturnal symbiont, the Hawaiian bobtail squid, while hunting¹⁵⁶. Bioluminescence is produced by the oxidation of a substrate (a luciferin) by an enzyme (a luciferase), which usually requires energy and oxygen. Cloning of the bioluminescence genes from *P. pyralis*¹⁵⁷, *V. fischeri*¹⁵⁸ and *Photorhabdus luminescens*¹⁵⁹, has let researchers use light production as a real-time non-invasive and non-destructive surrogate measure of microbial numbers in a wide variety of different culture environments, including within laboratory animals¹⁶⁰. This has proven particularly useful for studying microorganisms which take several weeks to grow on selective media, such as the bacterium *Mycobacterium tuberculosis*^{161,162}. As bioluminescence requires microbial metabolites, such as ATP and reduced flavin mononucleotide (FMNH₂), tagging microorganisms with luciferases means only live, metabolically active cells are detected.

Of the available bioluminescent reporter systems, the most widely used in bacteriology research is the bacterial luminescence reaction, encoded by the *lux* gene operon. The reaction involves the oxidation of a long chain aldehyde and FMNH₂, resulting in the production of oxidised flavin (FMN), a long chain fatty acid, and the emission of light at 490 nm¹⁶³. The reaction is catalysed by bacterial luciferase, a 77 kDa enzyme made up of an alpha and a beta subunit encoded by the *luxA* and *luxB* genes, respectively. The *luxC*, *D* and *E* genes encode the subunits of a multi-enzyme complex responsible for regenerating the aldehyde substrate from the fatty acid produced by the reaction. A significant advantage of the bacterial bioluminescence system is the ability to express the biosynthetic enzymes for substrate synthesis, allowing light to be produced constitutively. One of the underlying motivations for using

lux-tagged bacteria is the reduction in the number of animals needed for *in vivo* experiments, an ethical and legislative requirement in many countries. Using a technique known as biophotonic imaging, tagged bacteria can be non-invasively and non-destructively visualised and quantified on multiple occasions from within the same group of infected animals, whereas culture based techniques need groups of animals to be euthanised at each time point of interest¹⁶⁰. However, very little is known about the impact of constitutive light expression on tagged bacteria. We hypothesise that light production will impose a metabolic burden on the tagged bacteria, with the actual fitness costs dependent on the host bacterial species, the site of insertion of the bioluminescence genes and their expression levels.

We have previously made a lux-tagged derivative of *Citrobacter rodentium*⁹³, a bacterium that infects laboratory mice using the same virulence mechanisms as the life-threatening pathogens, enteropathogenic *Escherichia coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC) use to infect humans^{88,164}. *C. rodentium* ICC180 contains a single chromosomally-located copy of the *lux* operon from *P. luminescens*, alongside a gene for resistance to the antibiotic kanamycin. We have previously non-invasively tracked ICC180 during infection of mice⁹⁴, demonstrating that *C. rodentium* rapidly spreads between infected and uninfected animals and that bacteria shed from infected mice are 1,000 times more infectious than laboratory grown bacteria¹³⁸. While we have shown that ICC180 can reach similar numbers within the gastro-intestinal tracts of infected mice and causes similar pathology when compared to its non-bioluminescent parent strain ICC169⁹³, we have never fully investigated the impact of constitutive light expression on the fitness of ICC180.

In this study we set out to determine whether *C. rodentium* ICC180 has a competitive disadvantage when competed against its non-bioluminescent parent ICC169 in a range of *in vitro* and *in vivo* environments. We also sequenced the genome and associated plasmids of ICC180 to determine whether there were any other genetic differences between the two strains, perhaps as a result of the transposon mutagenesis technique¹⁶⁵ used to generate ICC180. Finally, we compared the growth profiles of the two strains using the BIOLOG Phenotypic Microarray (PM) system, a rapid 96-well microtitre plate assay for phenotypically profiling microorganisms based on their growth under approximately 2000 different metabolic conditions¹⁶⁶.

3.2.2 Materials and methods

Bacterial strains and culture conditions.

The bacterial strains used in this study were *Citrobacter rodentium* ICC169 (spontaneous nalidixic acid resistant mutant)¹⁴² and ICC180 (nalidixic acid and kanamycin resistant)⁹³. Bacteria were revived and grown from frozen stocks stored at -80°C in order to prevent adaptation of *C. rodentium* over multiple laboratory subcultures. Bacteria were grown at 37°C with shaking at 200 revolutions per minute (RPM) in LB-Lennox media (Fort Richard Laboratories Ltd., Auckland, New Zealand) or in defined minimal media (modified Davis & Mingioli media¹⁶⁷), containing ammonium sulphate [1 g l⁻¹], potassium dihydrogen phosphate [4.5 g l⁻¹], dipotassium hydrogen phosphate anhydrous [10.5 g l⁻¹], sodium citrate dihydrate [5 g l⁻¹], magnesium sulfate heptahydrate [24.65 mg l⁻¹], thiamine [0.5 mg l⁻¹], supplemented with 1% glucose) at 37°C. Antibiotics (kanamycin [50 µg ml⁻¹], nalidixic acid [50 µg ml⁻¹]) were only added to the media if they were required for selection. All chemicals and antibiotics were obtained from Sigma-Aldrich (Australia).

Genome sequencing and analysis.

Genomic DNA was prepared from bacteria grown overnight in LB-Lennox broth. Whole genome sequencing was performed using the Illumina HiSeq platform by BGI (Hong Kong). A total of 3,414,820 paired-end 90 bp reads were generated for ICC169 and 3,369,194 for ICC180. Data was quality trimmed using DynamicTrim¹⁶⁸ (minimum Phred score 25) and filtering of reads shorter than 45 bp after quality trimming was performed using LengthSort¹⁶⁸; both programmes are part of the SolexaQA software package¹⁶⁸. After filtering, 2,444,336 paired reads were retained for ICC169 and 2,383,491 for ICC180. All remaining high quality and properly paired reads were mapped to the reference strain *C. rodentium* ICC168 (Genbank accession number FN543502.1¹⁰²) using the default settings in BWA¹⁶⁹. On average, 95% of all high quality reads mapped uniquely to ICC168 (94.8% for ICC169 and 95.2% for ICC180) and single nucleotide polymorphisms (SNPs) and indels that were present only in ICC180 at 100% were identified using Samtools mpileup¹⁷⁰. SNPs and indels were confirmed by PCR and sequencing. In addition, the reads were also analysed using BreSeq version 0.24rc6¹⁷¹, which identified predicted mutations that were statistically valid. To locate the insertion site of the *lux* operon and kanamycin resistance (Km^R) gene, we first performed de novo assembly on quality trimmed data for ICC180 using EDENA v3.0¹⁷². All assembled contigs were mapped to the *C. rodentium* reference strain ICC168 using Geneious¹⁷³ and contigs unmapped to ICC168 were BLAST searched against the *lux* operon and Km^R gene. We located both the *lux* operon and Km^R gene on an unmapped contig 117,921 bp long. To identify the position of this contig, we broke the contig into two segments based on the location of *lux* operon and Km^R gene positions on the contig, and performed additional reference mapping to ICC168 to identify the insertion site. To determine changes

to the plasmids present in *C. rodentium*, reads were also mapped to the sequenced plasmids pCROD1 (Genbank accession number FN543503.1), pCROD2 (Genbank accession number FN543504.1), pCROD3 (Genbank accession number FN543505.1), and pCRP3 (Genbank accession number NC_003114).

Phenotypic microarrays.

Phenotypic microarrays were performed by BIOLOG Inc. (California, USA) as described previously¹⁶⁶. Assays were performed in duplicate using plates PM1-20 (Supplementary Table 1). The data was exported and analysed in the software package R as previously described¹⁷⁴. Briefly, growth curves were transformed into Signal Values (SVs)¹⁷⁵ summarising the growth over time while correcting for background signal. Principal component analyses showed a clear separation by genotype, suggesting reproducible differences in metabolism between the two strains. A histogram of log signal values displayed a clear bimodal distribution, which we interpreted as representing non-respiring cells ('off', low SV) and respiring cells ('on', high SV), respectively. Normal distributions were fitted to these two distributions using the R MASS package, and these models were then used to compute log-odds ratios for each well describing the probability that each observation originated from the 'on' or 'off' distribution. Wells which were at least 4 times more likely to come from the 'on' distribution than the 'off' in both replicates were considered to be actively respiring. In order to determine the significance of observed differences between genotypes, we applied the moderated t-test implemented in the limma R/Bioconductor package¹⁷⁶. Wells with a Benjamini-Hochberg corrected P-value of less than 0.05, that is allowing for a false discovery rate of 5%, and which were called as actively respiring for at least one genotype, were retained for further analysis. The data was also analysed using the DuctApe software suite¹⁷⁷. Growth curves were analysed using the dphenome module, with the background signal subtracted from each well. Based on the results of an elbow test (Supplementary Fig.1), 7 clusters were chosen for k-means clustering. An Activity Index (AV) was created based on the clustering, ranging from 0 (minimal activity) to 6 (maximal activity). AV data was visualised using the plot and ring commands of the dphenome module (Supplementary Fig.2).

In vitro growth experiments.

Briefly, for individual growth curves, 10 ml of either LB-Lennox or defined minimal medium was inoculated with 20 µl of a culture grown overnight in LB-Lennox broth. Cultures were grown at 37°C with shaking at 200 RPM and samples were removed at regular intervals to measure bioluminescence, using a VICTOR X Light Plate reader (Perkin Elmer), and viable counts, by plating onto LB-Lennox Agar (Fort Richard Laboratories Ltd., Auckland, New Zealand). Overnight cultures were plated to

determine the initial inocula. Experiments were performed on seven separate occasions and results used to calculate area under curve (AUC) values for each strain. For the competition experiments, 10 µl of a culture grown overnight in LB-Lennox broth was used to inoculate 1 ml of defined minimal medium, with the mixed culture tubes receiving 5 µl of each strain. Inoculated tubes were incubated overnight at 37°C with shaking at 200 RPM, followed by serial dilution in sterile phosphate buffered saline (PBS) for plating onto LB Agar containing either nalidixic acid or kanamycin. The ratio of colonies that grew on media containing each antibiotic was used to determine the proportion of each strain remaining. Experiments were performed on eight separate occasions and the results used to calculate AUC values and competitive indices (CI). CI's were calculated as follows: $CI = [\text{strain of interest output}/\text{competing strain output}]/[\text{strain of interest input}/\text{competing strain input}]^{178,179}$.

Infection of *Galleria mellonella*.

5th instar *Galleria mellonella* larvae (waxworms) were obtained from a commercial supplier (Biosuppliers.com, Auckland, New Zealand). Bacteria were grown overnight in LB-Lennox broth and used to infect waxworms which were pale in colour and weighed approximately 100-200 mg. Waxworms were injected into one of the last set of prolegs with 20 µl of approximately 10^8 colony forming units (CFU) of bacteria using a 1ml fine needle insulin syringe. Waxworms were injected with either ICC169, ICC180 or a 1:1 mix and incubated at 37°C. Throughout the course of a 24 h infection, individual waxworms were inspected for phenotypic changes and scored using a standardised method for assessing waxworm health (the Caterpillar Health Index [CHI]) which we have developed. Briefly, waxworms were monitored for movement, cocoon formation, melanisation, and survival. Together, these data form a numerical scale, with lower CHI scores corresponding with more serious infections and higher scores with healthier waxworms. Scores were used to calculate AUC values. Bioluminescence (given as relative light units [RLU]) was measured at regular intervals from waxworms infected with ICC180 (Supplementary Fig.3). Waxworms were placed into individual wells of a dark OptiPlate-96 well microtitre plate (Perkin Elmer) and bioluminescence measured for 1 second to provide relative light units (RLU)/second using the VICTOR X Light Plate reader. Waxworms infected with ICC169 were used as a control. Following death, or at 24 h, waxworms were homogenised in PBS and plated onto LB-Lennox Agar containing the appropriate antibiotics. Independent experiments were performed three times using 10 waxworms per group.

Infection of Mice.

Female 6-7 week old C57BL/6Elite mice were provided by the Vernon Jansen Unit (University of Auckland) from specific-pathogen free (SPF) stocks. All animals were housed in individually HEPA-filtered cages with sterile bedding and free access to sterilised food water. Experiments were performed in accordance with the New Zealand Animal Welfare Act (1999) and institutional guidelines provided by the University of Auckland Animal Ethics Committee, which reviewed and approved these experiments under applications R1003 and R1496. Bacteria grown overnight in LB-Lennox broth were spun at 4500 RPM for 5 minutes, and resuspended in a tenth of the volume of sterile PBS, producing a 10x concentrated inoculum. Animals were orally inoculated using a gavage needles with 200 µl of either ICC169, ICC180, or a 1:1 mix (containing approximately 10^8 CFU of bacteria) and biophotonic imaging used to determine correct delivery of bacteria to the stomach. The number of viable bacteria used as an inoculum was determined by retrospective plating onto LB-Lennox Agar containing either nalidixic acid or kanamycin. Stool samples were recovered aseptically at various time points after inoculation, and the number of viable bacteria per gram of stool was determined after homogenisation at 0.1 g ml⁻¹ in PBS and plating onto LB-Lennox Agar containing the appropriate antibiotics. The number and ratio of colonies growing on each antibiotic was used to calculate AUC values and CI's as described above. Independent experiments were performed twice using 6 animals per group.

***In vivo* bioluminescence imaging.**

Biophotonic imaging was used to noninvasively measure the bioluminescent signal emitted by *C. rodentium* ICC180 from anaesthetised mice to provide information regarding the localisation of the bacterium. Prior to being imaged, the abdominal area of each mouse was shaved, using a Vidal Sasoon handheld facial hair trimmer, to minimise any potential signal impedance by melanin within pigmented skin and fur. Mice were anaesthetised with gaseous isoflurane and bioluminescence (given as photons second⁻¹ cm⁻² steradian [sr]⁻¹) was measured using the IVIS® Kinetic camera system (Perkin Elmer). A photograph (reference image) was taken under low illumination before quantification of photons emitted from ICC180 at a binning of four over 1 minute using the Living Image software (Perkin Elmer). The sample shelf was set to position D (field of view, 12.5 cm). For anatomic localisation, a pseudocolor image representing light intensity (blue, least intense to red, most intense) was generated using the Living Image software and superimposed over the gray-scale reference image. Bioluminescence in specific regions of individual mice also was quantified using the region of interest tool in the Living Image software program (given as photons second-1) and used to calculate AUC values for each individual animal.

Statistical analyses.

Data was analysed using GraphPad Prism 6. Data was tested for normality using the D'Agostino-Pearson test; data which failed normality was analysed using a non-parametric test, while data which passed normality was analysed using a parametric test. One-tailed tests were used to test the hypothesis that constitutively expressing light gives ICC180 a differential fitness cost compared to the non-bioluminescent parent strain ICC169. When comparing multiple experimental groups, Dunn's post hoc multiple comparison test was applied.

3.2.3 Results

Bioluminescent *Citrobacter rodentium* strain ICC180 has three altered chromosomal genes and a large deletion in plasmid pCROD1 in addition to insertion of the lux operon and kanamycin resistance gene.

We determined the whole genome draft sequences of *C. rodentium* ICC169 and ICC180 using Illumina sequence data. Compared with sequenced type strain ICC168 (Genbank accession number FN543502.1), both strains have a substitution of a guanine (G) to an adenine (A) residue at 2,475,894 bp, resulting in an amino acid change from serine (Ser) to phenylalanine (Phe) within gyrA, the DNA gyrase subunit, and conferring resistance to nalidixic acid. The sequencing data indicate that the lux operon and kanamycin resistance gene (a 7,759 bp fragment) has inserted at 5,212,273 bp, disrupting the coding region of a putative site-specific DNA recombinase (Figure 1). In addition to the presence of the *lux* operon and kanamycin resistance gene, we found that the genome of ICC180 differs from ICC169 by two single nucleotide polymorphisms (SNPs), a single base pair insertion (of a G residue at 3,326,092 bp which results in a frameshift mutation within ROD_31611, a putative membrane transporter) and a 90 bp deletion in deoR (deoxyribose operon repressor) (Table 1). All four plasmids previously described for *C. rodentium* were present in ICC180, however the largest of these plasmids, pCROD1, shows evidence of extensive deletion events and is missing 41 out of 60 genes (Supplementary Table 2).

Table 1. SNPs and indels that differ between the bioluminescent *C. rodentium* derivative ICC180 and its parent strain ICC169. Sequencing revealed three points of difference between ICC180 and ICC169. Two SNPs are present, each cytosine substitutions, and one guanine insertion inducing a frameshift mutation. Sequencing data was analysed using BreSeq¹⁷¹.

Position	Base change	Amino acid change	Gene	Function
2,936,285	T→C	D471G (GAC→GGC)	<i>ctsIV</i>	T6SS protein CtsIV
3,999,002	T→C	E89G (GAG→GGG)	<i>pflD</i>	Formate acetyltransferase 2
3,326,092	CAG→CAGG	Frameshift	<i>ROD_31611</i>	Major Facilitator Superfamily transporter

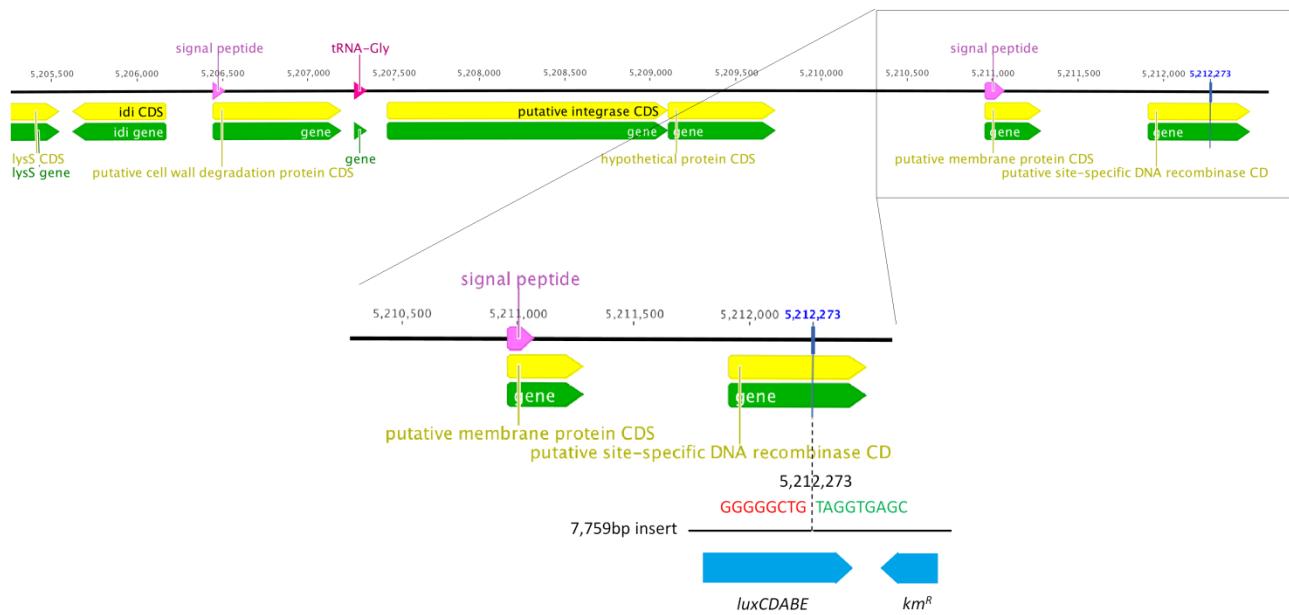


Figure 1. Whole genome sequencing shows that the *lux* operon and kanamycin resistance gene have inserted at position 5,212,273 in the chromosome of *C. rodentium* ICC180, disrupting a putative site-specific DNA recombinase.

Constitutive light expression does not greatly impact the metabolism of *C. rodentium* ICC180.

C. rodentium ICC169 and its bioluminescent derivative ICC180 were grown on two separate occasions using PM plates 1-20. We analysed the data using the DuctApe software suite which calculates an activity index (AV) for each strain in response to each well (Supplementary Fig.2). Next, the growth curve data were transformed into Signal Values (SVs) as previously described¹⁷⁴, summarising the growth of each strain over time for each well. Wells which were considered to be actively respiring were analysed using the moderated t-test implemented in the limma R/Bioconductor package¹⁷⁶. Those wells with a Benjamini-Hochberg corrected P-value of less than 0.05 are shown in Table 2 (with corresponding growth curves in Supplementary Fig. 3). Our results indicate that the growth of the two strains significantly differed ($p = <0.05$) in 26/1,920 wells. Of these >80% are from the PM11-20 plates, which belong to the chemical category, suggesting that the expression of bioluminescence is near-neutral in almost every non-toxic environment. The bioluminescent strain ICC180 is able to use D-glucosamine, cytidine and Ala-His as nitrogen sources, and inositol hexaphosphate as a phosphate source, and grew significantly better than ICC169 in the presence of 11 chemicals: the antibiotics kanamycin, paromomycin, geneticin, spiramycin, rolitetracycline, doxycycline, cefoxitin; the

quaternary ammonium salt dequalinium chloride; coumarin; iodonitrotetrazolium violet; and the acetaldehyde dehydrogenase inhibitor disulphiram (Table 2). That the expression of a kanamycin resistance gene also improves growth of ICC180 in the presence of related aminoglycosides is reassuring. In contrast, the wildtype strain ICC169 was able to use the nitrogen peptide Lys-Asp and grew significantly better in the presence of 8 chemicals: the metal chelators, EDTA and EGTA, sodium nitrate, the antibiotics rifampicin and phenethicillin, the fungicide oxycarboxin, the cyclic polypeptide colistin, the nucleoside analogue cytosine-1-b-D-arabinofuranoside and (Table 2). The fact that significant differences in growth rate were observed for so few conditions, provided robust and comprehensive evidence that light production is near-neutral in *C. rodentium* ICC180.

Table 2. Phenotypic microarray (PM) wells in which the growth of bioluminescent *C. rodentium* derivative ICC180 significantly differs from its non-bioluminescent parent strain ICC169.

PM Class	Substrate	Adjusted p value	Improved growth by ICC169	Improved growth by ICC180	Comment
Nitrogen	D-glucosamine	0.0159		✓	
	Cytidine	0.0280		✓	
	Ala-His	0.0316		✓	
Phosphate	Inositol hexaphosphate	0.0280		✓	
Nitrogen peptides	Lys-Asp	0.0306	✓		
Chemicals	Kanamycin	0.0076		✓	Conferred by KanR gene
	Paromomycin	0.0048		✓	Aminoglycoside -- the kanamycin cassette will be mediating resistance
	Geneticin	0.0048		✓	Aminoglycoside -- the kanamycin cassette will be mediating resistance
	Dequalinium chloride	0.0116		✓	Quaternary ammonium salt
	Spiramycin	0.0088		✓	Macrolide -- acts at ribosomal 50S, c.f. aminoglycosides at 30S
	Rolitetracycline	0.0316		✓	Tetracycline; prevents tRNA binding at 30S A-site
	Doxycycline	0.0210		✓	Tetracycline; prevents tRNA binding at 30S A-site
	Coumarin	0.0333		✓	Fragrant organic compound found in many plants

	Iodonitro tetrazolium violet (INT)	0.0087		✓	Electron acceptor, reduced by succinate dehydrogenase (and by superoxide radicals)
	EDTA	0.0048	✓		Metal chelator
	EGTA	0.0210	✓		Metal chelator
	Rifampicin	0.0048	✓		RNA polymerase inhibitor
	Colistin	0.0048	✓		Cyclic polypeptide; disrupts outer membrane
	Oxycarboxin	0.0121	✓		Fungicide
	Phenethicillin	0.0048	✓		Beta-lactam
	Cytosine-1-b-D-arabinofuranoside	0.0123	✓		Nucleoside analogue (anti-cancer/-viral)
	Sodium Nitrate	0.0306	✓		
	Cefoxitin	0.0316		✓	Beta-lactam
	Disulphiram	0.0349		✓	Inhibits acetaldehyde dehydrogenase

The growth of ICC180 is not impaired in a rich laboratory medium, when compared to its non-bioluminescent parent strain, but does exhibit an increased lag phase when grown in a restricted medium.

We grew ICC180 and ICC169 in rich (LB-Lennox) and restricted (minimal A salts with 1% glucose supplementation) laboratory media. For ICC180, we found that bioluminescence strongly correlated with the bacterial counts recovered throughout the growth period in both LB-Lennox (Spearman's $r = 0.9293$ [95% CI = 0.8828 - 0.9578], $p = <0.0001$) and the restricted medium (Spearman's $r = 0.9440$ [95% CI = 0.9001 - 0.9689], $p = <0.0001$) (Fig. 2A & B, 3A & B). We also found that the growth of each strain was comparable in LB-Lennox medium, with no significant difference between the bacterial counts recovered over 8 hours (Fig. 2B), as demonstrated by the calculated AUC values (Fig. 2C).

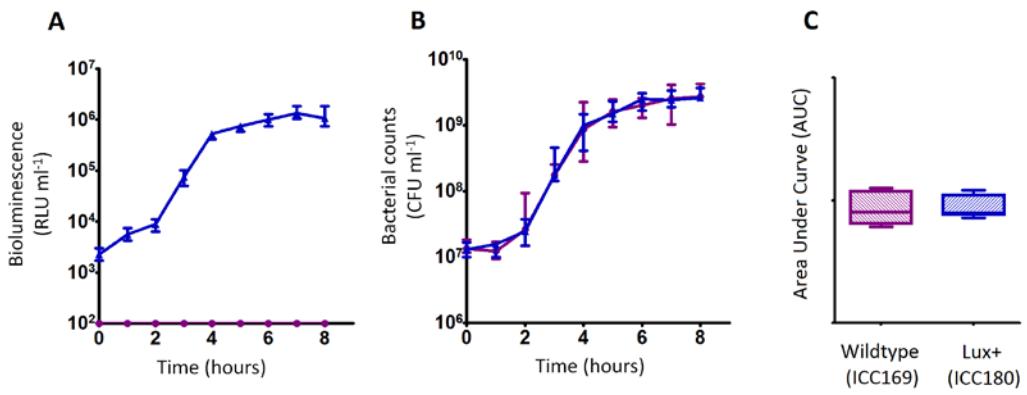


Figure 2. *C. rodentium* ICC180 is not impaired during growth in a rich laboratory medium when compared to its non-bioluminescent parent strain ICC169. Wildtype *C. rodentium* ICC169 (shown as purple circles) and its bioluminescent derivative ICC180 (shown as blue triangles) were grown in LB-Lennox broth and monitored for changes in bioluminescence (given as relative light units [RLU] ml⁻¹) (A) and bacterial counts (given as colony forming units [CFU] ml⁻¹) (B). Bacterial count data was used to calculate area under curve (AUC) values for each strain (C). Data (medians with ranges where appropriate) is presented from experiments performed on eight separate occasions.

We also found no significant difference between the bacterial counts recovered from ICC180 and ICC169 growing in the restricted medium for 14 hours (mean CFU 5.67×10^8 [SD 2.31×10^8] and 8.84×10^8 [SD 2.93×10^8], respectively). However, we did find a significant difference between the AUC values calculated from the bacterial counts recovered over the course of 14 hours ($p = 0.0078$, one-tailed Wilcoxon matched-pairs signed rank test) (Fig. 3C). We calculated the slopes of the growth curves and found that there was no difference in the rates of growth of the two strains during exponential phase. Instead, we found a significant difference between the slopes calculated during the first 4 hours of growth (1/slope values: ICC169 = 1.48×10^{-7} [SD 9.98×10^{-8}], ICC180 = 2.47×10^{-7} [SD 1.10×10^{-7}]; $p = 0.0041$, one-tailed Paired *t* test), suggesting ICC180 spends longer in lag phase than ICC169 when grown in the restricted medium used.

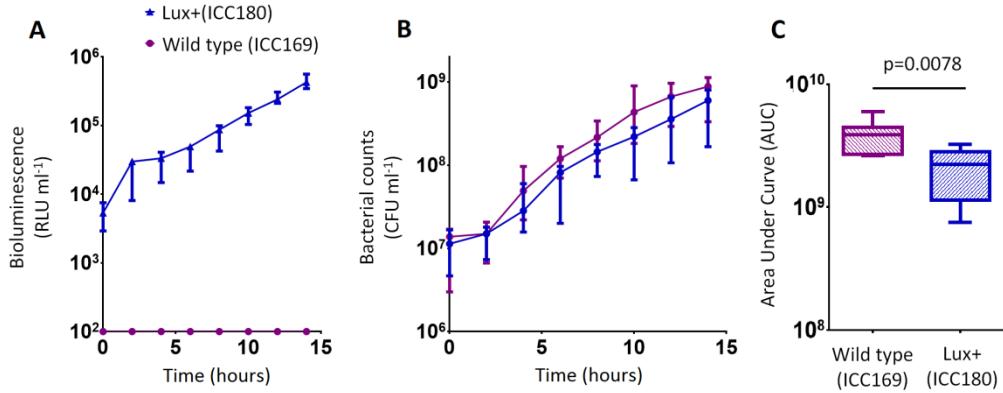


Figure 3. *C. rodentium* ICC180 is mildly impaired during growth in a defined minimal laboratory medium when compared to its non-bioluminescent parent strain ICC169. Wildtype *C. rodentium* ICC169 (shown as purple circles) and its bioluminescent derivative ICC180 (shown as blue triangles) were grown in minimal A salts supplemented with 1% glucose and monitored for changes in bioluminescence (given as relative light units [RLU] ml⁻¹) (A) and bacterial counts (given as colony forming units [CFU] ml⁻¹) (B). Bacterial count data was used to calculate area under curve (AUC) values for each strain, which were found to be significantly different ($p=0.0078$; Wilcoxon Matched pairs-signed rank test) (C). Data (medians with ranges where appropriate) is presented from experiments performed on eight separate occasions.

ICC180 is not impaired in the *Galleria mellonella* infection model.

We infected larvae of the Greater Wax Moth *G. mellonella* (waxworms) with ICC169 and ICC180 in single and 1:1 mixed infections. We monitored the waxworms over a 24-48 hour period for survival and disease symptoms. The Caterpillar Health Index (CHI) is a numerical scoring system which measures degree of melanisation, silk production, motility, and mortality. We found that the majority of infected waxworms succumb to *C. rodentium* infection (Fig. 4A), which is reflected by the concurrent decrease in CHI score (Fig. 4B). This is in contrast to waxworms injected with PBS, who all survived and consistently scored 9-10 on the CHI scale throughout the experiments. We also found that the survival and symptoms of waxworms infected with each strain were comparable, with no significant difference between the survival curves (Fig. 4A), and calculated AUC values for the CHI scores (Fig. 4C). However, when we directly compared ICC169 and ICC180 in mixed infections of approximately 1:1, we found a significant difference in the relative abundance of the bacteria recovered from waxworms at either time of death or 24 hours, whichever occurred first ($p = 0.001$, one-tailed Wilcoxon matched-pairs signed rank test). Despite a slightly lower infectious dose, higher numbers of ICC180 were consistently recovered from infected waxworms (Fig. 4D).

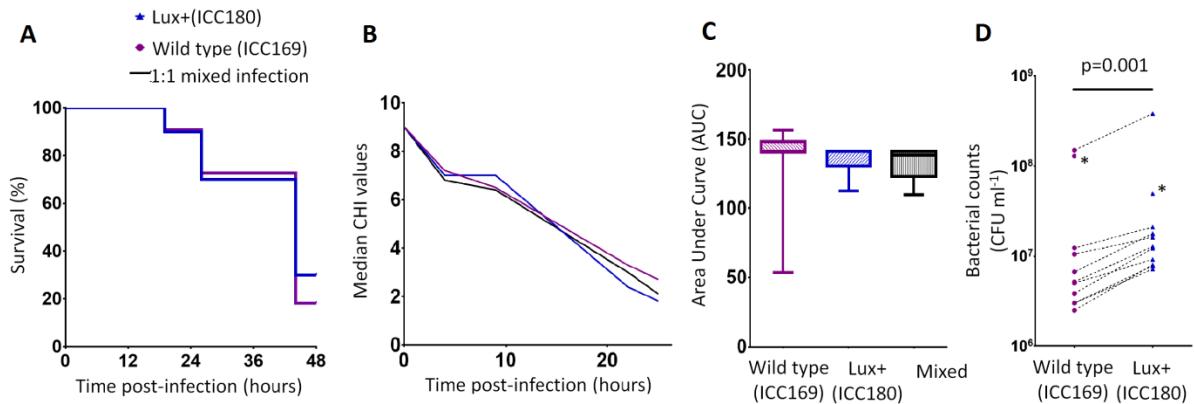


Figure 4. Bioluminescent *C. rodentium* ICC180 is not impaired in the *Galleria mellonella* infection model. Groups of larvae ($n = 10$) of the Greater Wax Moth *Galleria mellonella* were infected with ICC169 and ICC180 in single and 1:1 mixed infections and monitored for survival (%) (A) and for disease symptoms using the Caterpillar Health Index (CHI), a numerical scoring system which measures degree of melanisation, silk production, motility, and mortality (given as median CHI values) (B). Survival curves (A) and calculated area under curve (AUC) data of CHI scores reveals no difference between waxworm response to infection from either strain (C). Waxworms infected with a 1:1 mix of ICC169 and ICC180 were homogenised at 24-hours, or at time of death if earlier. Actual infecting doses for each strain were determined by retrospective plating, and are indicated by *. The bacterial burden of ICC180 and ICC169 in individual caterpillars (indicated by the dotted line), was calculated after plating onto differential media and found to be significantly different ($p=0.001$; one-tailed Wilcoxon matched pairs-signed rank test) (D). Data (medians with ranges where appropriate) is presented from experiments performed on 3 separate occasions, except (A) and (D), where the results of a representative experiment are shown.

ICC180 is impaired in mixed but not in single infections in mice when compared to its non-bioluminescent parent strain.

We orally gavaged groups of female 6-8 week old C57Bl/6 mice ($n=6$) with $\sim 5 \times 10^9$ CFU of ICC169 and ICC180, either individually or with a 1:1 ratio of each strain. We followed the infection dynamics by obtaining bacterial counts from stool samples (Fig. 5) and by monitoring bioluminescence from ICC180 using biophotonic imaging (Fig. 6). We found that the growth of each strain was comparable during single infections, with no significant difference between the bacterial counts recovered throughout the infection (Fig. 5A), as demonstrated by the calculated AUC values (Fig. 5B).

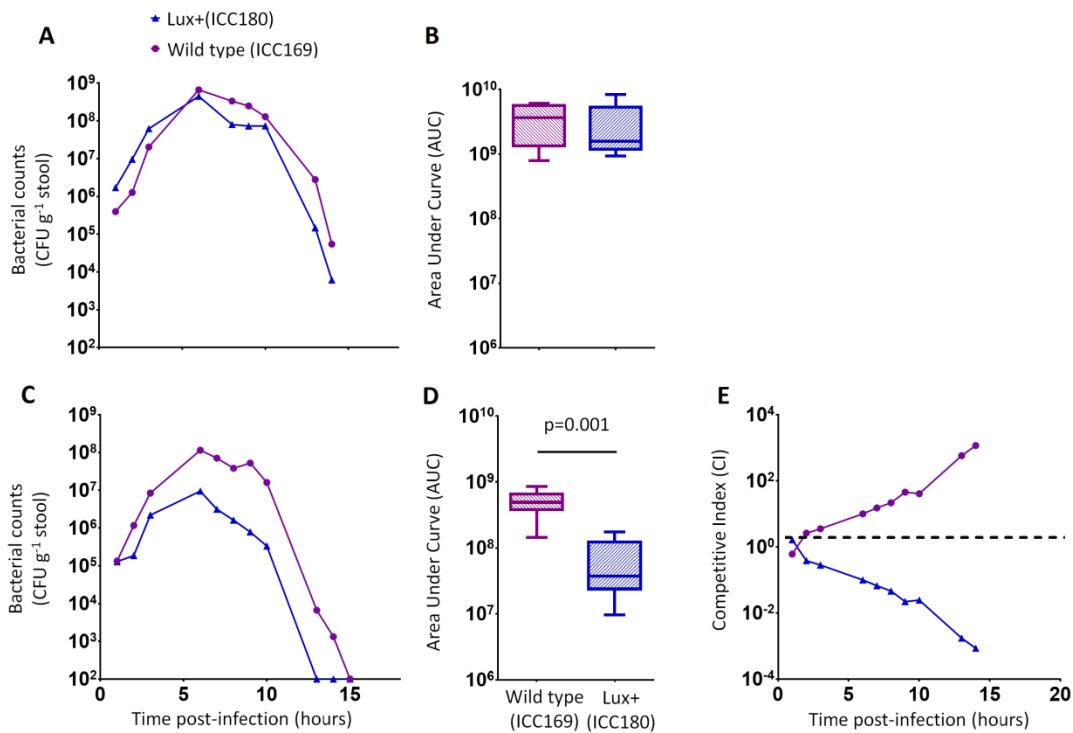


Figure 5. *C. rodentium* ICC180 is impaired during mixed, but not in single, infections in mice when compared to its non-bioluminescent parent strain ICC169. Groups of female 6-8 week old C57Bl/6 mice ($n=6$) were orally-gavaged with $\sim 5 \times 10^9$ CFU of wildtype *C. rodentium* ICC169 (shown as purple circles) and its bioluminescent derivative ICC180 (shown as blue triangles) in single infections (A, B) or 1:1 mixed infections (C, D) and monitored for changes in bacterial counts (given as colony forming units [CFU] g⁻¹ stool) (A, B). Bacterial count data was used to calculate area under curve (AUC) values for each strain in single (B) and mixed (D) infections, and were found to be significantly different only for the mixed infections ($p=0.001$; one-tailed Wilcoxon Matched pairs-signed rank test). This is reflected in the competitive indices (CI) calculated from the bacterial counts recovered during mixed infections, with ICC180 showing a growing competitive disadvantage from day 2 post-infection (E). Data (medians with ranges where appropriate) is presented from experiments performed on two separate occasions.

In contrast, we found a significant difference between the AUC values calculated from the bacterial counts recovered from ICC180 and ICC169 during mixed infections ($p = 0.001$, one-tailed Wilcoxon matched-pairs signed rank test) (Fig. 5D). Our data demonstrates that when in direct competition with ICC169, ICC180 is shed at consistently lower numbers from infected animals (Fig. 5C). At the peak of infection (days 6-8), this equates to over a 10-fold difference, with mice shedding a median of 1.195×10^8 CFU (SD 4.544×10^7) for ICC169 compared to 9.98×10^6 CFU (SD 1.544×10^7) for ICC180. This disadvantage is reflected in the Competitive Indices we calculated from bacterial counts recovered at each time point, which for ICC180 decreased steadily throughout the course of the infection (Fig. 5E). Despite this disadvantage, ICC180 is never completely outcompeted and remains detectable in the stools of infected animals until the clearance of infection (Fig. 5C), and by biophotonic imaging until day 10-13 post-infection (Fig. 6A).

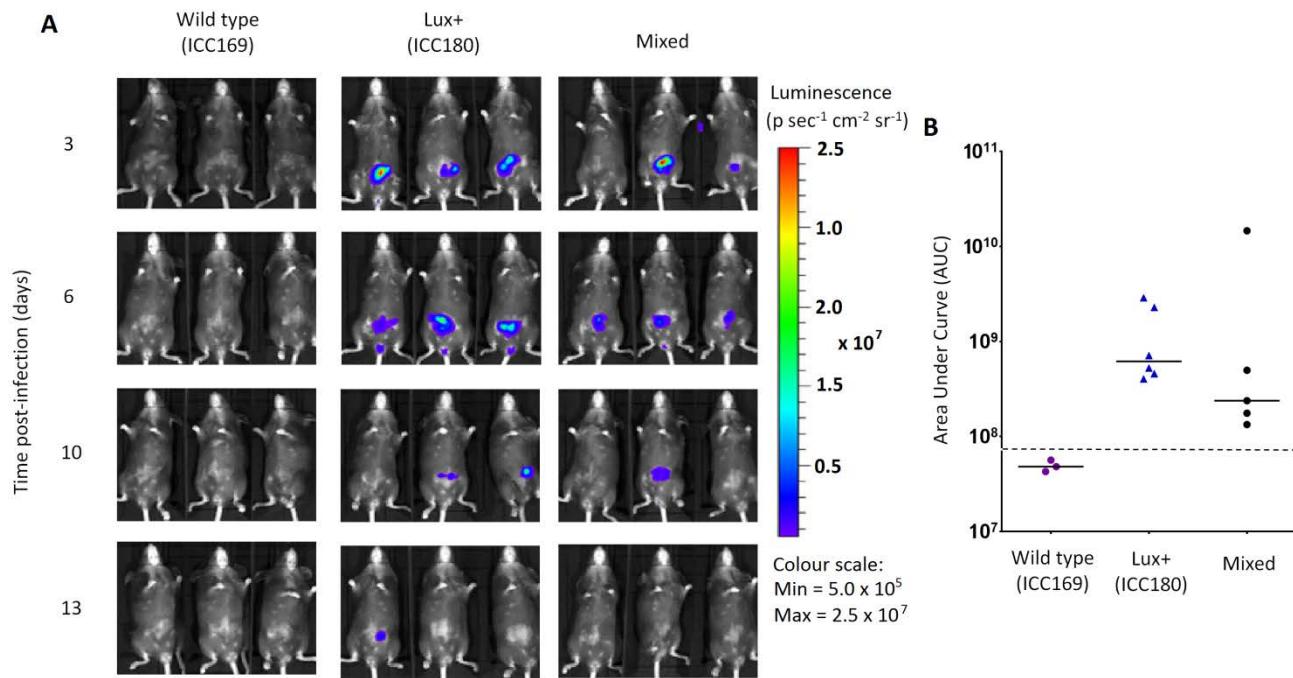


Figure 6. Despite having a fitness disadvantage in mixed infections of mice, ICC180 is still visible by biophotonic imaging. Groups of female 6-8 week old C57Bl/6 mice (n=6) were orally-gavaged with $\sim 5 \times 10^9$ CFU of wildtype *C. rodentium* ICC169 and its bioluminescent derivative ICC180 in single infections or 1:1 mixed infections. Mice were anaesthetised with gaseous isoflurane and bioluminescence (given as photons second $^{-1}$ cm $^{-2}$ sr $^{-1}$) from ICC180 measured using the IVIS® Kinetic camera system (Perkin Elmer). The images show changes in peak bioluminescence over time with variations in colour representing light intensity at a given location and superimposed over a grey-scale reference image (A). Red represents the most intense light emission, whereas blue corresponds to the weakest signal. Bioluminescence from the abdominal region of individual mice was quantified (as photons second $^{-1}$) using the region of interest tool in the Living Image software program and used to calculate area under curve (AUC) values for each individual animal over the course of the infection (B). Dotted line represents background. Experiments were performed on two separate occasions. Three representative animals are shown; no light was detected from animals infected with ICC169 alone, while lower levels of light were detected from animals infected with a mix of ICC169 and ICC180.

3.3.4 Discussion

Bioluminescently-labelled bacteria have gained popularity as a powerful tool for investigating microbial pathogenicity *in vivo*, and for preclinical drug and vaccine development^{180–183}. Individual infected and/or treated animals can be followed over time, in contrast to the large numbers of animals that are euthanised at specific time points of interest for quantifying bacterial loads using labour-intensive plate count methods. Most widely used is the *lux* operon of the terrestrial bacterium *P. luminescens*, which encodes for the luciferase enzyme which catalyses the bioluminescence reaction, and for a multi-enzyme complex responsible for regenerating the required substrate. As FMNH₂ is also required for light production, it is generally hypothesised that light production is likely to impose a metabolic burden on tagged bacteria.

The impact of expression of the *lux* operon has been reported for a number of microbial species. Sanz and colleagues created strains of *Bacillus anthracis* that emit light during germination, by introducing plasmids with *lux* operon expression driven by the *sspB* promoter¹⁸⁴. The authors noted that the bioluminescent strains were less efficient at germinating, resulting in an increase in the dose required to cause a lethal infection in mice inoculated by either the subcutaneous or intranasal route. Despite the reduced virulence, bioluminescent *B. anthracis* was still capable of successfully mounting an infection, and the use of biophotonic imaging revealed new infection niches which would have been difficult to accurately measure using traditional plating methods. Similarly, a clinical M75 isolate of *Streptococcus pyogenes* with the *lux* operon chromosomally inserted at the *spy0535* gene was found to have significantly attenuated maximal growth *in vitro*, as well as reduced survival in an intranasal mouse model¹⁸⁵. The bioluminescent *Listeria monocytogenes* Xen32 strain was shown to cause reduced mortality after oral inoculation of BALB/cJ mice, however subsequent investigation revealed that the chromosomally-located *lux* operon had inserted into the *flaA* gene, disrupting the ability of Xen32 to produce flagella. This suggests that the virulence attenuation observed is likely due to the location of the *lux* operon rather than the metabolic cost of light production¹⁸⁶.

In this study, we compared a bioluminescent-derivative of the mouse enteropathogen *C. rodentium*, strain ICC180, with its non-bioluminescent parent strain ICC169, using the BIOLOG Phenotypic Microarray (PM) system, which tests microbial growth under approximately 2000 different metabolic conditions. Rather surprisingly, our results demonstrated that the expression of bioluminescence in ICC180 is near-neutral in almost every non-toxic environment tested, suggesting that light production is not metabolically costly to *C. rodentium*. This supports the “free lunch hypothesis” proposed by Falls and colleagues, namely that cells have an excess of metabolic power available to them¹⁸⁷. Interestingly, ICC180 grew significantly better than its non-bioluminescent parent strain in the presence of a number of different chemicals, including several antibiotics, supporting previous findings that bacteria have many pleiotropic ways to resist toxins¹⁸⁸. In the case of the artificial electron acceptor

iodonitrotetrazolium violet, we hypothesise that light production may be altering the redox balance of the cell, thus making the dye less toxic.

We also compared the ability of ICC180 and ICC169 to directly compete with one another during infection of their natural host, laboratory mice, as well as larvae of the Greater Wax Moth *G. mellonella* (waxworms). Wax worms are becoming an increasingly popular surrogate host for infectious diseases studies due to legislative requirements in many countries to replace the use of animals in scientific research. Wax worms have a well-developed innate immune system involving a cellular immune response in the form of haemocytes, and a humoral immune response in the form of antimicrobial peptides in the hemolymph¹⁸⁹. Detection of bacterial cell wall components leads to activation of the prophenoloxidase cascade, which is similar to the complement system in mammals¹⁹⁰, and subsequent endocytosis of bacteria by haemocytes. The haemocytes function in a similar way to mammalian neutrophils, and kill bacteria via NADPH oxidase and production of reactive oxygen species¹⁹¹. Again, we observed no fitness costs to constitutive light production by ICC180. Interestingly, we recovered significantly more ICC180 from wax worms infected with both ICC180 and ICC169. Similar to the response to iodonitrotetrazolium violet, an altered redox balance caused by light production could make reactive oxygen species generated by the wax worm immune response, less toxic.

In contrast, our data shows that the non-bioluminescent parent strain ICC169 has a clear competitive advantage over ICC180 during infection of adult C57Bl/6 mice, with the bioluminescent strain shed from infected animals at consistently lower numbers. Surprisingly though, this competitive advantage is not sufficient for the parent strain to entirely outcompete and displace its bioluminescent derivative, which remains present in the gastrointestinal tract until clearance of both strains by the immune system. This suggests that there are sufficient niches within the gastrointestinal tract for the two strains to coexist.

It is important to note that in addition to light production, ICC180 differs from its non-bioluminescent parent strain ICC169 by lacking a putative site-specific DNA recombinase, disrupted by insertion of the *lux* operon. *C. rodentium* ICC180 was constructed by random transposon mutagenesis of ICC169 with a mini-Tn5 vector containing an unpromoted *lux* operon and kanamycin-resistance gene. Previous characterisation of the site of insertion of the *lux* operon suggested that the transposon had inserted within a homologue of the *xylE* gene. However, whole genome sequencing has revealed that this was incorrect and the *lux* operon has inserted at 5,212,273 bp, disrupting the coding region of the putative site-specific DNA recombinase. Whole genome sequencing also revealed that ICC180 differs from ICC169 by 2 non-synonymous SNPs, a single base pair insertion and a 90 bp deletion. It is unclear if these changes occurred during the process of transposon mutagenesis, and are merely ‘hitch-hikers’, or

after laboratory passage. The single base pair insertion revealed by sequencing is of a G residue at 3,326,092 bp which results in a frameshift mutation within a putative membrane transporter, while the 90 bp deletion is within the deoxyribose operon repressor gene *deoR*. The DeoR protein represses the *deoCABD* operon, which is involved in the catabolism of deoxyribonucleotides. One SNP is the substitution of an aspartic acid (D) for a glycine (G) at residue 471 of CtsIV, a Type 6 secretion system protein involved in ATP binding. The other SNP is the substitution of a glutamic acid (E) for a glycine (G) at residue 89 of the formate acetyltransferase 2 gene *pflD*, which is involved in carbon utilisation under anaerobic conditions. Modelling suggests that once mutated, residue 89 will be unable to make several key contacts, suggesting the function of PflD will be affected. As we have not introduced these genetic differences into the non-bioluminescent parent strain, we cannot be certain whether the fitness costs we observed are a result of any single or combination of these differences, or expression of the lux operon. In addition, at 54 kb the largest *C. rodentium* plasmid pCROD1 is dramatically altered in ICC180, missing 41 out of 60 of genes. This is in contrast to previous results which indicated that pCROD1 is entirely absent in ICC180¹³⁴. We do not anticipate that the loss of a large part of this plasmid will have any significant impact however, as it has been shown that pCROD1 is frequently lost in *C. rodentium*, and that strains lacking pCROD1 do not show any attenuation of virulence in a C57BL/6 mouse model¹³⁴.

The bioluminescent *C. rodentium* strain ICC180 has a clear disadvantage when directly competed with its parent stain in mice. However, the fact that it reaches similar numbers, and causes similar pathology^{93, 94}, during single infections is reassuring. Our phenotypic microarray data suggests that constitutive light expression is surprisingly neutral in *C. rodentium* and supports the view that bioluminescent versions of microbes can be used as a substitute for their non-bioluminescent parents, at least in theory. In reality, the actual fitness costs will likely depend on the host bacterial species, whether the lux operon is located on a multi-copy plasmid or integrated into the chromosome (and if chromosomal, the site of insertion of the operon), and the levels of expression of the *lux* genes.

3.3.5 Supplementary Tables

Supplementary Table 1. BIOLOG Phenotypic Microarray assays.

Substrates and positions within the 96 well plates of the Biolog PM plates used in this study

PM1 – Carbon Sources	A1, Negative Control; A2, L-Arabinose; A3, N-Acetyl-D Glucosamine; A4, D-Saccharic Acid; A5, Succinic Acid; A6, D-Galactose; A7, L-Aspartic Acid; A8, L-Proline; A9, D-Alanine; A10, D-Trehalose; A11, D-Mannose; A12, Dulcitol; B1, D-Serine; B2, D-Sorbitol; B3, Glycerol; B4, L-Fucose; B5, D-Glucuronic Acid; B6, D-Gluconic Acid; B7, D,L- α -Glycerol Phosphate; B8, D-Xylose; B9, L-Lactic Acid; B10, Formic Acid; B11, D-Mannitol; B12, L-Glutamic Acid; C1, D-Glucose-6-Phosphate; C2, D-Galactonic Acid- γ -Lactone; C3, D,L-Malic Acid; C4, D-Ribose; C5, Tween 20; C6, L-Rhamnose; C7, D-Fructose; C8, Acetic Acid; C9, α -D-Glucose; C10, Maltose; C11, D-Melibiose; C12, Thymidine; D1, L-Asparagine; D2, D-Aspartic Acid; D3, D-Glucosaminic Acid; D4, 1,2- Propanediol; D5, Tween 40; D6, α -Keto-Glutaric Acid; D7, α -Keto-Butyric Acid; D8, α -Methyl-D Galactoside; D9, α -D-Lactose; D10, Lactulose; D11, Sucrose; D12, Uridine; E1, L-Glutamine; E2, m-Tartaric Acid; E3, D-Glucose-1-Phosphate; E4, D-Fructose-6-Phosphate; E5, Tween 80; E6, α -Hydroxy Glutaric Acid- α -Lactone; E7, α -Hydroxy Butyric Acid; E8, α -Methyl-DGlucoside; E9, Adonitol; E10, Maltotriose; E11, 2-Deoxy Adenosine; E12, Adenosine; F1, Glycyl-L-Aspartic Acid; F2, Citric Acid; F3, m-Inositol; F4, D-Threonine; F5, Fumaric Acid; F6, Bromo Succinic Acid; F7, Propionic Acid; F8, Mucic Acid; F9, Glycolic Acid; F10, Glyoxylic Acid; F11, D-Cellobiose; F12, Inosine; G1, Glycyl-L-Glutamic Acid; G2, Tricarballylic Acid; G3, L-Serine; G4, L-Threonine; G5,L-Alanine; G6, L-Alanyl-Glycine; G7, Acetoacetic Acid; G8, N-Acetyl- β -D-Mannosamine; G9, Mono Methyl Succinate; G10, Methyl Pyruvate; G11, D-Malic Acid; G12, L-Malic Acid; H1, Glycyl-L-Proline; H2, p-Hydroxy Phenyl Acetic Acid; H3, m-Hydroxy Phenyl Acetic Acid; H4, Tyramine; H5, D-Psicose; H6, L-Lyxose; H7, Glucuronamide; H8, Pyruvic Acid; H9, L-Galactonic Acid- γ -Lactone; H10, D-Galacturonic Acid; H11, Phenylethylamine; H12, 2-Amino Ethanol.
PM2 – Carbon Sources	A1, Negative Control; A2, Chondroitin Sulfate C; A3, α -Cyclodextrin; A4, β -Cyclodextrin; A5, γ -Cyclodextrin; A6, Dextrin; A7, Gelatin; A8, Glycogen; A9, Inulin; A10, Laminarin; A11, Mannan; A12, Pectin; B1, N-Acetyl-D Galactosamine; B2, N-Acetyl Neuraminic Acid; B3, β -D-Allose; B4, Amygdalin; B5, D-Arabinose; B6, D-Arabitol; B7, L-Arabinol; B8, Arbutin; B9, 2-Deoxy-D Ribose; B10, i-Erythritol; B11, D-Fucose; B12,3-0- β -D-Galactopyranosyl-D Arabinose; C1, Gentiobiose; C2, L-Glucose; C3, Lactitol; C4, D-Melezitose; C5, Maltitol; C6, α -Methyl-D Glucoside; C7, β -Methyl-D Galactoside; C8, 3-Methyl Glucose; C9, β -Methyl-D Glucuronic Acid; C10, α -Methyl-D Mannoside; C11, β -Methyl-D Xyloside; C12, Palatinose; D1, D-Raffinose; D2, Salicin; D3, Sedoheptulosan; D4, L-Sorbose; D5, Stachyose; D6, D-Tagatose; D7, Turanose; D8, Xylitol; D9, N-Acetyl-D Glucosaminitol; D10, γ -Amino Butyric Acid; D11, δ -Amino Valeric Acid; D12, Butyric Acid; E1, Capric Acid; E2, Caproic Acid; E3, Citraconic Acid; E4, Citramalic Acid; E5, D-Glucosamine; E6, 2-Hydroxy Benzoic Acid; E7, 4-Hydroxy Benzoic Acid; E8, β -Hydroxy Butyric Acid; E9, δ -Hydroxy Butyric Acid; E10, a-Keto-Valeric Acid; E11, Itaconic Acid; E12, 5-Keto-D Gluconic Acid; F1, D-Lactic Acid Methyl Ester; F2, Malonic Acid; F3, Melibionic Acid; F4, Oxalic Acid; F5, Oxalomalic Acid; F6, Quinic Acid; F7, D-Ribono-1,4-Lactone; F8, Sebacic Acid; F9, Sorbic Acid; F10, Succinamic Acid; F11, D-Tartaric Acid; F12, L-Tartaric Acid; G1, Acetamide; G2, L-Alaninamide; G3, N-Acetyl-L Glutamic Acid; G4, L-Arginine; G5, Glycine; G6, L-Histidine; G7, L-Homoserine; G8, Hydroxy-L Proline; G9, L-Isoleucine; G10, L-Leucine; G11, L-Lysine; G12, L-Methionine; H1, L-Ornithine; H2, L-Phenylalanine; H3, L-Pyroglutamic Acid; H4, L-Valine; H5, D,L-Carnitine; H6, Sec-Butylamine; H7, D,L-Octopamine; H8, Putrescine; H9, Dihydroxy Acetone; H10, 2,3-Butanediol; H11, 2,3-Butanone; H12, 3-Hydroxy 2-Butanone.
PM3 – Nitrogen Sources	A1, Negative Control; A2, Ammonia; A3, Nitrite; A4, Nitrate; A5, Urea; A6, Biuret; A7, L-Alanine; A8, L-Arginine; A9, L-Asparagine; A10, L-Aspartic Acid; A11, L-Cysteine; A12, L-Glutamic Acid; B1, L-Glutamine; B2, Glycine; B3, L-Histidine; B4, L-Isoleucine; B5, L-Leucine; B6, L-Lysine; B7, L-Methionine; B8, L-Phenylalanine; B9, L-Proline; B10, L-Serine; B11, L-Threonine; B12, L-Tryptophan; C1, L-Tyrosine; C2, L-Valine; C3, D-Alanine; C4, D-Asparagine; C5, D-Aspartic Acid; C6, D-Glutamic Acid; C7, D-Lysine; C8, D-Serine; C9, D-Valine; C10, L-Citrulline; C11, L-Homoserine; C12, L-Ornithine; D1, N-Acetyl-L Glutamic Acid; D2, N-Phthaloyl-L Glutamic Acid; D3, L-Pyroglutamic Acid; D4, Hydroxylamine; D5, Methylamine; D6, N-Amylamine; D7, N-Butylamine; D8, Ethylamine; D9, Ethanolamine; D10, Ethylenediamine; D11, Putrescine; D12, Agmatine; E1, Histamine; E2, β -Phenylethylamine; E3, Tyramine; E4, Acetamide; E5, Formamide; E6, Glucuronamide; E7, D,L-Lactamide; E8, D-Glucosamine; E9, D-Galactosamine; E10, D-Mannosamine; E11, N-Acetyl-D Glucosamine; E12, N-Acetyl-D Galactosamine; F1, N-Acetyl-D Mannosamine; F2, Adenine; F3, Adenosine; F4, Cytidine; F5, Cytosine; F6, Guanine; F7, Guanosine; F8, Thymine; F9, Thymidine; F10, Uracil; F11, Uridine; F12, Inosine; G1, Xanthine; G2, Xanthosine; G3, Uric Acid; G4, Alloxan; G5, Allantoin; G6, Parabanic Acid; G7, D,L- α -Amino-N Butyric Acid; G8, β -Amino-N Butyric Acid; G9, ε -

	Amino-N Caproic Acid; G10, D,L- α -Amino Caprylic Acid; G11, δ -Amino-N Valeric Acid; G12, α -Amino-N Valeric Acid; H1, Ala-Asp; H2, Ala-Gln; H3, Ala-Glu; H4, Ala-Gly; H5, Ala-His; H6, Ala-Leu; H7, Ala-Thr; H8, Gly-Asn; H9, Gly-Gln; H10, Gly-Glu; H11, Gly-Met; H12, Met-Ala.
PM4 – Phosphorus and Sulfur Sources	A1, Negative Control; A2, Phosphate; A3, Pyrophosphate; A4, Trimetaphosphate; A5, Tripolyphosphate; A6, Triethyl Phosphate; A7, Hypophosphate; A8, Adenosine-2'-monophosphate; A9, Adenosine-3'-monophosphate; A10, Adenosine-5'-monophosphate; A11, Adenosine-2',3'-cyclic monophosphate; A12, Adenosine-3',5'-cyclic monophosphate; B1, Thiophosphate; B2, Dithiophosphate; B3, D,L- α -Glycerol Phosphate; B4, β -Glycerol Phosphate; B5, Carbamyl Phosphate; B6, D-2-Phospho Glyceric Acid; B7, D-3-Phospho Glyceric Acid; B8, Guanosine-2'-monophosphate; B9, Guanosine-3'-monophosphate; B10, Guanosine-5'-monophosphate; B11, Guanosine-2',3'-cyclic monophosphate; B12, Guanosine-3',5'-cyclic monophosphate; C1, Phosphoenol Pyruvate; C2, Phospho Glycolic Acid; C3, D-Glucose-1-Phosphate; C4, D-Glucose-6-Phosphate; C5, 2-Deoxy-D Glucose 6-Phosphate; C6, D-Glucosamine-6-Phosphate; C7, 6-Phospho Gluconic Acid; C8, Cytidine-2'-monophosphate; C9, Cytidine-3'-monophosphate; C10, Cytidine-5'-monophosphate; C11, Cytidine-2',3'-cyclic monophosphate; C12, Cytidine-3',5'-cyclic monophosphate; D1, D-Mannose-1-Phosphate; D2, D-Mannose-6-Phosphate; D3, Cysteamine-S Phosphate; D4, Phospho-L Arginine; D5, O-Phospho-D Serine; D6, O-Phospho-L Serine; D7, O-Phospho-L Threonine; D8, Uridine-2'-monophosphate; D9, Uridine-3'-monophosphate; D10, Uridine-5'-monophosphate; D11, Uridine-2',3'-cyclic monophosphate; D12, Uridine-3',5'-cyclic monophosphate; E1, O-Phospho-D Tyrosine; E2, O-Phospho-L Tyrosine; E3, Phosphocreatine; E4, Phosphoryl Choline; E5, O-Phosphoryl Ethanolamine; E6, Phosphono Acetic Acid; E7, 2-Aminoethyl Phosphonic Acid; E8, Methylene Diphosphonic Acid; E9, Thymidine-3'-monophosphate; E10, Thymidine-5'-monophosphate; E11, Inositol Hexaphosphate; E12, Thymidine 3',5'-cyclic monophosphate; F1, Negative Control; F2, Sulfate; F3, Thiosulfate; F4, Tetrathionate; F5, Thiophosphate; F6, Dithiophosphate; F7, L-Cysteine; F8, D-Cysteine; F9, L-Cysteinyl Glycine; F10, L-Cysteic Acid; F11, Cysteamine; F12, L-Cysteine Sulfinic Acid; G1, N-Acetyl-L Cysteine; G2, S-Methyl-L Cysteine; G3, Cystathionine; G4, Lanthionine; G5, Glutathione; G6, D,L-Ethionine; G7, L-Methionine; G8, D-Methionine; G9, Glycyl-L Methionine; G10, N-Acetyl-D,L Methionine; G11, L- Methionine Sulfoxide; G12, L-Methionine Sulfone; H1, L-Djenkolic Acid; H2, Thiourea; H3, 1-Thio- β -D Glucose; H4, D,L-Lipoamide; H5, Taurocholic Acid; H6, Taurine; H7, Hypotaurine; H8, p-Amino Benzene Sulfonic Acid; H9, Butane Sulfonic Acid; H10, 2-Hydroxyethane Sulfonic Acid; H11, Methane Sulfonic Acid; H12, Tetramethylene Sulfone.
PM5 – Nutrient Supplements	A1, Negative Control; A2, Positive Control; A3, L-Alanine; A4, L-Arginine; A5, L-Asparagine; A6, L-Aspartic Acid; A7, L-Cysteine, A8, L-Glutamic Acid; A9, Adenosine-3',5'-cyclic monophosphate; A10, Adenine; A11, Adenosine; A12, 2'-Deoxy Adenosine; B1, L-Glutamine; B2, Glycine; B3, L-Histidine; B4, L-Isoleucine; B5, L-Leucine; B6, L-Lysine; B7, L-Methionine; B8, L-Phenylalanine; B9, Guanosine-3',5'-cyclic monophosphate; B10, Guanine; B11, Guanosine; B12, 2'-Deoxy Guanosine; C1, L-Proline; C2, L-Serine; C3, L-Threonine; C4, L-Tryptophan; C5, L-Tyrosine; C6, L-Valine; C7, L-isoleucine + L-Valine; C8, trans-4-Hydroxy L-Proline; C9, (5) 4-AminoImidazole-4(5)-Carboxamide; C10, Hypoxanthine; C11, Inosine; C12, 2'-Deoxy Inosine; D1, L-Ornithine; D2, L-Citrulline; D3, Chorismic Acid; D4, (-)Shikimic Acid; D5, L-Homoserine Lactone; D6, D-Alanine; D7, D-Aspartic Acid; D8, D-Glutamic Acid; D9, D,L- α , ϵ -Diaminopimelic Acid; D10, Cytosine; D11, Cytidine; D12, 2'-Deoxy Cytidine; E1, Putrescine; E2, Spermidine; E3, Spermine; E4, Pyridoxine; E5, Pyridoxal; E6, Pyridoxamine; E7, β -Alanine; E8, D-Pantothenic Acid; E9, Orotic Acid; E10, Uracil; E11, Uridine; E12, 2'-Deoxy Uridine; F1, Quinolinic Acid; F2, Nicotinic Acid; F3, Nicotinamide; F4, β -Nicotinamide Adenine Dinucleotide; F5, δ -Amino Levulinic Acid; F6, Hematin F7, Deferoxamine; Mesylate; F8, D-(+)-Glucose; F9, N-Acetyl D-Glucosamine; F10, Thymine; F11, Glutathione (reduced form); F12, Thymidine; G1, Oxaloacetic Acid; G2, D-Biotin; G3, CyanoCobalamine; G4, p-Amino Benzoic Acid; G5, Folic Acid; G6, Inosine +Thiamine; G7, Thiamine; G8, Thiamine Pyrophosphate; G9, Riboflavin; G10, Pyrrolo-Quinoline Quinone; G11, Menadione; G12, m-Inositol; H1, Butyric Acid; H2, D,L- α -Hydroxy Butyric Acid; H3, α -Keto Butyric Acid; H4, Caprylic Acid; H5, D,L- α -Lipoic Acid (oxidized form); H6, D,L-Mevalonic Acid; H7, D,L-Carnitine; H8, Choline; H9, Tween 20; H10, Tween 40; H11, Tween 60; H12, Tween 80.
PM6 – Peptide Nitrogen sources	A1, Negative Control; A2, Positive Control: L-Glutamine; A3, Ala-Ala; A4, Ala-Arg; A5, Ala-Asn; A6, Ala-Glu; A7, Ala-Gly; A8, Ala-His; A9, Ala-Leu; A10, Ala-Lys; A11, Ala-Phe; A12, Ala-Pro; B1, Ala-Ser; B2, Ala-Thr; B3, Ala-Trp; B4, Ala-Tyr; B5, Arg-Ala; B6, Arg-Arg; B7, Arg-Asp; B8, Arg-Gln; B9, ; rg-Glu; B10, Arg-Ile; B11, Arg-Leu; B12, Arg-Lys; C1, Arg-Met; C2, Arg-Phe; C3, Arg-Ser; C4, Arg-; rp; C5, Arg-Tyr; C6, Arg-Val; C7, Asn-Glu; C8, Asn-Val; C9, Asp-Asp; C10, Asp-Glu; C11, Asp-Leu; C12, Asp-Lys; D1, Asp-Phe; D2, Asp-Trp; D3, Asp-Val; D4, Cys-Gly; D5, Gln-Gln ;D6, Gln-Gly; D7, Glu-Asp; D8, Glu-Glu; D9, Glu-Gly; D10, Glu-Ser; D11, Glu-Trp; D12, Glu-Tyr; E1, Glu-Val; E2, Gly-Ala; E3, Gly-Arg; E4, Gly-Cys; E5, Gly-Gly; E6, Gly-His; E7, Gly-Leu; E8, Gly-Lys; E9, Gly-Met; E10, Gly-Phe; E11, Gly-Pro; E12, Gly-Ser; F1, Gly-Thr; F2, Gly-Trp; F3, Gly-Tyr; F4, Gly-Val; F5, His-Asp; F6, His-Gly; F7, His-Leu; F8, His-Lys; F9, His-Met; F10, His-Pro; F11 ,His-Ser; F12, His-Trp; G1, His-Tyr; G2, His-Val; G3, Ile-Ala; G4, Ile-Arg; G5, Ile-Gln; G6, Ile-Gly; G7, Ile-His; G8, Ile-Ile; G9, Ile-Met; G10, Ile-Phe; G11, Ile-Pro; G12, Ile-Ser; H1, lle-Trp; H2, lle-Tyr; H3, lle-Val; H4, Leu-Ala; H5, Leu-Arg; H6, Leu-Asp; H7, Leu-Glu; H8, Leu-Gly; H9, Leu-Ile; H10, Leu-Leu; H11, Leu-Met; H12, Leu-Phe.

PM7 – Peptide Nitrogen sources	A1, Negative Control; A2, Positive Control: L-Glutamine; A3, Leu-Ser; A4, Leu-Trp; A5, Leu-Val; A6, Lys-Ala; A7, Lys-Arg; A8, Lys-Glu; A9, Lys-Ile; A10, Lys-Leu; A11, Lys-Lys; A12, Lys-Phe; B1, Lys-Pro; B2, Lys-Ser; B3, Lys-Thr; B4, Lys-Trp; B5 ,Lys-Tyr; B6, Lys-Val; B7, Met-Arg; B8, Met-Asp; B9, Met-Gln; B10, Met-Glu; B11, Met-Gly; B12, Met-His; C1, Met-Ile; C2, Met-Leu; C3, Met-Lys; C4, Met-Met; C5, Met-Phe; C6, Met-Pro; C7, Met-Trp; C8, Met-Val; C9, Phe-Ala; C10, Phe-Gly; C11, Phe-Ile; C12, Phe-Phe; D1, Phe-Pro; D2, Phe-Ser; D3, Phe-Trp; D4, Pro-Ala; D5, Pro-Asp; D6, Pro-Gln; D7, Pro-Gly; D8, Pro-Hyp; D9, Pro-Leu; D10, Pro-Phe; D11, Pro-Pro; D12, Pro-Tyr; E1, Ser-Ala; E2, Ser-Gly; E3, Ser-His; E4, Ser-Leu; E5, Ser-Met; E6, Ser-Phe; E7, Ser-Pro; E8, Ser-Ser; E9, Ser-Tyr; E10, Ser-Val; E11, Thr-Ala; E12, Thr-Arg; F1, Thr-Glu; F2, Thr-Gly; F3, Thr-Leu; F4, Thr-Met; F5, Thr-Pro; F6, Trp-Ala; F7, Trp-Arg; F8, Trp-Asp; F9, Trp-Glu; F10, Trp-Gly; F11, Trp-Leu; F12, Trp-Lys; G1, Trp-Phe; G2, Trp-Ser; G3, Trp-Trp; G4, Trp-Tyr; G5, Tyr-Ala; G6, Tyr-Gln; G7, Tyr-Glu; G8, Tyr-Gly G9, Tyr-His; G10, Tyr-Leu; G11, Tyr-Lys; G12, Tyr-Phe; H1, Tyr-Trp; H2, Tyr-Tyr; H3, Val-Arg; H4, Val-Asn; H5, Val-Asp; H6, Val-Gly; H7, Val-His; H8, Val-Ile; H9, Val-Leu; H10, Val-Tyr; H11, Val-Val; H12, Y-Glu-Gly.
PM8 – Peptide Nitrogen sources	A1, Negative Control; A2, Positive Control: L-Glutamine; A3, Ala-Asp; A4, Ala-Gln; A5, Ala-Ile; A6, Ala-Met; A7, Ala-Val; A8, Asp-Ala; A9, Asp-Gln; A10, Asp-Gly; A11, Glu-Ala; A12, Gly-Asn; B1, Gly-Asp; B2, Gly-Ile; B3, His-Ala; B4, His-Glu; B5, His-His; B6, Ile-Asn; B7, Ile-Leu; B8, Leu-Asn; B9, Leu-His; B10, Leu-Pro; B11, Leu-Tyr; B12, Lys-Asp; C1, Lys-Gly; C2, Lys-Met; C3, Met-Thr; C4, Met-Tyr; C5, Phe-Asp; C6, Phe-Glu; C7, Gln-Glu; C8, Phe-Met; C9, Phe-Tyr; C10, Phe-Val; C11, Pro-Arg; C12, Pro-Asn; D1, Pro-Glu; D2, Pro-Ile; D3, Pro-Lys; D4, Pro-Ser; D5, Pro-Trp; D6, Pro-Val; D7, Ser-Asn; D8, Ser-Asp; D9, Ser-Gln; D10, Ser-Glu; D11, Thr-Asp; D12, Thr-Gln; E1, Thr-Phe; E2, Thr-Ser; E3, Trp-Val; E4, Tyr-Ile; E5, Tyr-Val; E6, Val-Ala; E7, Val-Gln; E8, Val-Glu; E9, Val-Lys; E10, Val-Met; E11, Val-Phe; E12, Val-Pro; F1, Val-Ser; F2, β-Ala-Ala; F3, β-Ala-Gly; F4, β-Ala-His; F5, Met-β-Ala; F6, β-Ala-Phe; F7, D-Ala-D-Ala; F8, D-Ala-Gly; F9, D-Ala-Leu; F10, D-Leu-D-Leu; F11, D-Leu-Gly; F12, D-Leu-Tyr; G1, Y-Glu-Gly; G2, Y-D-Glu-Gly; G3, Gly-D-Ala; G4, Gly-D-Asp; G5, Gly-D-Ser; G6, Gly-D-Thr; G7, Gly-D-Val; G8, Leu-β-Ala; G9, Leu-D-Leu; G10, Phe-β-Ala; G11, Ala-Ala-Ala; G12, D-Ala-Gly-Gly; H1, Gly-Gly-Ala; H2, Gly-Gly-D-Leu; H3, Gly-Gly-Gly; H4, Gly-Gly-Ile; H5, Gly-Gly-Leu; H6, Gly-Gly-Phe; H7, Val-Tyr-Val; H8, Gly-Phe-Phe; H9, Leu-Gly-Gly; H10, Leu-Leu-Leu; H11, Phe-Gly-Gly; H12, Tyr-Gly-Gly.
PM9 – Osmolytes	A1, NaCl 1%; A2, NaCl 2%; A3, NaCl 3%; A4, NaCl 4%; A5, NaCl 5%; A6, NaCl 5.5%; A7, NaCl 6%; A8, NaCl 6.5%; A9, NaCl 7%; A10, NaCl 8%; A11, NaCl 9%; A12, NaCl 10%; B1, NaCl 6%; B2, NaCl 6% +Betaine; B3, NaCl 6% +N-N Dimethyl Glycine; B4, NaCl 6% + Sarcosine; B5, NaCl 6% + Dimethyl sulphonyl propionate; B6, NaCl 6% + MOPS; B7, NaCl 6% + Ectoine; B8, NaCl 6% + Choline; B9, NaCl 6% + Phosphoryl Choline; B10, NaCl 6% + Creatine; B11, NaCl 6% + Creatinine; B12, NaCl 6% + L- Carnitine; C1, NaCl 6% + KCl; C2, NaCl 6% + L-Proline; C3, NaCl 6% + N-Acetyl L-Glutamine; C4, NaCl 6% + β-Glutamic Acid; C5 ,NaCl 6% + γ-Amino -N Butyric Acid; C6, NaCl 6% + Glutathione; C7, NaCl 6% + Glycerol; C8, NaCl 6% +Trehalose; C9, NaCl 6% + TrimethylamineN-oxide; C10, NaCl 6% + Trimethylamine; C11, NaCl 6% + Octopine; C12, NaCl 6% + Trigonelline; D1, Potassium chloride 3%; D2, Potassium chloride 4%; D3, Potassium chloride 5%; D4, Potassium chloride 6%; D5, Sodium sulphate 2%; D6, Sodium sulphate 3%; D7, Sodium sulphate 4%; D8, Sodium sulphate 5%; D9, Ethylene glycol 5%; D10, Ethylene glycol 10%; D11, Ethylene glycol 15%; D12, Ethylene glycol 20%; E1, Sodium formate 1%; E2, Sodium formate 2%; E3, Sodium formate 3%; E4, Sodium formate 4%; E5, Sodium formate 5%; E6, Sodium formate 6%; E7, Urea 2%; E8, Urea 3%; E9, Urea 4%; E10, Urea 5%; E11, Urea 6%; E12, Urea 7%; F1, Sodium Lactate 1%; F2, Sodium Lactate 2%; F3, Sodium Lactate 3%; F4, Sodium Lactate 4%; F5, Sodium Lactate 5%; F6, Sodium Lactate 6%; F7, Sodium Lactate 7%; F8, Sodium Lactate 8%; F9, Sodium Lactate 9%; F10, Sodium Lactate 10%; F11, Sodium Lactate 11%; F12, Sodium Lactate 12%; G1, Sodium Phosphate pH 7 20mM; G2, Sodium Phosphate pH 7 50mM; G3, Sodium Phosphate pH 7 100mM; G4, Sodium Phosphate pH 7 200mM; G5, Sodium Benzoate pH 5.2 20mM; G6, Sodium Benzoate pH 5.2 50mM; G7, Sodium Benzoate pH5.2 100mM; G8, Sodium Benzoate pH 5.2 200mM; G9, Ammonium sulfate pH8 10mM; G10, Ammonium sulfate pH 8 20mM; G11, Ammonium sulfate pH 8 50mM; G12, Ammonium sulfate pH8 100mM; H1, Sodium Nitrate 10mM; H2, Sodium Nitrate 20mM; H3, Sodium Nitrate 40mM; H4, Sodium Nitrate 60mM; H5, Sodium Nitrate 80mM; H6, Sodium Nitrate 100mM; H7, Sodium Nitrite 10mM; H8, Sodium Nitrite 20mM; H9, Sodium Nitrite 40mM; H10, Sodium Nitrite 60mM; H11, Sodium Nitrite 80mM; H12, Sodium Nitrite 100mM.
PM10 – pH	A1, pH 3.5; A2, pH 4; A3, pH 4.5; A4, pH 5; A5, pH 5.5; A6, pH 6; A7, pH 7; A8, pH 8; A9, pH 8.5; A10, pH 9; A11, pH 9.5; A12, pH 10; B1, pH 4.5; B2, pH 4.5 + L-Alanine; B3, pH 4.5 + L-Arginine; B4, pH 4.5 + L-Asparagine; B5, pH 4.5 + L-Aspartic Acid; B6, pH 4.5 + L-Glutamic Acid; B7, pH 4.5 + L-Glutamine; B8, pH 4.5 + Glycine; B9, pH 4.5 + L-Histidine; B10, pH 4.5 + L-Isoleucine; B11, pH 4.5 + L-Leucine; B12, pH 4.5 + L-Lysine; C1, pH 4.5 + L-Methionine; C2, pH 4.5 + L-Phenylalanine; C3, pH 4.5 + L-Proline; C4, pH 4.5 + L-Serine; C5, pH 4.5 + L-Threonine; C6, pH 4.5 + L-Tryptophan; C7, pH 4.5 + L-Citrulline; C8, pH 4.5 + L-Valine; C9, pH 4.5 + HydroxyL-Proline; C10, pH 4.5 + L-Ornithine; C11, pH 4.5 + L-Homoarginine; C12, pH 4.5 + L-Homoserine; D-1, pH 4.5 + Anthranilic Acid; D2, pH 4.5 + L-Norleucine; D3, pH 4.5 + L-Norvaline; D4, pH 4.5 + α- Amino-N Butyric Acid; D5, pH 4.5 + p-

	Amino Benzoic Acid; D6, pH 4.5 + L-Cysteic Acid; D7, pH 4.5 + D-Lysine; D8, pH 4.5 + 5-Hydroxy Lysine; D9, pH 4.5 + 5-Hydroxy Tryptophan; D10, pH 4.5 + D,L-Diamino pimelic Acid; D11, pH 4.5 + Trimethylamine N-oxide; D12, pH 4.5 + Urea; E1, pH 9.5; E2, pH 9.5 + L-Alanine; E3, pH 9.5 + L-Arginine; E4, pH 9.5 + L-Asparagine; E5, pH 9.5 + L-Aspartic Acid; E6, pH 9.5 + L-Glutamic Acid; E7, pH 9.5 + L-Glutamine; E8, pH 9.5 + Glycine; E9, pH 9.5 + L-Histidine; E10, pH 9.5 + L-Isoleucine; E11, pH 9.5 + L-Leucine; E12, pH 9.5 + L-Lysine; F1, pH 9.5 + L-Methionine; F2, pH 9.5 + L-Phenylalanine; F3, pH 9.5 + L-Proline; F4, pH 9.5 + L-Serine; F5, pH 9.5 + L-Threonine; F6, pH 9.5 + L-Tryptophan; F7, pH 9.5 + L-Tyrosine; F8, pH 9.5 + L-Valine; F9, pH 9.5 + Hydroxy L-Proline; F10, pH 9.5 + L-Ornithine; F11, pH 9.5 + L-Homoarginine; F12, pH 9.5 + L-Homoserine; G1, pH 9.5 + Anthranilic acid; G2, pH 9.5 + L-Norleucine; G3, pH 9.5 + L-Norvaline; G4, pH 9.5 + Agmatine; G5, pH 9.5 + Cadaverine; G6, pH 9.5 + Putrescine; G7, pH 9.5 + Histamine; G8, pH 9.5 + Phenylethylamine; G9, pH 9.5 + Tyramine; G10, pH 9.5 + Creatine; G11, pH 9.5 + Trimethylamine N-oxide; G12, pH 9.5 + Urea; H1, X-Caprylate; H2, X- α -DGlucoside; H3, X- β -DGlcusoside; H4, X- α -DGalactoside; H5, X- β -DGalactoside; H6, X- α - DGlcuronide; H7, X- β - DGlcuronide; H8, X- β -DGlcosaminide; H9, X- β -DGlcatosaminide; H10, X- α -DMannoside; H11, X-PO4; H12, X-SO4.
PM11C – chemical	A1, Amikacin (1); A2, Amikacin (2); A3, Amikacin (3); A4, Amikacin (4); A5, Chlortetracycline (1); A6, Chlortetracycline (2) ;A7, Chlortetracycline (3); A8, Chlortetracycline (4); A9, Lincomycin (1); A10, Lincomycin (2); A11, Lincomycin (3); A12, Lincomycin (4); B1, Amoxicillin (1); B2, Amoxicillin (2); B3, Amoxicillin (3); B4, Amoxicillin (4); B5, Cloxacillin (1); B6, Cloxacillin (2); B7, Cloxacillin (3); B8, Cloxacillin (4); B9, Lomefloxacin (1); B10, Lomefloxacin (2); B11, Lomefloxacin (3); B12, Lomefloxacin (4); C1, Bleomycin (1); C2, Bleomycin (2); C3, Bleomycin (3); C4, Bleomycin (4); C5, Colistin (1); C6, Colistin (2); C7, Colistin (3); C8, Colistin (4); C9, Minocycline (1); C10, Minocycline (2); C11, Minocycline (3); C12, Minocycline (4); D1, Capreomycin (1); D2, Capreomycin (2); D3, Capreomycin (3); D4, Capreomycin (4); D5, Demeclocycline (1); D6, Demeclocycline (2); D7, Demeclocycline (3); D8, Demeclocycline (4); D9, Nafcillin (1); D10, Nafcillin (2); D11, Nafcillin (3); D12, Nafcillin (4); E1, Cefazolin (1); E2, Cefazolin (2); E3, Cefazolin (3); E4, Cefazolin (4); E5, Enoxacin (1); E6, Enoxacin (2); E7, Enoxacin (3); E8, Enoxacin (4); E9, Nalidixic acid (1); E10, Nalidixic acid (2); E11, Nalidixic acid (3); E12, Nalidixic acid (4); F1, Chloramphenicol (1); F2, Chloramphenicol (2); F3, Chloramphenicol (3); F4, Chloramphenicol (4); F5, Erythromycin (1); F6, Erythromycin (2); F7, Erythromycin (3); F8, Erythromycin (4); F9, Neomycin (1); F10, Neomycin (2); F11, Neomycin (3); F12, Neomycin (4); G1, Ceftriaxone (1); G2, Ceftriaxone (2); G3, Ceftriaxone (3); G4, Ceftriaxone (4); G5, Gentamicin (1); G6, Gentamicin (2); G7, Gentamicin (3); G8, Gentamicin (4); G9, Potassium tellurite (1); G10, Potassium tellurite (2); G11, Potassium tellurite (3); G12, Potassium tellurite (4); H1, Cephalothin (1); H2, Cephalothin (2); H3, Cephalothin (3); H4, Cephalothin (4); H5, Kanamycin (1); H6, Kanamycin (2); H7, Kanamycin (3); H8, Kanamycin (4); H9, Ofloxacin (1); H10, Ofloxacin (2); H11, Ofloxacin (3); H12, Ofloxacin (4).
PM12B – chemical	A1, Penicillin G (1); A2, Penicillin G (2); A3, Penicillin G (3); A4, Penicillin G (4); A5, Tetracycline (1); A6, Tetracycline (2); A7, Tetracycline (3); A8, Tetracycline (4); A9, Carbenicillin (1); A10, Carbenicillin (2); A11, Carbenicillin (3); A12, Carbenicillin (4); B1, Oxacillin (1); B2, Oxacillin (2); B3, Oxacillin (3); B4, Oxacillin (4); B5, Penimepicycline (1); B6, Penimepicycline (2); B7, Penimepicycline (3); B8, Penimepicycline (4); B9, Polymyxin B (1); B10, Polymyxin B (2); B11, Polymyxin B (3); B12, Polymyxin B (4); C1, Paromomycin (1); C2, Paromomycin (2); C3, Paromomycin (3); C4, Paromomycin (4); C5, Vancomycin (1); C6, Vancomycin (2); C7, Vancomycin (3); C8, Vancomycin (4); C9, D,L-Serinehydroxamate (1); C10, D,L-Serine hydroxamate (2); C11, D,L-Serine hydroxamate (3); C12, D,L-Serine hydroxamate (4); D1, Sisomicin (1); D2, Sisomicin (2); D3, Sisomicin (3); D4, Sisomicin (4); D5, Sulfamethazine (1); D6, Sulfamethazine (2); D7, Sulfamethazine (3); D8, Sulfamethazine (4); D9, Novobiocin (1); D10, Novobiocin (2); D11, Novobiocin (3); D12, Novobiocin (4); E1, 2,4-Diamino-6,7-diisopropylpteridine (1); E2, 2,4-Diamino-6,7-diisopropylpteridine (2); E3, 2,4-Diamino-6,7-diisopropylpteridine (3); E4, 2,4-Diamino-6,7-diisopropylpteridine (4); E5, Sulfadiazine (1); E6, Sulfadiazine (2); E7, Sulfadiazine (3); E8, Sulfadiazine (4); E9, Benzethoniumchloride (1); E10, Benzethoniumchloride (2); E11, Benzethoniumchloride (3); E12, Benzethoniumchloride (4); F1, Tobramycin (1); F2, Tobramycin (2); F3, Tobramycin (3); F4, Tobramycin (4); F5, Sulfathiazole (1); F6, Sulfathiazole (2); F7, Sulfathiazole (3); F8, Sulfathiazole (4); F9, 5-Fluoroorotic acid (1); F10, 5-Fluoroorotic acid (2); F11, 5-Fluoroorotic acid (3); F12, 5-Fluoroorotic acid (4); G1, Spectinomycin (1); G2, Spectinomycin (2); G3, Spectinomycin (3); G4, Spectinomycin (4); G5, Sulfamethoxazole (1); G6, Sulfamethoxazole (2); G7, Sulfamethoxazole (3); G8, Sulfamethoxazole (4); G9, L-Aspartic- β -hydroxamate (1); G10, L-Aspartic- β -hydroxamate (2); G11, L-Aspartic- β -hydroxamate (3); G12, L-Aspartic- β -hydroxamate (4); H1, Spiramycin (1); H2, Spiramycin (2); H3, Spiramycin (3); H4, Spiramycin (4); H5, Rifampicin (1); H6, Rifampicin (2); H7, Rifampicin (3); H8, Rifampicin (4); H9, Dodecyltrimethyl ammonium bromide (1); H10, Dodecyltrimethyl ammonium bromide (2); H11, Dodecyltrimethyl ammonium bromide (3); H12. Dodecyltrimethyl ammonium bromide (4).
PM13B –	A1, Ampicillin (1); A2, Ampicillin (2); A3, Ampicillin (3); A4, Ampicillin (4); A5, Dequalinium chloride (1); A6, Dequalinium chloride (2); A7, Dequalinium chloride (3); A8, Dequalinium chloride (4); A9, Nickel chloride (1); A10, Nickel chloride (2); A11, Nickel chloride (3); A12, Nickel chloride (4); B1,

chemical	Azlocillin (1); B2, Azlocillin (2); B3, Azlocillin (3); B4, Azlocillin (4); B5, 2, 2'-Dipyridyl (1); B6, 2, 2'-Dipyridyl (2); B7, 2, 2'-Dipyridyl (3); B8, 2, 2'-Dipyridyl (4); B9, Oxolinic acid (1); B10, Oxolinic acid (2); B11, Oxolinic acid (3); B12, Oxolinic acid (4); C1, 6-Mercaptopurine (1); C2, 6-Mercaptopurine (2); C3, -Mercaptopurine (3); C4, 6-Mercaptopurine (4); C5, Doxycycline (1); C6, Doxycycline (2); C7, Doxycycline (3); C8, Doxycycline (4); C9, Potassium chromate (1); C10, Potassium chromate (2); C11, Potassium chromate (3); C12, Potassium chromate (4); D1, Cefuroxime (1); D2, Cefuroxime (2); D3, Cefuroxime (3); D4, Cefuroxime (4); D5, 5-Fluorouracil (1); D6, 5-Fluorouracil (2); D7, 5-Fluorouracil (3); D8, 5-Fluorouracil (4); D9, Rolitetracycline (1); D10, Rolitetracycline (2); D11, Rolitetracycline (3); D12, Rolitetracycline (4); E1, Cytosine-1- β D-arabinofuranoside (1); E2, Cytosine-1- β D-arabinofuranoside (2); E3, Cytosine-1- β D-arabinofuranoside (3); E4, Cytosine-1- β D-arabinofuranoside (4); E5, Geneticin (G418) (1); E6, Geneticin (G418) (2); E7, Geneticin (G418) (3); E8, Geneticin (G418) (4); E9, Ruthenium red (1); E10, Ruthenium red (2); E11, Ruthenium red (3); E12, Ruthenium red (4); F1, Cesium chloride (1); F2, Cesium chloride (2); F3, Cesium chloride (3); F4, Cesium chloride (4); F5, Glycine (1); F6, Glycine (2); F7, Glycine (3); F8, Glycine (4); F9, Thallium (I) acetate (1); F10, Thallium (I) acetate (2); F11, Thallium (I) acetate (3); F12, Thallium (I) acetate (4); G1, Cobalt chloride (1); G2, Cobalt chloride (2); G3, Cobalt chloride (3); G4, Cobalt chloride (4); G5, Manganese chloride (1); G6, Manganese chloride (2); G7, Manganese chloride (3); G8, Manganese chloride (4); G9, Trifluoperazine (1); G10, Trifluoperazine (2); G11, Trifluoperazine (3); G12, Trifluoperazine (4); H1, Cupric chloride (1); H2, Cupric chloride (2); H3, Cupric chloride (3); H4, Cupric chloride (4); H5, Moxalactam (1); H6, Moxalactam (2); H7, Moxalactam (3); H8, Moxalactam (4); H9, Tylosin (1); H10, Tylosin (2); H11, Tylosin (3); H12, Tylosin (4).
PM14A – chemical	A1, Acriflavine (1); A2, Acriflavine (2); A3, Acriflavine (3); A4, Acriflavine (4); A5, Furaltadone (1); A6, Furaltadone (2); A7, Furaltadone (3); A8, Furaltadone (4); A9, Sanguinarine (1); A10, Sanguinarine (2); A11, Sanguinarine (3); A12, Sanguinarine (4); B1, 9-Aminoacridine (1); B2, 9-Aminoacridine (2); B3, 9-Aminoacridine (3); B4, 9-Aminoacridine (4); B5, Fusaric acid (1); B6, Fusaric acid (2); B7, Fusaric acid (3); B8, Fusaric acid (4); B9, Sodium arsenate (1); B10, Sodium arsenate (2); B11, Sodium arsenate (3); B12, Sodium arsenate (4); C1, Boric Acid (1); C2, Boric Acid (2); C3, Boric Acid (3); C4, Boric Acid (4); C5, 1-Hydroxypyridine-2-thione (1); C6, 1-Hydroxypyridine-2-thione (2); C7, 1-Hydroxypyridine-2-thione (3); C8, 1-Hydroxypyridine-2-thione (4); C9, Sodium cyanate (1); C10, Sodium cyanate (2); C11, Sodium cyanate (3); C12, Sodium cyanate (4); D1, Cadmium chloride (1); D2, Cadmium chloride (2); D3, Cadmium chloride (3); D4, Cadmium chloride (4); D5, Iodoacetate (1); D6, Iodoacetate (2); D7, Iodoacetate (3); D8, Iodoacetate (4); D9, Sodium dichromate (1); D10, Sodium dichromate (2); D11, Sodium dichromate (3); D12, Sodium dichromate (4); E1, Cefoxitin (1); E2, Cefoxitin (2); E3, Cefoxitin (3); E4, Cefoxitin (4); E5, Nitrofurantoin (1); E6, Nitrofurantoin (2); E7, Nitrofurantoin (3); E8, Nitrofurantoin (4); E9, Sodium metaborate (1); E10, Sodium metaborate (2); E11, Sodium metaborate (3); E12, Sodium metaborate (4); F1, Chloramphenicol (1); F2, Chloramphenicol (2); F3, Chloramphenicol (3); F4, Chloramphenicol (4); F5, Piperacillin (1); F6, Piperacillin (2); F7, Piperacillin (3); F8, Piperacillin (4); F9, Sodium metavanadate (1); F10, Sodium metavanadate (2); F11, Sodium metavanadate (3); F12, Sodium metavanadate (4); G1, Chelerythrine (1); G2, Chelerythrine (2); G3, Chelerythrine (3); G4, Chelerythrine (4); G5, Carbenicillin (1); G6, Carbenicillin (2); G7, Carbenicillin (3); G8, Carbenicillin (4); G9, Sodium nitrite (1); G10, Sodium nitrite (2); G11, Sodium nitrite (3); G12, Sodium nitrite (4); H1, EGTA (1); H2, EGTA (2); H3, EGTA (3); H4, EGTA (4); H5, Promethazine (1); H6, Promethazine (2); H7, Promethazine (3); H8, Promethazine (4); H9, Sodium orthovanadate (1); H10, Sodium orthovanadate (2); H11, Sodium orthovanadate (3); H12, Sodium orthovanadate (4).
PM15B – chemical	A1, Procaine (1); A2, Procaine (2); A3, Procaine (3); A4, Procaine (4); A5, Guanidine hydrochloride (1); A6, Guanidine hydrochloride (2); A7, Guanidine hydrochloride (3); A8, Guanidine hydrochloride (4); A9, Cefmetazole (1); A10, Cefmetazole (2); A11, Cefmetazole (3); A12, Cefmetazole (4); B1, D-Cycloserine (1); B2, D-Cycloserine (2); B3, D-Cycloserine (3); B4, D-Cycloserine (4); B5, EDTA (1); B6, EDTA (2); B7, EDTA (3); B8, EDTA (4); B9, 5,7-Dichloro- 8-hydroxyquinaldine (1); B10, 5,7-Dichloro- 8-hydroxyquinaldine (2); B11, 5,7-Dichloro- 8-hydroxyquinaldine (3); B12, 5,7-Dichloro- 8-hydroxyquinaldine (4); C1, 5,7-Dichloro-8-hydroxyquinoline (1); C2, 5,7-Dichloro-8-hydroxyquinoline (2); C3, 5,7-Dichloro-8-hydroxyquinoline (3); C4, 5,7-Dichloro-8-hydroxyquinoline (4); C5, Fusidic acid (1); C6, Fusidic acid (2); C7, Fusidic acid (3); C8, Fusidic acid (4); C9, 1,10-Phenanthroline (1); C10, 1,10-Phenanthroline (2); C11, 1,10-Phenanthroline (3); C12, 1,10-Phenanthroline (4); D1, Phleomycin (1); D2, Phleomycin (2); D3, Phleomycin (3); D4, Phleomycin (4); D5, Domiphen bromide (1); D6, Domiphen bromide (2); D7, Domiphen bromide (3); D8, Domiphen bromide (4); D9, Nordihydroguaiacetic acid (1); D10, Nordihydroguaiacetic acid (2); D11, Nordihydroguaiacetic acid (3); D12, Nordihydroguaiacetic acid (4); E1, Alexidine (1); E2, Alexidine (2); E3, Alexidine (3); E4, Alexidine (4); E5, 5-Nitro-2-furaldehyde semicarbazone (1); E6, 5-Nitro-2-furaldehyde semicarbazone (2); E7, 5-Nitro-2-furaldehyde semicarbazone (3); E8, 5-Nitro-2-furaldehyde semicarbazone (4); E9, Methyl viologen (1); E10, Methyl viologen (2); E11, Methyl viologen (3); E12, Methyl viologen (4); F1, 3, 4-Dimethoxybenzyl alcohol (1); F2, 3, 4-Dimethoxybenzyl alcohol (2); F3, 3, 4-Dimethoxybenzyl alcohol (3); F4, 3, 4-Dimethoxybenzyl alcohol (4); F5, Oleandomycin (1); F6, Oleandomycin (2); F7, Oleandomycin (3); F8, Oleandomycin (4); F9, Puromycin (1); F10, Puromycin (2); F11, Puromycin (3); F12, Puromycin (4); G1,

	CCCP (1); G2, CCCP (2); G3, CCCP (3); G4, CCCP (4); G5, Sodium azide (1); G6, Sodium azide (2); G7, Sodium azide (3); G8, Sodium azide (4); G9, Menadione (1); G10, Menadione (2); G11, Menadione (3); G12, Menadione (4); H1, 2-Nitroimidazole (1); H2, 2-Nitroimidazole (2); H3, 2-Nitroimidazole (3); H4, 2-Nitroimidazole (4); H5, Hydroxyurea (1); H6, Hydroxyurea (2); H7, Hydroxyurea (3); H8, Hydroxyurea (4); H9, Zinc chloride (1); H10, Zinc chloride (2); H11, Zinc chloride (3); H12, Zinc chloride (4).
PM16A – chemical	A1, Cefotaxime (1); A2, Cefotaxime (2); A3, Cefotaxime (3); A4, Cefotaxime (4); A5, Phosphomycin (1); A6, Phosphomycin (2); A7, Phosphomycin (3); A8, Phosphomycin (4); A9, 5-Chloro-7-iodo-8-hydroxyquinoline (1); A10, 5-Chloro-7-iodo-8-hydroxyquinoline (2); A11, 5-Chloro-7-iodo-8-hydroxyquinoline (3); A12, 5-Chloro-7-iodo-8-hydroxyquinoline (4); B1, Norfloxacin (1); B2, Norfloxacin (2); B3, Norfloxacin (3); B4, Norfloxacin (4); B5, Sulfanilamide (1); B6, Sulfanilamide (2); B7, Sulfanilamide (3); B8, Sulfanilamide (4); B9, Trimethoprim (1); B10, Trimethoprim (2); B11, Trimethoprim (3); B12, Trimethoprim (4); C1, Dichlofluanid (1); C2, Dichlofluanid (2); C3, Dichlofluanid (3); C4, Dichlofluanid (4); C5, Protamine sulfate (1); C6, Protamine sulfate (2); C7, Protamine sulfate (3); C8, Protamine sulfate (4); C9, Cetylpyridinium chloride (1); C10, Cetylpyridinium chloride (2); C11, Cetylpyridinium chloride (3); C12, Cetylpyridinium chloride (4); D1, 1-Chloro -2,4-dinitrobenzene (1); D2, 1-Chloro -2,4-dinitrobenzene (2); D3, 1-Chloro -2,4-dinitrobenzene (3); D4, 1-Chloro -2,4-dinitrobenzene (4); D5, Diamide (1); D6, Diamide (2); D7, Diamide (3); D8, Diamide (4); D9, Cinoxacin (1); D10, Cinoxacin (2); D11, Cinoxacin (3); D12, Cinoxacin (4); E1, Streptomycin (1); E2, Streptomycin (2); E3, Streptomycin (3); E4, Streptomycin (4); E5, 5-Azacytidine (1); E6, 5-Azacytidine (2); E7, 5-Azacytidine (3); E8, 5-Azacytidine (4); E9, Rifamycin SV (1); E10, Rifamycin SV (2); E11, Rifamycin SV (3); E12, Rifamycin SV (4); F1, Potassium tellurite (1); F2, Potassium tellurite (2); F3, Potassium tellurite (3); F4, Potassium tellurite (4); F5, Sodium selenite (1); F6, Sodium selenite (2); F7, Sodium selenite (3); F8, Sodium selenite (4); F9, Aluminum sulfate (1); F10, Aluminum sulfate (2); F11, Aluminum sulfate (3); F12, Aluminum sulfate (4); G1, Chromium chloride (1); G2, Chromium chloride (2); G3, Chromium chloride (3); G4, Chromium chloride (4); G5, Ferric chloride (1); G6, Ferric chloride (2); G7, Ferric chloride (3); G8, Ferric chloride (4); G9, L-Glutamic-ghydroxamate (1); G10, L-Glutamic-ghydroxamate (2); G11, L-Glutamic-ghydroxamate (3); G12, L-Glutamic-ghydroxamate (4); H1, Glycine hydroxamate (1); H2, Glycine hydroxamate (2); H3, Glycine hydroxamate (3); H4, Glycine hydroxamate (4); H5, Chloroxylenol (1); H6, Chloroxylenol (2); H7, Chloroxylenol (3); H8, Chloroxylenol (4); H9, Sorbic acid (1); H10, Sorbic acid (2); H11, Sorbic acid (3); H12, Sorbic acid (4).
PM17A – chemical	A1, D-Serine (1); A2, D-Serine (2); A3, D-Serine (3); A4, D-Serine (4); A5, β -ChloroL-alanine hydrochloride (1); A6, β -ChloroL-alanine hydrochloride (2); A7, β -ChloroL-alanine hydrochloride (3); A8, β -ChloroL-alanine hydrochloride (4); A9, Thiosalicylic acid (1); A10, Thiosalicylic acid (2); A11, Thiosalicylic acid (3); A12, Thiosalicylic acid (4); B1, Sodium salicylate (1); B2, Sodium salicylate (2); B3, Sodium salicylate (3); B4, Sodium salicylate (4); B5, Hygromycin B (1); B6, Hygromycin B (2); B7, Hygromycin B (3); B8, Hygromycin B (4); B9, Ethionamide (1); B10, Ethionamide (2); B11, Ethionamide (3); B12, Ethionamide (4); C1, 4-Aminopyridine (1); C2, 4-Aminopyridine (2); C3, 4-Aminopyridine (3); C4, 4-Aminopyridine (4); C5, Sulfachloropyridazine (1); C6, Sulfachloropyridazine (2); C7, Sulfachloropyridazine (3); C8, Sulfachloropyridazine (4); C9, Sulfamonomethoxine (1); C10, Sulfamonomethoxine (2); C11, Sulfamonomethoxine (3); C12, Sulfamonomethoxine (4); D1, Oxycarboxin (1); D2, Oxycarboxin (2); D3, Oxycarboxin (3); D4, Oxycarboxin (4); D5, 3-Amino-1,2,4-triazole (1); D6, 3-Amino-1,2,4-triazole (2); D7, 3-Amino-1,2,4-triazole (3); D8, 3-Amino-1,2,4-triazole (4); D9, Chlorpromazine (1); D10, Chlorpromazine (2); D11, Chlorpromazine (3); D12, Chlorpromazine (4); E1, Niaproof (1); E2, Niaproof (2); E3, Niaproof (3); E4, Niaproof (4); E5, Compound 48/80 (1); E6, Compound 48/80 (2); E7, Compound 48/80 (3); E8, Compound 48/80 (4); E9, Sodium tungstate (1); E10, Sodium tungstate (2); E11, Sodium tungstate (3); E12, Sodium tungstate (4); F1, Lithium chloride (1); F2, Lithium chloride (2); F3, Lithium chloride (3); F4, Lithium chloride (4); F5, DL-Methionine hydroxamate (1); F6, DL-Methionine hydroxamate (2); F7, DL-Methionine hydroxamate (3); F8, DL-Methionine hydroxamate (4); F9, Tannic acid (1); F10, Tannic acid (2); F11, Tannic acid (3); F12, Tannic acid (4); G1, Chlorambucil (1); G2, Chlorambucil (2); G3, Chlorambucil (3); G4, Chlorambucil (4); G5, Cefamandole nafate (1); G6, Cefamandole nafate (2); G7, Cefamandole nafate (3); G8, Cefamandole nafate (4); G9, Cefoperazone (1); G10, Cefoperazone (2); G11, Cefoperazone (3); G12, Cefoperazone (4); H1, Cefsulodin (1); H2, Cefsulodin (2); H3, Cefsulodin (3); H4, Cefsulodin (4); H5, Caffeine (1); H6, Caffeine (2); H7, Caffeine (3); H8, Caffeine (4); H9, Phenylarsine oxide (1); H10, Phenylarsine oxide (2); H11, Phenylarsine oxide (3); H12, Phenylarsine oxide (4).
PM18C – chemical	A1, Ketoprofen (1); A2, Ketoprofen (2); A3, Ketoprofen (3); A4, Ketoprofen (4); A5, Sodium pyrophosphate decahydrate (1); A6, Sodium pyrophosphate decahydrate (2); A7, Sodium pyrophosphate decahydrate (3); A8, Sodium pyrophosphate decahydrate (4); A9, Thiamphenicol (1); A10, Thiamphenicol (2); A11, Thiamphenicol (3); A12, Thiamphenicol (4); B1, Trifluorothymidine (1); B2, Trifluorothymidine (2); B3, Trifluorothymidine (3); B4, Trifluorothymidine (4); B5, Pipemidic Acid (1); B6, Pipemidic Acid (2); B7, Pipemidic Acid (3); B8, Pipemidic Acid (4); B9, Azathioprine (1); B10, Azathioprine (2); B11, Azathioprine (3); B12, Azathioprine (4); C1, Poly-L-lysine (1); C2, Poly-L-lysine

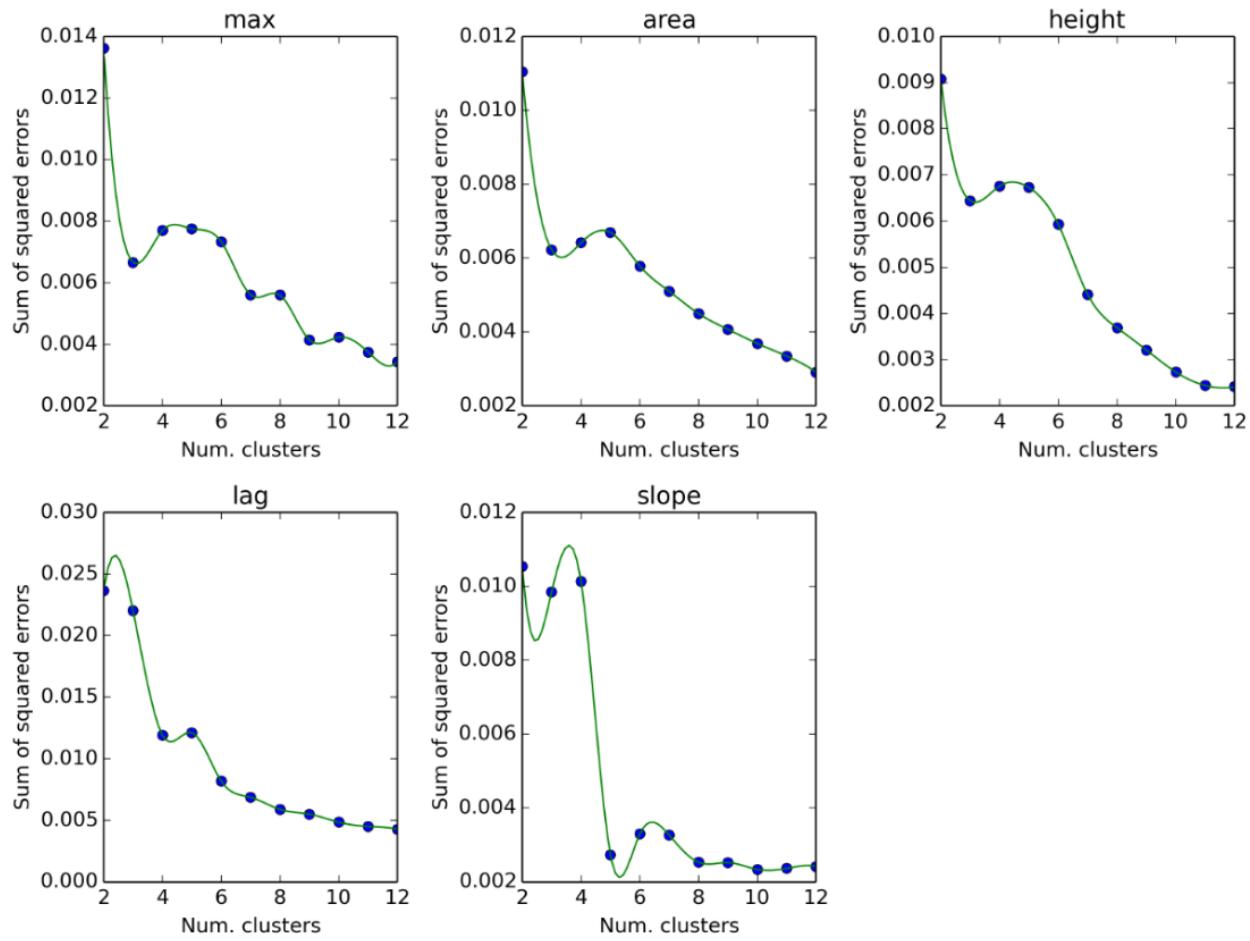
	(2); C3, Poly-L-lysine (3); C4, Poly-L-lysine (4); C5, Sulfisoxazole (1); C6, Sulfisoxazole (2); C7, Sulfisoxazole (3); C8, Sulfisoxazole (4); C9, Pentachlorophenol (1); C10, Pentachlorophenol (2); C11, Pentachlorophenol (3); C12, Pentachlorophenol (4); D1, Sodium m-arsenite (1); D2, Sodium m-arsenite (2); D3, Sodium m-arsenite (3); D4, Sodium m-arsenite (4); D5, Sodium bromate (1); D6, Sodium bromate (2); D7, Sodium bromate (3); D8, Sodium bromate (4); D9, Lidocaine (1); D10, Lidocaine (2); D11, Lidocaine (3); D12, Lidocaine (4); E1, Sodium metasilicate (1); E2, Sodium metasilicate (2); E3, Sodium metasilicate (3); E4, Sodium metasilicate (4); E5, Sodium m-periodate (1); E6, Sodium m-periodate (2); E7, Sodium m-periodate (3); E8, Sodium m-periodate (4); E9, Antimony (III) chloride (1); E10, Antimony (III) chloride (2); E11, Antimony (III) chloride (3); E12, Antimony (III) chloride (4); F1, Semicarbazide (1); F2, Semicarbazide (2); F3, Semicarbazide (3); F4, Semicarbazide (4); F5, Tinidazole (1); F6, Tinidazole (2); F7, Tinidazole (3); F8, Tinidazole (4); F9, Aztreonam (1); F10, Aztreonam (2); F11, Aztreonam (3); F12, Aztreonam (4); G1, Triclosan (1); G2, Triclosan (2); G3, Triclosan (3); G4, Triclosan (4); G5, 3,5-Diamino-1,2,4-triazole (Guanazole) (1); G6, 3,5-Diamino-1,2,4-triazole (Guanazole) (2); G7, 3,5-Diamino-1,2,4-triazole (Guanazole) (3); G8, 3,5-Diamino-1,2,4-triazole (Guanazole) (4); G9, Myricetin (1); G10, Myricetin (2); G11, Myricetin (3); G12, Myricetin (4); H1, 5-fluoro-5'-deoxyuridine (1); H2, 5-fluoro-5'-deoxyuridine (2); H3, 5-fluoro-5'-deoxyuridine (3); H4, 5-fluoro-5'-deoxyuridine (4); H5, 2-Phenylphenol (1); H6, 2-Phenylphenol (2); H7, 2-Phenylphenol (3); H8, 2-Phenylphenol (4); H9, Plumbagin (1); H10, Plumbagin (2); H11, Plumbagin (3); H12, Plumbagin (4).
PM19 – chemical	A1, Josamycin (1); A2, Josamycin (2); A3, Josamycin (3); A4, Josamycin (4); A5, Gallic acid (1); A6, Gallic acid (2); A7, Gallic acid (3); A8, Gallic acid (4); A9, Coumarin (1); A10, Coumarin (2); A11, Coumarin (3); A12, Coumarin (4); B1, Methyltriocetylammmonium chloride (1); B2, Methyltriocetylammnonium chloride (2); B3, Methyltriocetylammnonium chloride (3); B4, Methyltriocetylammnonium chloride (4); B5, Harmane (1); B6, Harmane (2); B7, Harmane (3); B8, Harmane (4); B9, 2,4-Dintrophenol (1); B10, 2,4-Dintrophenol (2); B11, 2,4-Dintrophenol (3); B12, 2,4-Dintrophenol (4); C1, Chlorhexidine (1); C2, Chlorhexidine (2); C3, Chlorhexidine (3); C4, Chlorhexidine (4); C5, Umbelliferone (1); C6, Umbelliferone (2); C7, Umbelliferone (3); C8, Umbelliferone (4); C9, Cinnamic acid (1); C10, Cinnamic acid (2); C11, Cinnamic acid (3); C12, Cinnamic acid (4); D1, Disulphiram (1); D2, Disulphiram (2); D3, Disulphiram (3); D4, Disulphiram (4); D5, Iodonitro Tetrazolium Violet (1); D6, Iodonitro Tetrazolium Violet (2); D7, Iodonitro Tetrazolium Violet (3); D8, Iodonitro Tetrazolium Violet (4); D9, Phenyl- methylsulfonylfluoride (PMSF) (1); D10, Phenylmethylsulfonylfluoride (PMSF) (2); D11, Phenyl- methylsulfonylfluoride (PMSF) (3); D12, Phenylmethylsulfonylfluoride (PMSF) (4); E1, FCCP (1); E2, FCCP (2); E3, FCCP (3); E4, FCCP (4); E5, D,L-Thioctic Acid (1); E6, D,L-Thioctic Acid (2); E7, D,L-Thioctic Acid (3); E8, D,L-Thioctic Acid (4); E9, Lawsone (1); E10, Lawsone (2); E11, Lawsone (3); E12, Lawsone (4); F1, Phenethicillin (1); F2, Phenethicillin (2); F3, Phenethicillin (3); F4, Phenethicillin (4); F5, Blasticidin S (1); F6, Blasticidin S (2); F7, Blasticidin S (3); F8, Blasticidin S (4); F9, Sodium caprylate (1); F10, Sodium caprylate (2); F11, Sodium caprylate (3); F12, Sodium caprylate (4); G1, Lauryl sulfobetaine (1); G2, Lauryl sulfobetaine (2); G3, Lauryl sulfobetaine (3); G4, Lauryl sulfobetaine (4); G5, Dihydrostreptomycin (1); G6, Dihydrostreptomycin (2); G7, Dihydrostreptomycin (3); G8, Dihydrostreptomycin (4); G9, Hydroxylamine (1); G10, Hydroxylamine (2); G11, Hydroxylamine (3); G12, Hydroxylamine (4); H1, Hexammine cobalt (III) chloride (1); H2, Hexammine cobalt (III) chloride (2); H3, Hexammine cobalt (III) chloride (3); H4, Hexammine cobalt (III) chloride (4); H5, Thioglycerol (1); H6, Thioglycerol (2); H7, Thioglycerol (3); H8, Thioglycerol (4); H9, Polymyxin B (1); H10, Polymyxin B (2); H11, Polymyxin B (3); H12, Polymyxin B (4).
PM20B – chemical	A1, Amitriptyline (1); A2, Amitriptyline (2); A3, Amitriptyline (3) A4, Amitriptyline (4); A5, Apramycin (1); A6, Apramycin(2); A7, Apramycin (3); A8, Apramycin (4); A9, Benserazide (1); A10, Benserazide (2); A11, Benserazide (3); A12, Benserazide (4); B1, Orphenadrine (1); B2, Orphenadrine (2); B3, Orphenadrine (3); B4, Orphenadrine (4); B5, D,L-Propranolol (1); B6, D,L-Propranolol (2); B7, D,L-Propranolol (3), B8, D,L-Propranolol (4); B9, Tetrazolium Violet (1); B10, Tetrazolium Violet (2); B11, Tetrazolium Violet (3); B12, Tetrazolium Violet (4); C1, Thioridazine (1); C2, Thioridazine (2); C3, Thioridazine (3); C4, Thioridazine (4); C5, Atropine (1); C6, Atropine (2); C7, Atropine (3); C8, Atropine (4); C9, Ornidazole (1); C10, Ornidazole (2); C11, Ornidazole (3); C12, Ornidazole (4); D1, Proflavine (1); D2, Proflavine (2); D3, Proflavine (3); D4, Proflavine (4); D5, Ciprofloxacin (1); D6, Ciprofloxacin (2); D7, Ciprofloxacin (3); D8, Ciprofloxacin (4); D9, 18-Crown-6 Ether (1); D10, 18-Crown-6 Ether (2); D11, 18-Crown-6 Ether (3); D12, 18-Crown-6 Ether (4); E1, Crystal violet (1); E2, Crystal violet (2); E3, Crystal violet (3); E4, Crystal violet (4); E5, Dodine (1); E6, Dodine (2); E7, Dodine (3); E8, Dodine (4); E9, Hexachlorophene (1); E10, Hexachlorophene (2); E11, Hexachlorophene (3); E12, Hexachlorophene (4); F1, 4-Hydroxycoumarin (1); F2, 4-Hydroxycoumarin (2); F3, 4-Hydroxycoumarin (3); F4, 4-Hydroxycoumarin (4); F5, Oxytetracycline (1); F6, Oxytetracycline (2); F7, Oxytetracycline (3); F8, Oxytetracycline (4); F9, Pridinol (1); F10, Pridinol (2); F11, Pridinol (3); F12, Pridinol (4); G1, Captan (1); G2, Captan (2); G3, Captan (3); G4, Captan (4); G5, 3,5-Dinitrobenzene (1); G6, 3,5-Dinitrobenzene (2); G7, 3,5-Dinitrobenzene (3); G8, 3,5-Dinitrobenzene (4); G9, 8-Hydroxyquinoline (1); G10, 8-Hydroxyquinoline (2); G11, 8-Hydroxyquinoline (3); G12, 8-Hydroxyquinoline (4); H1, Patulin (1); H2, Patulin (2); H3, Patulin (3); H4, Patulin (4); H5, Tolyfluanid (1); H6, Tolyfluanid (2); H7, Tolyfluanid

	(3); H8, Tolyfluauad (4); H9, Troleandomycin (1); H10, Troleandomycin (2); H11, Troleandomycin (3); H12, Troleandomycin (4).
--	--

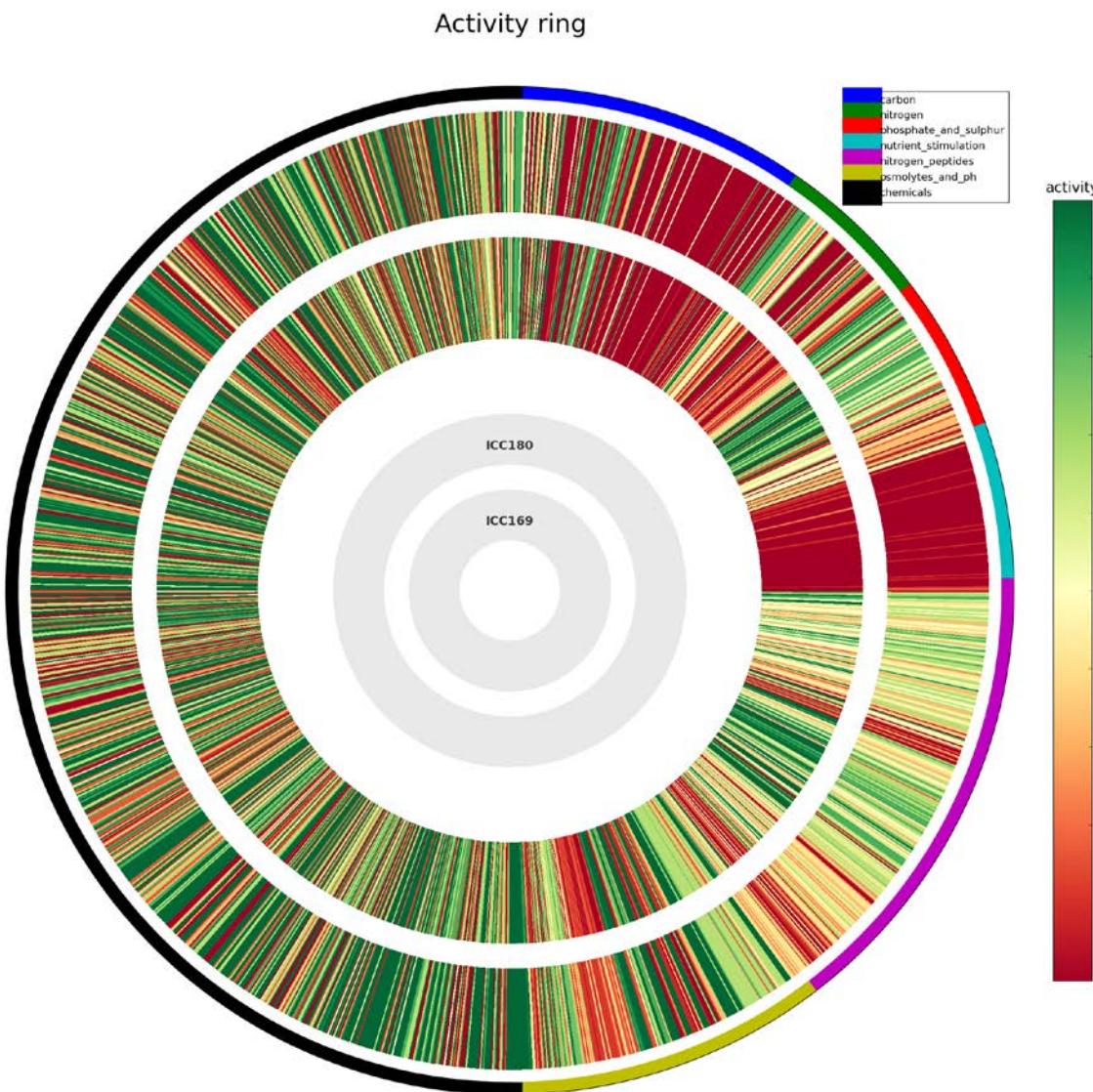
Supplementary Table 2. Genes missing from pCROD1 of *C. rodentium* ICC180
List of genes missing from plasmid pCROD1 of *C. rodentium* ICC180, as determined by sequencing

Gene	Location	Function
<i>ROD_RS25055</i>	240..494	Replication regulatory protein repA2
<i>ROD_RS25060</i>	797..1654	Replication protein
<i>ROD_RS25065</i>	2593..3246	Hypothetical protein
<i>ROD_RS25070</i>	3339..3596	Antitoxin
<i>ROD_RS25075</i>	3598..3930	Hypothetical protein
<i>ROD_RS25080</i>	4318..4614	Transposase
<i>ROD_RS25085</i>	5726..6955	Autotransporter strand-loop-strand
<i>ROD_RS25090</i>	6939..11720	Autotransporter
<i>ROD_RS25095</i>	12358..12558	Hypothetical protein
<i>ROD_RS25100</i>	12814..13043	Transposase
<i>ROD_RS25105</i>	14045..14563	Fimbrial protein
<i>ROD_RS25110</i>	14636..17053	Fimbrial protein
<i>ROD_RS25115</i>	17046..17738	Fimbrial protein
<i>ROD_RS25120</i>	18251..18820	Hypothetical protein
<i>ROD_RS25125</i>	18945..19904	Hypothetical protein
<i>ROD_RS25130</i>	20068..20844	EAL domain-containing protein
<i>ROD_RS25135</i>	22328..26254	Autotransporter
<i>ROD_RS25140</i>	26743..26993	Toxin HigB-2
<i>ROD_RS25145</i>	27079..27339	Transcriptional regulator
<i>ROD_RS25150</i>	27957..30536	Usher protein
<i>ROD_RS25155</i>	30578..31048	Hypothetical protein
<i>ROD_RS25160</i>	33519..33941	Twitching motility protein PilT
<i>ROD_RS25165</i>	33938..34168	Virulence factor
<i>ROD_RS25170</i>	34837..35055	Hypothetical protein
<i>ROD_RS25175</i>	35057..35362	Hypothetical protein
<i>ROD_RS25180</i>	35364..35690	Hypothetical protein
<i>ROD_RS25185</i>	35680..36471	Resolvase
<i>ROD_RS25190</i>	36627..40730	Autotransporter
<i>ROD_RS25730</i>	41808..43358	Hypothetical protein
<i>ROD_RS25205</i>	43907..44329	Entry exclusion protein 2
<i>ROD_RS25210</i>	44567..45523	Hypothetical protein
<i>ROD_RS25215</i>	45875..46504	Serine recombinase
<i>ROD_RS25220</i>	46779..47306	Putative resolvase
<i>ROD_RS25225</i>	47600..48241	Chromosome partitioning protein ParA
<i>ROD_RS25230</i>	48333..48665	Molecular chaperone GroEL
<i>ROD_RS25235</i>	49280..50002	DNA repair protein
<i>ROD_RS25240</i>	50082..51653	Transposase
<i>ROD_RS25250</i>	52020..52697	Transposase
<i>ROD_RS25255</i>	52721..52750	Endonuclease
<i>ROD_RS25260</i>	53458..54144	Hypothetical protein
<i>ROD_RS25265</i>	54141..54449	Hypothetical protein

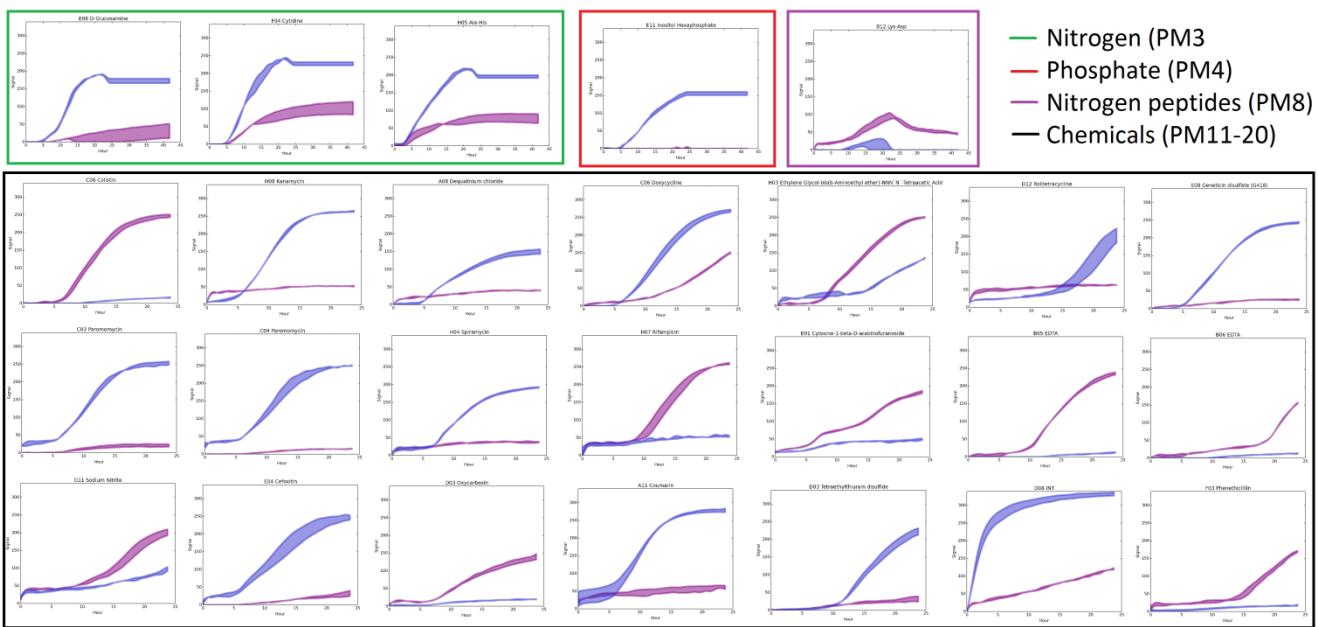
3.3.6 Supplementary Figures



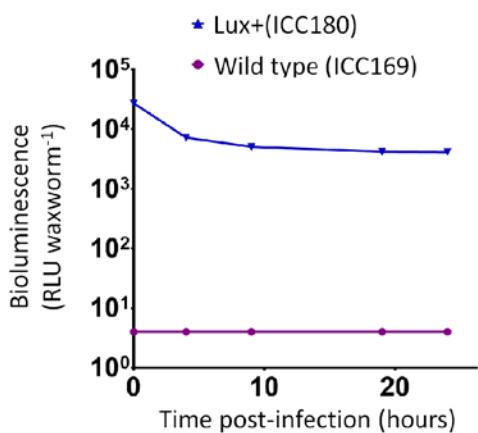
Supplementary Fig. 1. Elbow tests of phenotypic microarray array data to determine the number of clusters appropriate for k-means clustering. Data was analysed using the DuctApe software suite.



Supplementary Fig. 2. The growth of *C. rodentium* ICC180 compared to its non-bioluminescent parent strain ICC169 as assessed by phenotypic microarray (PM). Wildtype *C. rodentium* ICC169 and its bioluminescent derivative ICC180 were grown on two separate occasions using PM plates 1-20. Activity rings from the PM data are shown where the grey inner circles indicate the strains' order and the external circle indicates the PM categories (see Key). The activity index (AV) was calculated for each strain in response to each well and the values for ICC169 are shown as colour stripes going from red (AV = 0 [not active]) to green (AV = 6 [active]); 7 total k-means clusters.



Supplementary Fig. 3. Phenotypic microarray (PM) growth curves of *C. rodentium* ICC180 and its non-bioluminescent parent strain ICC169 which are significantly different. Wildtype *C. rodentium* ICC169 (shown as purple lines) and its bioluminescent derivative ICC180 (shown as blue lines) were grown on two separate occasions using PM plates 1-20 (categorised by colour [see Key]). Differences between the growth of ICC169 and ICC180 in each individual well were analysed using the moderated t-test provided by limma²⁸. Wells in which the differences had an adjusted p-value of less than 0.5 (stringent cut-off) are shown.



Supplementary Fig. 4. Infection of larvae of the Greater Wax Moth *Galleria mellonella* with bioluminescent *C. rodentium* ICC180 can be visualised by luminometry. Groups of larvae ($n = 10$) of the Greater Wax Moth *Galleria mellonella* were infected with $\sim 10^8$ CFU of *C. rodentium* ICC169 or ICC180 and monitored for bioluminescence using a plate luminometer. Data (medians with ranges) is presented from experiments performed on 3 separate occasions and is given as relative light units [RLU] waxworm⁻¹.

3.3 Conclusions

The thorough comparison between *C. rodentium* ICC169 and its bioluminescent derivative ICC180 showed surprisingly little effect of light expression on *C. rodentium*. Of note, the fitness cost observed following prolonged competition in the mouse environment shows ICC180 has a disadvantage compared to its parent strain, being shed at approximately 10-fold lower levels than ICC169. The data presented in this paper provides the benchmark for comparison between the *in vivo*-adapted *C. rodentium* and ICC169, which will be presented in Chapter 5. If the fitness disadvantage that the ancestral ICC180 strain has is no longer present when the *in vivo*-adapted strains are competed against ICC169, then this would suggest an advantage of the *in vivo*-adapted strains over the ancestral ICC180 strain.

Chapter 4: Experimental *in vivo* Evolution of *C. rodentium*

4.1 Introduction

While laboratory experiments investigating evolution have been around for many decades, these have seldom reached outside of laboratory flasks and artificial media. Although such approaches have provided evidence of important evolutionary principles, the fact remains: evolutionary processes and adaptation are sensitive to subtle changes. It therefore stands to reason that adaptation to an artificial environment cannot fully encapsulate the types and breadths of changes that may be happening within more complex environments. Similarly, one consequence of bacterial adaptation that is of crucial importance to humanity is the change observed in pathogenic species, resulting in, for example, evasion of our immune system and the antibiotics we use to treat infections, increases in virulence, or changes in host range and the emergence of new diseases. Studying the interaction of a pathogen with its host fundamentally requires a pathogen interacting with its host, and this relationship cannot be fully modelled in an *in vitro* system¹⁹². Therefore, I have chosen to study the *in vivo* adaptation of the natural mouse pathogen *Citrobacter rodentium* to its mouse host following sequential tightly controlled transmission from animal to animal. By using this model system I can experimentally test the effects of different conditions on the evolution of *C. rodentium*.

Antibiotics, a natural product of ongoing microbial warfare between bacterial and fungal species, are increasingly present in our day-to-day lives as a result of large scale use in human medicine, veterinary medicine, aquaculture, and the agricultural industry^{193,194}. In 2002, Wise estimated the total worldwide consumption of antibiotics to be between 100,000 and 200,000 ton per year¹⁹⁴. A study by Kümmerer and Henninger (2003) showed that, in a German hospital, 70% of the antibiotics consumed were excreted unchanged into the sewage system¹⁹⁵. In the agricultural industry, antibiotics are given to healthy livestock to improve overall growth of the animal and quality of the meat^{196,197}. These factors all contribute to the increasing environmental antibiotic concentrations, as studied extensively in many different countries^{198–203}, and with as yet unknown effects on microbial evolution.

Sub-inhibitory levels of antibiotics have been shown to affect gene expression with knock-on impacts on pathogenicity, as shown by Goncalves and colleagues in a mouse model of urinary tract infection using *Staphylococcus saprophyticus* and uropathogenic *E. coli*²⁰⁴. Studies have also shown that antibiotic use has a profound effect on the microbiota of mice as well as impacts on other factors including, but not limited to, the phageome (a measure of the variety of bacteriophages present in an animal)²⁰⁵. In a study by Modi and colleagues, laboratory mice were exposed to either ciprofloxacin (a quinolone, like nalidixic acid) or ampicillin²⁰⁵. The authors demonstrated that the phage metagenome is a potential source of antibiotic resistance genes. Phageomes from mice treated with the quinolone had genes

important in replication and DNA repair pathways, including homologous recombination, as well as genes involved in the bacterial SOS response system²⁰⁵.

The question remains: will antibiotic use increase or decrease variation and therefore adaptation? To investigate the effect of antibiotic use on the evolution of a pathogen, my *in vivo* evolution experiment has two distinct treatment groups: one group of five independent lineages which received untreated food and water, and five independent lineages which received drinking water supplemented with a low-dose of the quinolone antibiotic nalidixic acid. The strain of *C. rodentium* I am investigating is already resistant to nalidixic acid, however I hypothesise that antibiotic treatment of the mice will have an impact on the normal gut flora, and that the altered gut flora will influence the evolution trajectory of *C. rodentium*. Alternatively the impacts on the normal gut flora may have little to no effect.

4.2 Methods

A schematic of the *in vivo* evolution experiment is shown in Fig 4.1. I infected 10 individual lineages of C57BL/6 mice (five for each treatment) with *C. rodentium* ICC180 and allowed the infection to transmit to naïve mice through tightly controlled mouse-to-mouse exposure. The infection was spread between animals via natural transmission, through normal grooming processes and coprophagia of infectious stools. Five of the lineages of mice received nalidixic acid treatment in their drinking water and were designated N1, N2, N3, N4 and N5. The remaining five lineages of mice received no antibiotic and were designated W1, W2, W3, W4 and W5. Bacteria recovered from the stool of mice from each lineage were given the same designation.

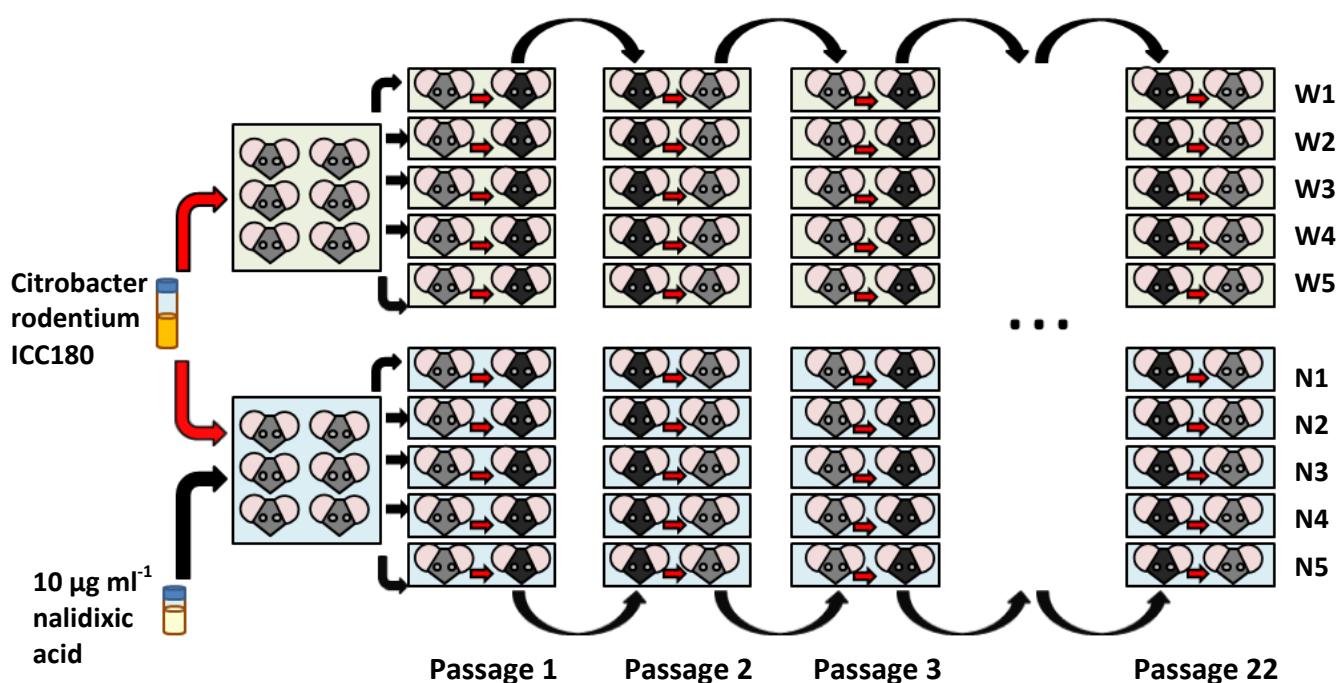


Figure 4.1. Schematic of the *Citrobacter rodentium* *in vivo* evolution experiment. A schematic of the *in vivo* mouse evolution experiment. 12 C57BL6 mice were divided into two treatment groups and artificially infected with *C. rodentium* ICC180 via oral gavage. The lower treatment group also received $10 \mu\text{g ml}^{-1}$ of nalidixic acid in their drinking water. At the peak of infection (approximately 7 days post-infection), animals with the highest bacterial shedding counts were divided into individual cages and housed with an uninfected animal (Passage 1 [P1]). These naïve animals were infected via coprophagia of infectious stools or via natural grooming processes. At the peak of the new infection (approximately 7 days post co-mingling), the newly infected animals were placed into fresh cages and cohoused with a new uninfected animal (P2). This process was continued over a period of five months, by which time 22 passages had occurred.

4.3 Results

4.3.1 Impact of nalidixic acid treatment on the mouse gut microbiota

To assess the effect of treatment with the antibiotic nalidixic acid on the gut microbiota of mice, we randomly allocated 12 C57BL/6 mice into two groups. One group received a low dose of nalidixic acid ($10 \mu\text{g ml}^{-1}$) in their drinking water, while the other group received untreated drinking water. Stools were recovered from each animal before and after 24 hours treatment, and snap frozen in a dry ice and ethanol bath. DNA was extracted from the stool samples using a MoBio PowerSoil® kit (Qiagen) and 16S sequencing analysis performed (Elahe Kia and Associate Professor Mike Taylor [University of Auckland]). The 16S pipeline was used as previously published²⁰⁶, with the following minor deviations: samples were rarefied to a depth of 2000 reads, and eukaryotic OTUs identified as mitochondrial sequences were not removed. From the extensive sets of operational taxonomic units (OTUs) identified in each treatment group, a filter of 0.1% was used to retain only those OTUs present in at least 0.1% of the total sequence counts. The complete list of filtered OTUs present in the mice in each treatment group before and after treatment can be found in appendix 1.

Prior to treatment with nalidixic acid, there was no statistically significant difference between the relative abundance of OTUs in each of the two groups. However, 24-hours treatment with nalidixic acid resulted in a statistically significant difference in relative abundance in nine OTUs (Table 4.1).

Table 4.1. OTUs with a significant difference in relative abundance after 24 hours treatment with nalidixic acid

OTU	Identified Taxonomy*	Untreated	Antibiotic-treated**	P value
27	Phylum: Bacteroidetes; Class: Bacteroidia Order: Bacteroidales	0.5	12.2	0.0028
30	Phylum: Bacteroidetes; Class: Bacteroidia Order: Bacteroidales	0	41.7	0.0021
50	Phylum: Bacteroidetes; Class: Bacteroidia Order: Bacteroidales	17.2	2.0	0.0039
60	Phylum: Firmicutes; Class: Clostridia Order: Clostridiales	0.7	33.7	0.0090
61	Phylum: Bacteroidetes; Class: Bacteroidia Order: Bacteroidales	0	13.2	0.0021
66	Phylum: Firmicutes; Class: Clostridia Order: Clostridiales; Family: Lachnospiraceae	0.2	9.2	0.0036
107	Phylum: Firmicutes; Class: Clostridia Order: Clostridiales; Family: Ruminococcaceae Genus: Ruminococcus	0	11.7	0.0073
117	Phylum: Firmicutes; Class: Clostridia Order: Clostridiales	0.2	4.7	0.0035
134	Phylum: Firmicutes; Class: Clostridia Order: Clostridiales	0	4.0	0.0073

* The identified taxonomy for a given OTU. Different OTUs may have the same identified taxonomy while still being distinct enough to be classified as a separate OTU.

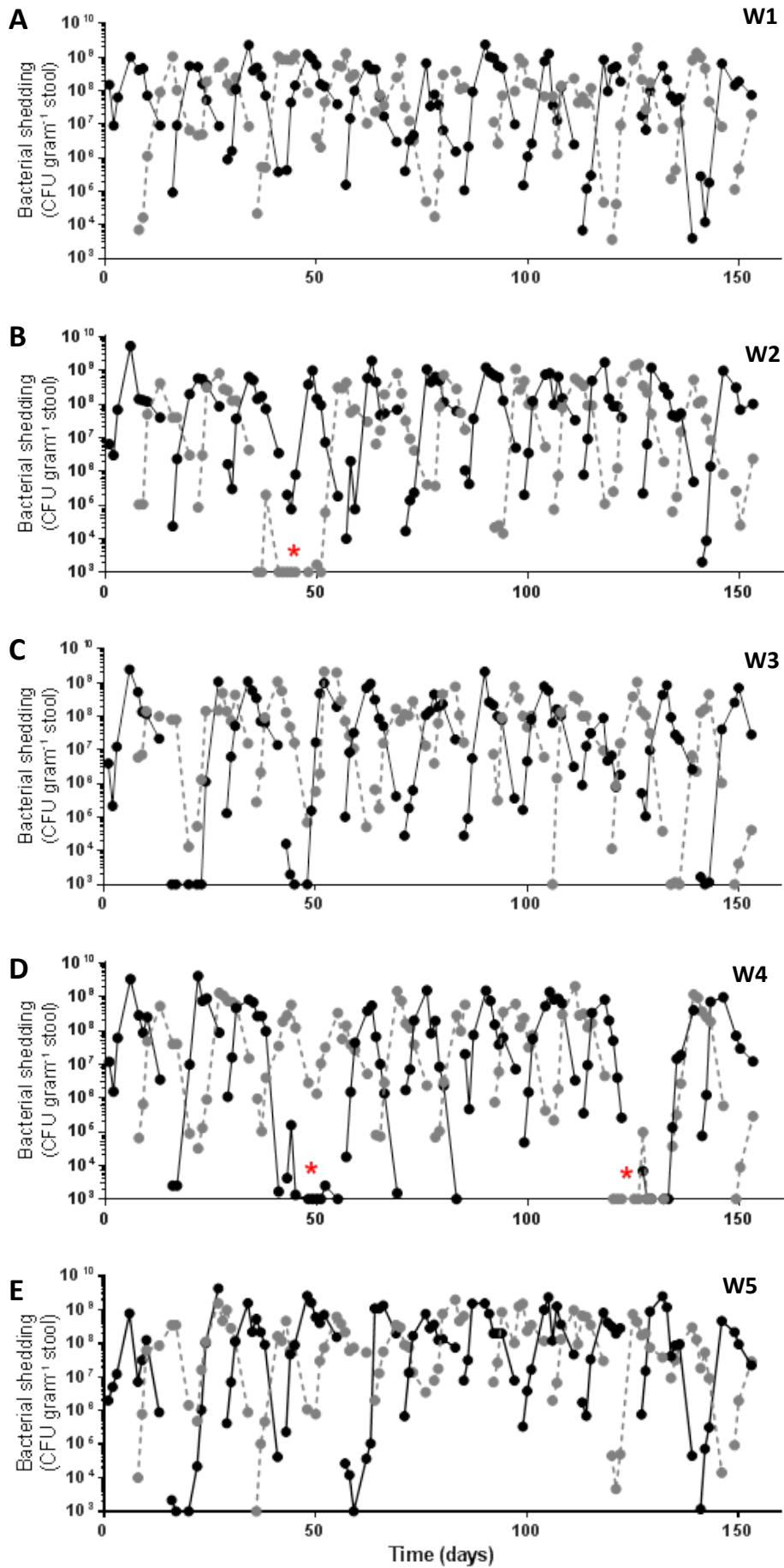
** Relative abundance after 24 hours treatment with 10 µg ml⁻¹ nalidixic acid in the drinking water. The group for each OTU with the highest relative abundance is highlighted in bold.

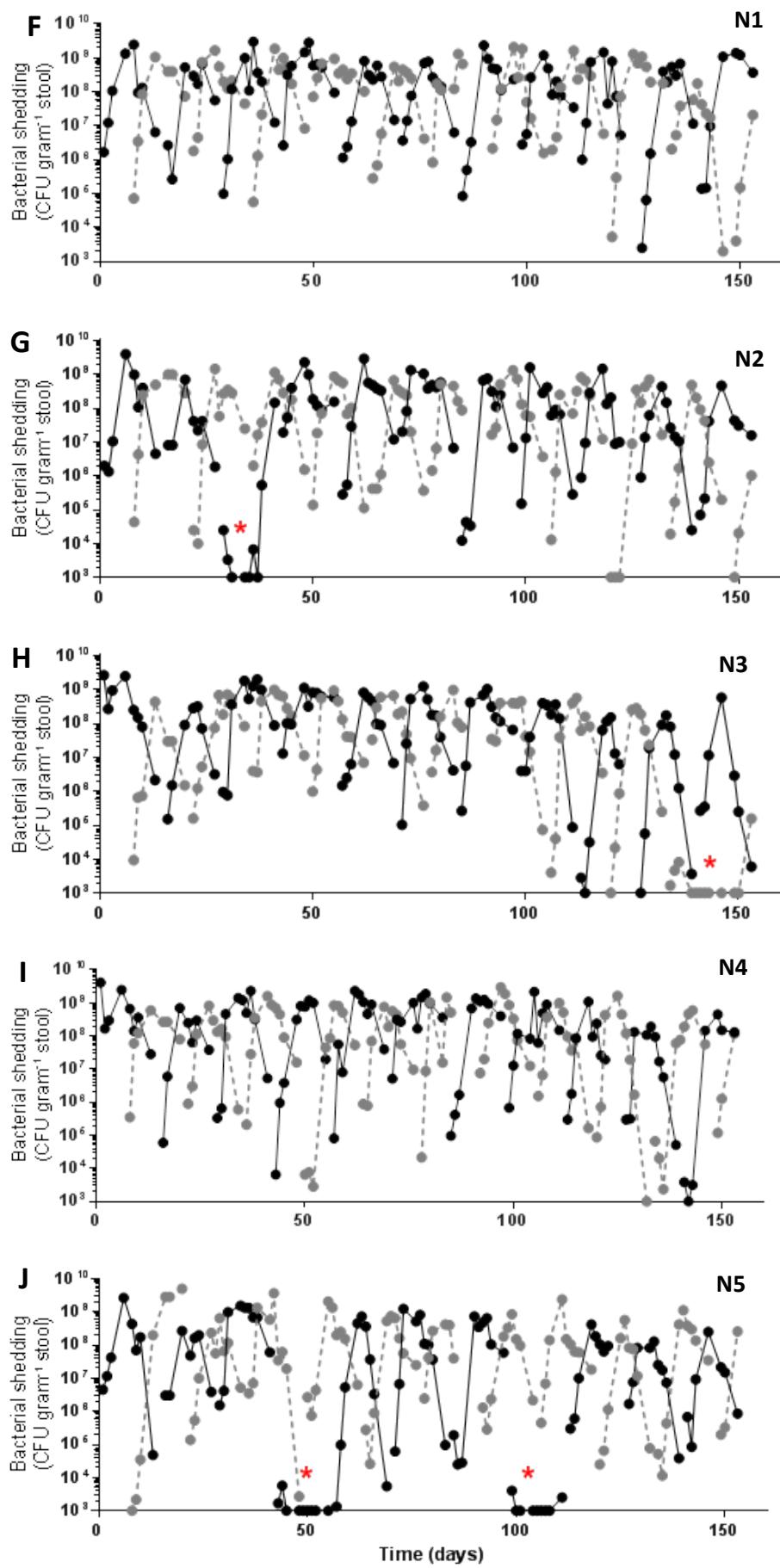
4.3.2 Bacterial shedding during the *in vivo* evolution experiment and the occurrence of transmission failure events

Throughout the *in vivo* evolution experiment, I monitored each individual animal, collecting their stool and plating it onto selective media in order to determine how much *C. rodentium* was shed from the animals in each infection chain. Fig. 4.2 shows the amount of *C. rodentium* shed from each infected animal (given as CFU gram^{-1} stool), with alternating black and grey lines indicating alternating mice; that is, a mouse depicted by a black line was cohoused and infected a mouse depicted by a grey line, which was subsequently infected and housed with a mouse depicted by a black line. At the beginning of an infection, the detectable number of *C. rodentium* recovered from stools is low, rapidly rising and then peaking at approximately 7 days post-infection. As the next successfully infected animal's *C. rodentium* shedding rises, the previous animal's infection begins to wane, lowering until the next co-mingling event at 14 days post-infection. At this stage, the animals were anaesthetised and blood taken by cardiac puncture. Tissues were taken from these animals and stored for future analysis.

For some lineages, the cohoused naïve animal failed to be successfully infected with *C. rodentium*, as shown by lines with consistently low *C. rodentium* recovered from stool samples (marked with a red star). These transmission failure events occurred three times in mice received untreated food and water (1 event in the W2 lineage, 2 events in the W4 lineage) and four times in mice receiving water supplemented with nalidixic acid (1 event in the N2 lineage, 1 event in the N3 lineage, and 2 events in the N5 lineage). This does not suggest that the nalidixic acid treatment increased the propensity or reduced the likelihood of a transmission failure event. Rather, these events appear to be sporadic, occurring at seemingly random times during the experiment, and perhaps reflect differences in the host mice and particular grooming and dietary habits. When transmission failed to occur, I artificially infected the next mouse in the infection chain by oral gavage with *C. rodentium* revived from stored stool samples from the previous animal in the infection chain. This artificial infection process never failed, suggesting that the transmission failure event was not due to the appearance of a transmission defective *C. rodentium* population in that lineage.

Figure 4.2. *C. rodentium* successfully transmits in controlled mouse-to-mouse infections over a period of 5 months (overleaf). Bacterial numbers shed (given as CFU gram^{-1} stool) from each individual mouse over the course of each transmission chain. Alternating black and grey lines indicate alternating mice; that is, the mouse depicted by a black line infected the subsequent mouse depicted by a grey line, and vice versa. In the cases where a mouse failed to be successfully infected (a transmission failure event, highlighted with a red star), the following mouse in the chain was artificially infected via oral gavage with bacteria revived from frozen stocks taken from the previous mouse. A-E, lineages which received untreated autoclaved food and water; F-J, lineages which received a low-dose of nalidixic acid ($10 \mu\text{g ml}^{-1}$) in their drinking water.





4.3.3 Comparison of *C. rodentium* ICC180 levels shed by mice during the *in vivo* evolution experiment

To assess whether there were any changes in the amounts of *C. rodentium* ICC180 shed from infected animals throughout the *in vivo* experiment, I calculated a value for the area under curve (AUC) from the amount of bacteria shed from each animal. How these values change over the course of each transmission chain is shown in Fig. 4.3. Transmission failure events are indicated as red crosses, while the linear regression slope detailed in Table 4.2 follows the trend of bacterial shedding over time. For the majority of lineages, there is no significant change in the AUC of *C. rodentium* shed from each animal. Lineages N2, N3, and N5 show a statistically significant downward trend of reduced AUC of *C. rodentium* shed over time (all $p < 0.05$). It is interesting to note that these three lineages all come from the treatment group that received drinking water supplemented with nalidixic acid. This may indicate a potential effect of the antibiotic treatment on overall infection dynamics over a prolonged period of time, occurring in 3/5 of the lineages receiving the drug. It is important to note that AUC data takes into account both the peak of bacterial shedding, but also duration of the peak and speed taken to reach the peak, therefore giving an indication of the numbers of *C. rodentium* shed throughout each individual infection, instead of just reflecting the fact that ICC180 reached sufficiently high numbers to propagate an infection.

Figure 4.3. *Citrobacter rodentium* is shed at relatively stable levels throughout the *in vivo* experimental evolution experiment (overleaf). The overall level of shedding of ICC180 by each individual animal is presented throughout the 5 month experiment as area under curve (AUC) values calculated from the amounts of bacteria shed per gram stool. A-E, evolution lineages from mice who received untreated food and water; F-J, evolution lineages from mice who received nalidixic acid ($10 \mu\text{g ml}^{-1}$) in their drinking water. The first and last passage data points have been removed as the AUC value for these animals is altered due to incomplete data on bacteria shed during the course of the initial and final infections. Transmission failure events are indicated as red crosses at the corresponding passage number. A linear regression trend line is shown for each lineage.

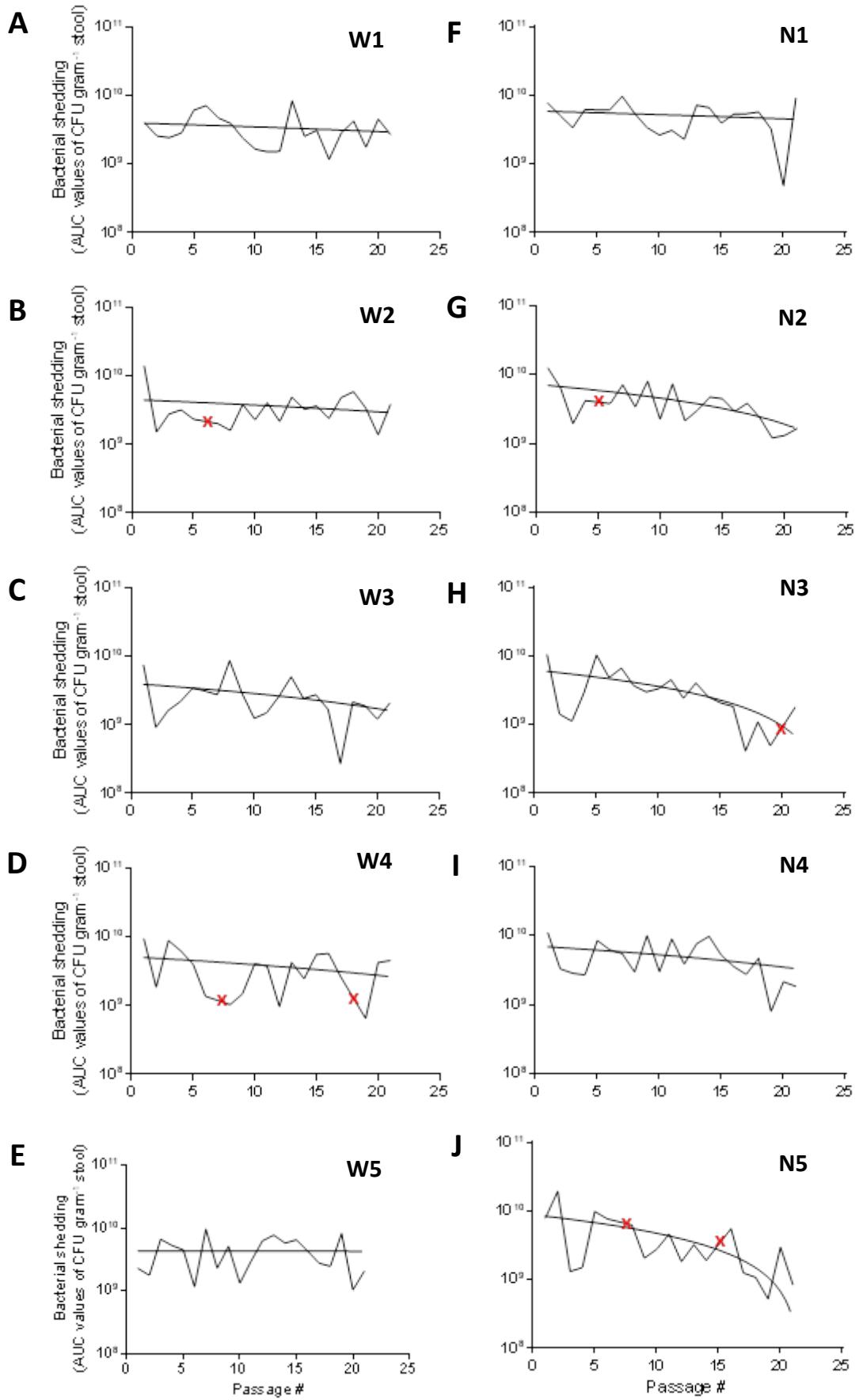


Table 4.2. Slopes of AUC of CFU shed in each transmission chain

Lineage	Slope*	P value**
W1	-5.11 x 10 ⁷ ± 7.06 x 10 ⁷	0.4781
W2	-7.69 x 10 ⁷ ± 1.02 x 10 ⁸	0.4592
W3	-1.15 x 10 ⁸ ± 7.10 x 10 ⁷	0.1206
W4	-1.19 x 10 ⁸ ± 9.44 x 10 ⁷	0.2243
W5	-3.31 x 10 ⁶ ± 9.51 x 10 ⁷	0.9726
N1	-6.85 x 10 ⁷ ± 8.33 x 10 ⁷	0.4214
N2	-2.67 x 10 ⁸ ± 8.62 x 10 ⁷	0.0062
N3	-2.65 x 10 ⁸ ± 9.20 x 10 ⁷	0.0100
N4	-1.78 x 10 ⁸ ± 1.05 x 10 ⁸	0.1061
N5	-4.08 x 10 ⁸ ± 1.41 x 10 ⁸	0.0103

* Slope of the linear regression trend line of AUC of *C. rodentium* shed over the course of the evolution experiment, as shown in Fig. 3.3. The value of the slope is presented ± the standard error.

** Statistically significant and non-significant deviations from 0 (a flat line, indicating no change) are shown in bold.

When the AUC values for *C. rodentium* shed from each animal for each lineage are pooled together, the majority of lineages appear similar (Fig 4.4). A statistically significant difference between the highest (N1) and the lowest (W3) can be detected (Kruskal-Wallis test with Dunn's post hoc multiple comparison test p<0.01). When comparing the two treatment groups (Fig 4.4B), there appears to be no change in the AUC of *C. rodentium* shed in the presence of nalidixic acid. This suggests that adapting in the presence of low-dose nalidixic acid does not impact on the total amount of *C. rodentium* ICC180 shed by infected animals over the 5 month period. The values for each lineage are variable, with wide whiskers on the box and whisker plot, perhaps reflecting variation in the host response to *C. rodentium* rather than fluidity of the lineages' infection attributes over the five month time period. Co-mingling and transmission of *C. rodentium* between animals was carried out at the peak of infection, which is approximately 7 days post-infection. The levels of *C. rodentium* ICC180 shed from each animal for each lineage over the course of the *in vivo* evolution experiment did not significantly differ between lineages (Fig 4.4C). Similarly, when comparing the two treatment groups (Fig 4.4D), there appears to be no change in the AUC of *C. rodentium* shed in the presence of nalidixic acid.

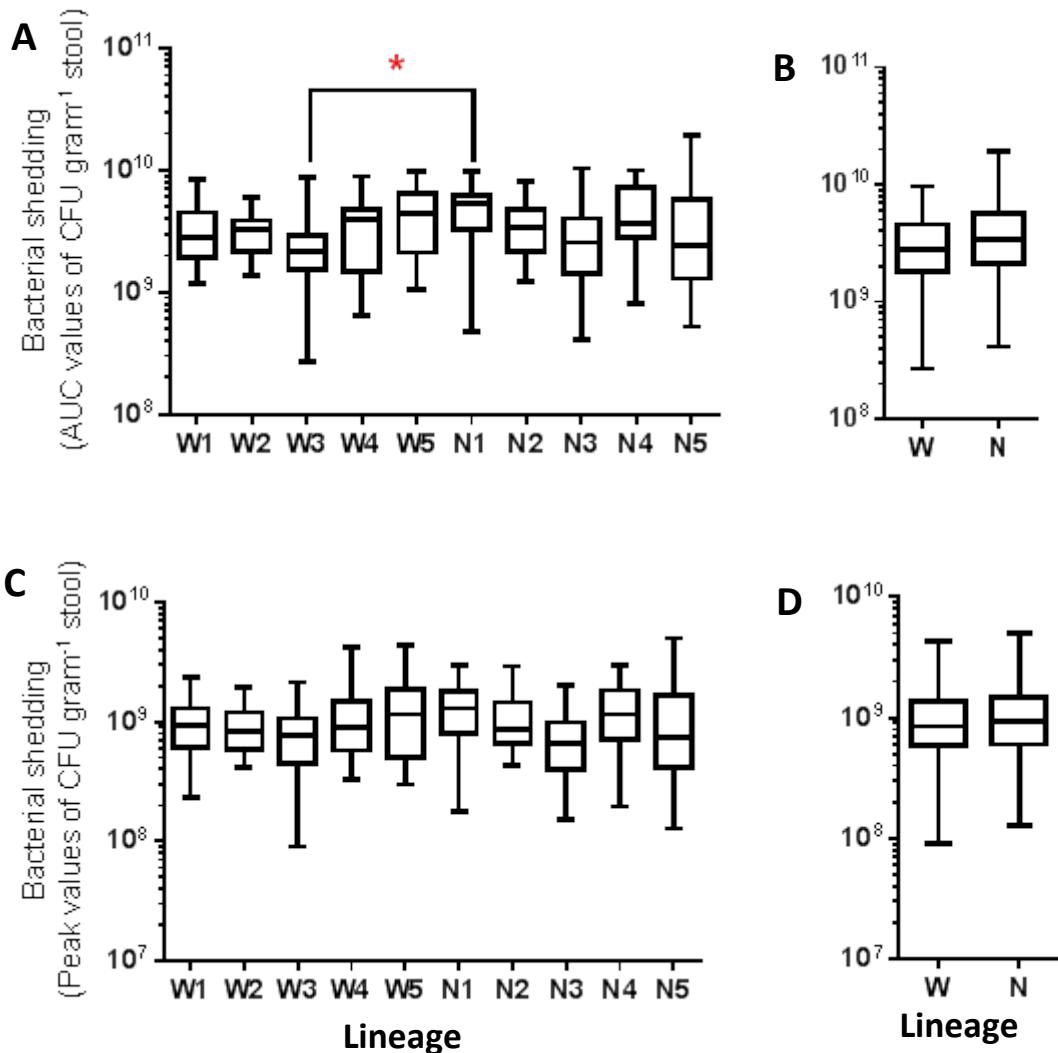


Figure 4.4. Administration of low-dose nalidixic acid in the drinking water does not change the overall amount or peak levels of *Citrobacter rodentium* shed over the 5-month evolution experiment. Total AUC values and peak levels for *C. rodentium* shed (A,C) for each lineage, and (B,D) for each treatment group, over the course of the 5 month evolution experiment, respectively. W denotes lineages which received untreated drinking water and food, N denotes lineages which received nalidixic acid ($10 \mu\text{g ml}^{-1}$) in their drinking water. * indicates a statistically significant difference between the lineages with the highest (N1) and the lowest (W3) AUC values (Kruskal-Wallis test with Dunn's post hoc multiple comparison test $p<0.01$).

4.3.4 Biophotonic imaging of *C. rodentium* ICC180 from within infected mice

C. rodentium ICC180 is a bioluminescent derivative of ICC169, containing a chromosomally located copy of the *lux* operon. The light produced by this strain can be detected *in vivo* by biophotonic imaging (BPI), using imaging systems such as the IVIS Kinetic (Perkin Elmer). BPI allows a real-time assessment of the location and numbers of bacteria inside a living host to be made by detecting the light emitted from bioluminescent bacteria. During the *in vivo* evolution experiment, I measured the light present in two specific regions of interest (ROI): the abdomen and the rectum (Fig 4.5). I did not observe ICC180 appearing in other locations in the mouse, indicating no obvious change in infection niche. Neither did I recover any non-bioluminescent colonies of *C. rodentium* ICC180, indicating that ICC180 maintained the bioluminescent phenotype throughout the *in vivo* evolution experiment.

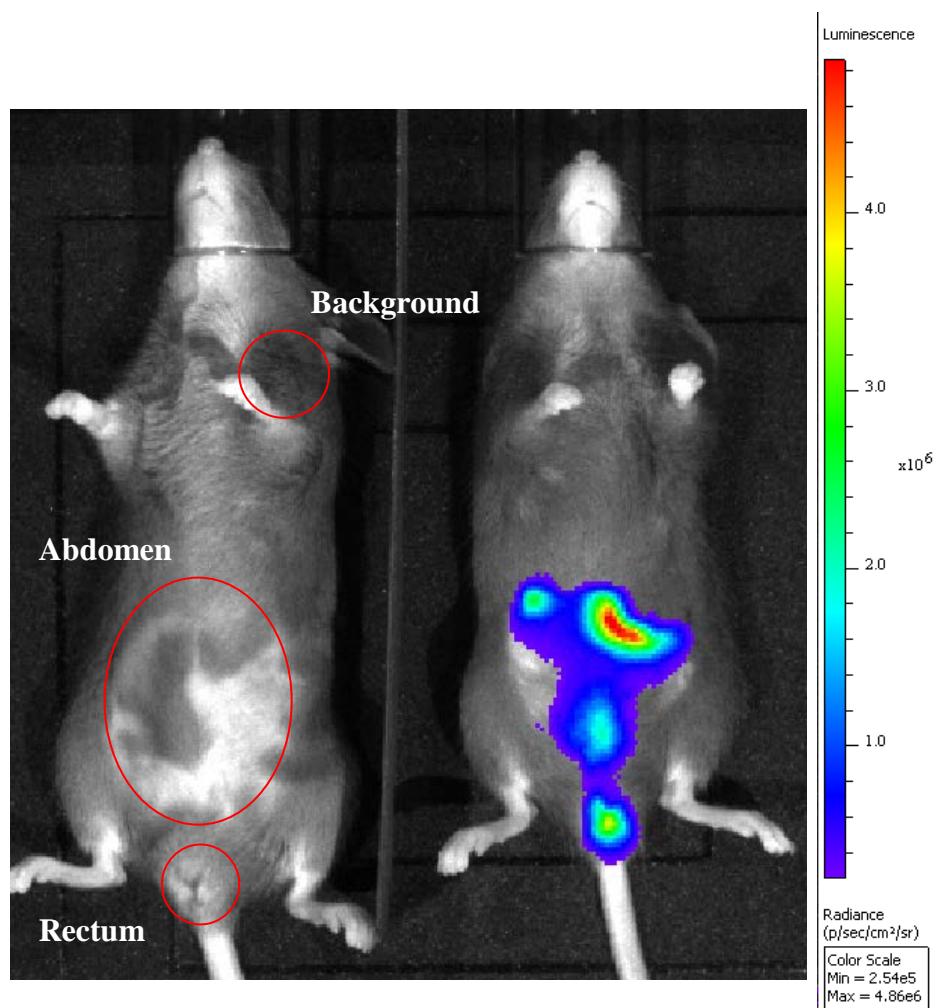


Figure 4.5. Representative images showing detectable bioluminescent *Citrobacter rodentium* from within anaesthetised mice and the regions of interest measured. Images were taken using the IVIS® Kinetic camera system (Perkin Elmer). Mice were anaesthetised using isoflurane then imaged at a binning of four over 1 minute using the Living Image software (Perkin Elmer). Bioluminescence (as photons second⁻¹ cm⁻² steradian⁻¹) is superimposed on the photograph as a pseudocolour image, with intensity of light going from blue (least intense) to red (most intense). The Living Image software was used to measure bioluminescence (as total flux [photons second⁻¹]) from specific regions of interest (ROI) corresponding to the abdomen and rectum. The upper shoulder was used to measure background luminescence.

The AUC of total flux measured from each region of interest for each lineage over the course of the *in vivo* evolution experiment is shown in Figure 4.6. A wide degree of variation in the AUC for each animal can be observed, following no particular pattern and appearing to indicate host variation rather than a change of *C. rodentium* infection. The overall trend over time is shown as linear regression lines, detailed in Table 4.3. The majority of lineages do not show any significant changes over time, indicating a stable detection of light from both ROI over the course of the experiment. Three lineages from the group receiving untreated water and food showed a statistically significant decrease in AUC of total flux: W1 showing a decrease of light detected from the rectum ($p=0.0398$), and W2 and W3 showing a decrease of light detected from the abdomen ($p=0.0265$ and $p=0.0239$, respectively).

Figure 4.6. Amount of light detected from each region of interest over the course of the evolution experiment does not change in the majority of lineages (overleaf). A-J show the area under the curve (AUC) of total flux (photons second⁻¹) for each infected animal for each region of interest (abdomen and rectum) over the course of the 5 month evolution experiment. A-E show lineages which received untreated drinking water and food, and F-J show lineages which received drinking water containing nalidixic acid. Filled circles indicate total flux AUC of the abdomen ROI, and filled triangles indicate total flux AUC of the rectum ROI. The dashed line indicates background luminescence. A linear regression trend line is shown to indicate the slope of total flux over the 5 month evolution experiment.

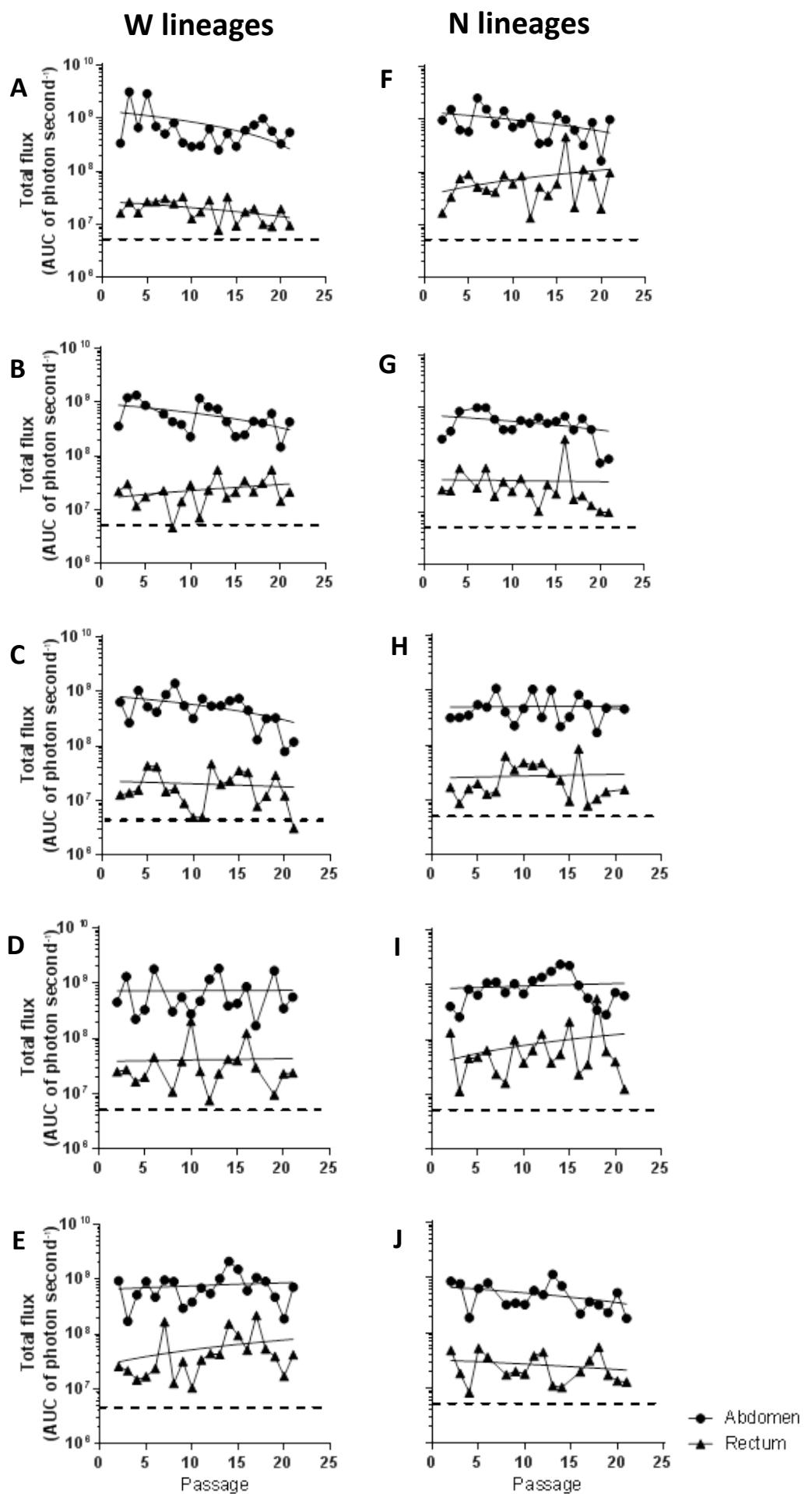


Table 4.3. Linear regression of total flux for abdominal and rectal regions of interest

	Abdominal region of interest		Rectal region of interest	
Lineage	Slope*	P value**	Slope*	P value**
W1	-5.36 x 10 ⁷ ± 2.90 x 10 ⁷	0.0810	-6.63 x 10 ⁵ ± 2.99 x 10 ⁵	0.0398
W2	-3.03 x 10 ⁷ ± 1.25 x 10 ⁷	0.0265	6.85 x 10 ⁵ ± 5.31 x 10 ⁵	0.2142
W3	-2.78 x 10 ⁷ ± 1.13 x 10 ⁷	0.0239	-2.50 x 10 ⁵ ± 5.41 x 10 ⁵	0.6497
W4	1.19 x 10 ⁶ ± 2.39 x 10 ⁷	0.9608	2.52 x 10 ⁵ ± 2.04 x 10 ⁶	0.9031
W5	1.08 x 10 ⁷ ± 1.82 x 10 ⁷	0.3495	2.61 x 10 ⁶ ± 2.24 x 10 ⁶	0.2596
N1	-3.92 x 10 ⁷ ± 1.98 x 10 ⁷	0.0633	3.69 x 10 ⁶ ± 3.69 x 10 ⁶	0.3302
N2	-1.75 x 10 ⁷ ± 9.64 x 10 ⁶	0.0876	-2.04 x 10 ⁵ ± 2.20 x 10 ⁶	0.9272
N3	1.35 x 10 ⁶ ± 1.20 x 10 ⁷	0.9121	2.08 x 10 ⁵ ± 9.06 x 10 ⁵	0.8212
N4	1.05 x 10 ⁷ ± 2.35 x 10 ⁷	0.6624	4.48 x 10 ⁶ ± 4.74 x 10 ⁶	0.3575
N5	-1.76 x 10 ⁷ ± 1.03 x 10 ⁷	0.1054	-5.52 x 10 ⁵ ± 6.27 x 10 ⁵	0.3917

* Linear regression of AUC values from total flux calculated for abdominal and rectal ROI for each infected animal over the 5 month evolution experiment. The value of the slope is presented ± the standard error.

** Statistically significant deviations from 0 (a flat line, indicating no change) are given in bold.

When the AUC of total flux detected from each ROI for mice from each lineage is pooled together, there is little difference between each of the lineages (Fig. 4.7). W4 has the highest overall values for AUC of light detected from the abdomen, significantly higher than almost all of the other lineages (N2, N3, N5, W1, W2, W3, W5; all p<0.01 [Kruskal-Wallis test with Dunn's post hoc multiple comparisons]). This difference is approximately ten fold, with W4 having a median AUC value of 3.910 x 10⁹ and the other strains having a median of 5.087 x 10⁸. Overall, there is no significant difference between the light detected from the abdomen from the two treatment groups (Fig 4.7C). For the AUC values calculated for light from the rectum, N1 is statistically different to N3, W1, W2, and W3 (p<0.05, p<0.001, p<0.05, p<0.001 respectively [Kruskal-Wallis test with Dunn's post hoc multiple comparisons]), and N4 is statistically different to W1 and W3 (p<0.01 and p<0.001 respectively [Kruskal-Wallis test with Dunn's post hoc multiple comparisons]). Overall, there is a significant difference between the light detected from the rectum of the two treatment groups (Fig. 4.7D), with more light detected from mice receiving drinking water with nalidixic acid, however this is a very modest difference which translates to approximately 0.1 of a log (median values of 3.381 x 10⁷ for the nalidixic acid-treated animals vs. 2.301 x 10⁷ for the untreated animals).

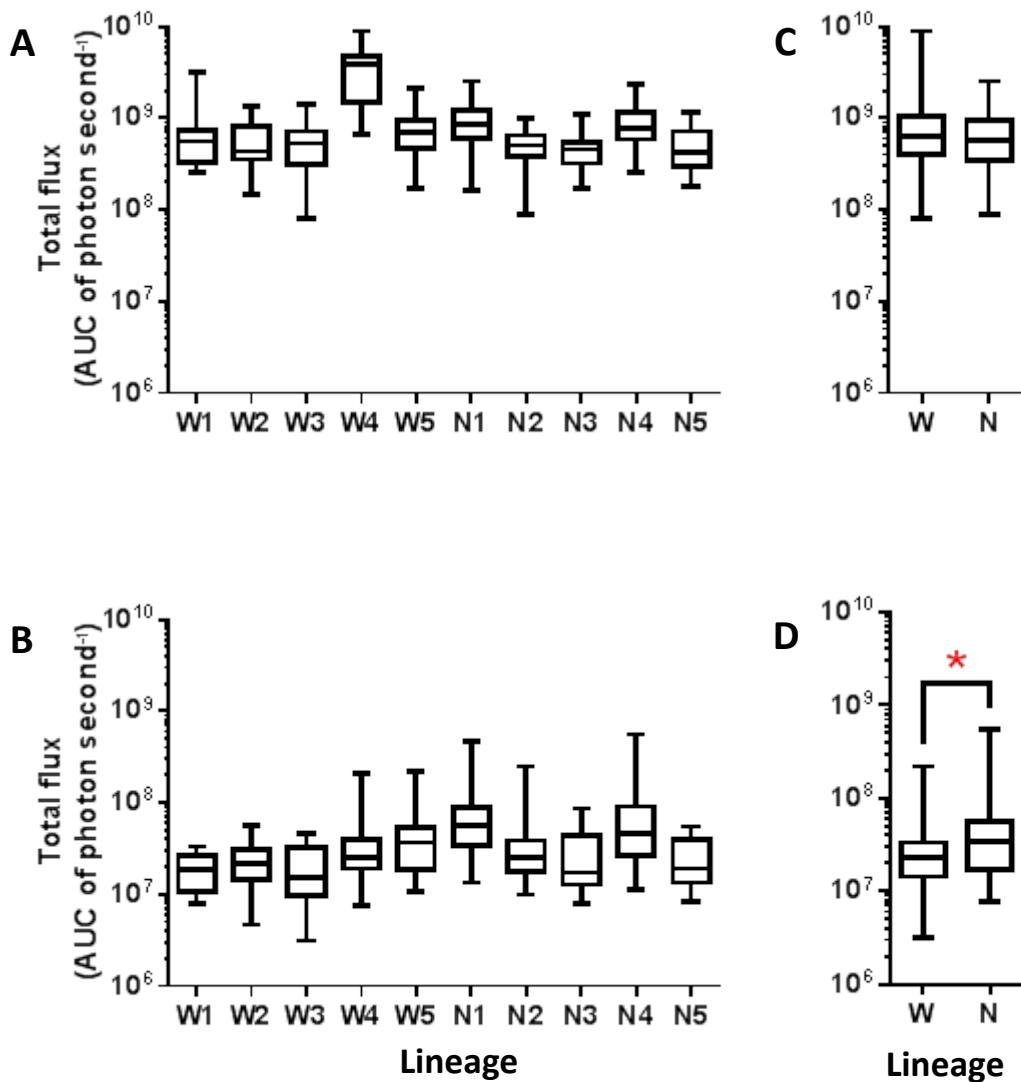
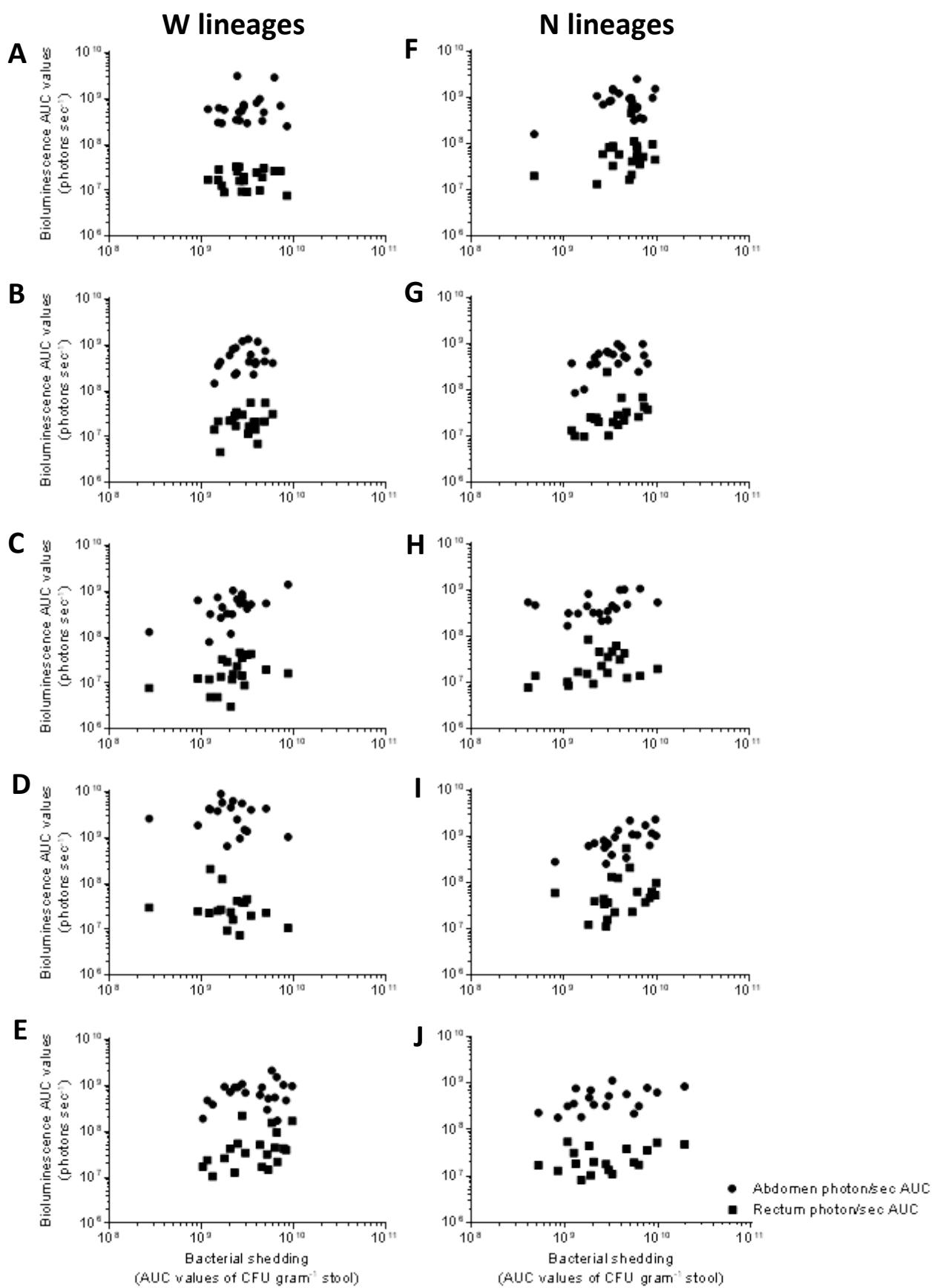


Figure 4.7. AUC values for light detected from the abdomen and rectum of mice vary between lineages. AUC of total flux (photons sec⁻¹) from defined regions of interest (ROI) from the abdomen (A) and rectum (B) of animals from each lineage of the course of the 5 month evolution experiment. Significant differences were observed for abdominal ROI between N2 and W4; N3 and W4; N5 and W4; W1 and W4; W2 and W4; W3 and W4; and W4 and W5 (all p<0.01, Kruskal-Wallis test with Dunn's post hoc multiple comparisons). Significant differences were observed for the rectal ROI between N1 and N3 (p<0.05), N1 and W1 (p<0.001), N1 and W2 (p<0.05), N1 and W3 (p<0.001), N4 and W1 (p<0.01), and N4 and W3 (p<0.001). Total AUC of abdomen (C) and rectum (D) ROI total flux for each treatment group. W denotes lineages which received untreated drinking water and food; N denotes lineages which received drinking water with nalidixic acid. (C) No statistically significant differences were observed for light measured from the abdominal ROI for the two treatment groups. (D) Lineages receiving untreated food and water have significantly lower light measured from the rectal ROI compared with lineages receiving nalidixic acid in their drinking water (Mann Whitney two-tailed test; p=0.0027).

Correlations of AUC of total flux from each ROI and CFU AUC of *C. rodentium* shed from each animal for each lineage is shown in Figure 4.8. Not all of the lineages show a correlation between AUC of light detected and AUC of bacteria shed, indicating a disconnect between bacteria shed in the stools and bacteria residing inside the hosts. Lineages N2 and W3 show a weak correlation between the rectum ROI and bacteria shed (0.6404 and 0.5594, respectively); and lineages N4, N5 and W3 show a weak correlation between abdomen ROI and bacteria shed (0.6346, 0.5191 and 0.5083, respectively). This observation is curious as the faecal matter would collect around the rectum and one would assume a stronger correlation between the rectum ROI and the bacteria shed in the stools.

Figure 4.8. The relationship between light detected from infected mice and the amount of *Citrobacter rodentium* recovered from stools varies between different evolutionary lineages (overleaf). A-J show the relationship between the AUC values for light detected from each ROI (abdomen and rectum) and the AUC values for CFU of *C. rodentium* shed from each animal. A-E shows lineages which received untreated drinking water and food; F-J show lineages which received drinking water with nalidixic acid. Filled circles indicate abdomen AUC of ROI data, and filled squares indicate rectum AUC of ROI data.



4.3.5 No detectable change in the severity of disease

Throughout the *in vivo* evolution experiment, I monitored all animals for general clinical signs including changes in activity, posture, coat quality, mucous production, and for signs of dehydration. All animals scored as normal, with animals rarely showing any symptoms other than loose, watery stools. I weighed each animal before and throughout the course of infection (Fig. 4.9A); there were no major deviations from the original body weight and all animals remained healthy and alert. Overall, animals receiving nalidixic acid in their drinking water weighed less (Fig. 4.9B) than those animals who did not receive antibiotics, however the median differences between the groups was very small (median calculated AUC of weight for control animals of 30.88 vs. 29.05 for animals receiving nalidixic acid).

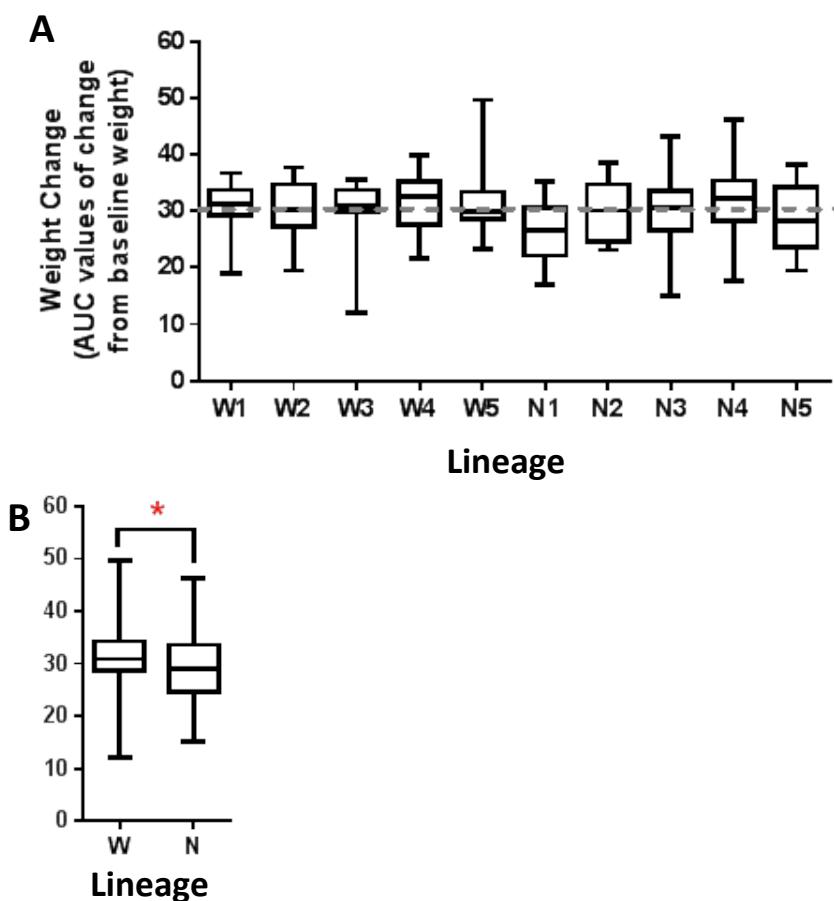


Figure 4.9. No change in disease severity as measured by change in mouse weight. (A) Area under the curve values of weight change from baseline for each animal from each lineage (W1-5 lineages receiving untreated food and water; N1-5 lineages receiving nalidixic acid treatment). The dashed grey line indicates the AUC of an animal with no weight change (AUC = 30). No significant difference between the lineages or from the no weight change value as measured using a Kruskal-Wallis test ($p=0.0911$) with Dunn's post hoc multiple comparison test. (B) Area under the curve values of weight change for each treatment group (W receiving untreated food and water; N receiving nalidixic acid). The red star denotes statistically significant differences (Mann Whitney test, $p<0.0284$).

4.4 Discussion

The *in vivo* evolution experiment described in this chapter is, to our knowledge, the longest and largest controlled experimental infectious disease transmission chain through a pathogen's natural mammalian host to date. The transmission chain of each of the *C. rodentium* lineages involves tightly controlled mouse-to-mouse exposure, ensuring no uncertainty as to the direction of infection or the chain of infection. Samples from each animal have been frozen and stored, and the bacteria can be revived at any time for analysis or reinfection purposes. The infection chain was followed for a five month period, by which point at least 20 passages had occurred, with the potential for *C. rodentium* isolated at the end of the infection chain to be revived and the *in vivo* evolution experiment to be continued. Due to a number of unknown variables in the *in vivo* experiment (for example, the bacterial growth rate within a mouse), the generation number of the *in vivo* experiment could not be estimated with any acceptable degree of confidence, and therefore a set time duration of five months was chosen for comparative purposes between the *in vivo* and *in vitro* evolution experiments.

4.4.1 High rates of successful transmission throughout the *in vivo* evolution experiment

I observed an overall high rate of transmission, owing to *C. rodentium*'s already documented high transmissibility, as shown by the ease with which the first recorded outbreaks spread within and presumably between animal houses^{78–80}. In the case of a transmission failure, the next animal in the infection chain was artificially infected with *C. rodentium* revived from frozen storage of stool samples from the previous point in the infection chain, and the effective passage number reduced by one to reflect the missed passage. Throughout the *in vivo* evolution experiment there were a low number of incidences of transmission failure, with no relationship between the number of transmission failure events and the treatment group. Each instance of artificial infection following transmission failure led to a successful infection, and transmission failure events showed no pattern of regularly and repeatedly occurring soon after each event, which in combination suggests that the transmission failure events were sporadic and more likely due to differences in the murine host and grooming habits than a reduction in infectiousness of a particular *C. rodentium* lineage.

4.4.2 No loss of bioluminescence was observed following *in vivo* evolution

I chose to carry out the *in vivo* evolution experiment using a bioluminescent strain of *C. rodentium* (ICC180) for multiple practical and ethical reasons. Importantly, the use of ICC180 allowed me to monitor the location of infecting *C. rodentium* within each animal in real time without the need to euthanise each animal to determine the location and number of bacteria within specific organs. Throughout the course of the *in vivo* experiment, no drastic change in *C. rodentium* location was

observed, with the infection predominantly occurring in the lower gastrointestinal tract. No evidence of extra-intestinal infection was observed, and no change in disease phenotype was observed. Diarrhoeal disease remained evident as a symptom of *C. rodentium* infection, with no noticeable increase or decrease in the severity of disease over the course of the experiment. Furthermore, I observed no loss of light production from *C. rodentium* ICC180, with strains detected and isolated from each animal consistently producing light and retaining resistance to kanamycin.

4.4.3 Presence of nalidixic acid associated with reduction in the number of *C. rodentium* shed from animals

Between the two treatment groups, some differences were observed. A trend of lower bacterial shedding was observed in 3/5 lineages (N2, N3, and N5) receiving nalidixic acid, while a trend of lower detectable light was observed in 3/5 lineages with untreated water (W1 with reduced light detected from the rectum, and W2 and W3 with reduced light detected from the abdomen). Interestingly, while lineages N2, N3 and N5 shed less bacteria, the light from within the animal remained stable, reflecting a stable bacterial burden for the infected animals. This suggests that, rather than experiencing a reduction in fitness, in these lineages fewer bacteria are sloughed off from the intestines leaving a greater number present in the infected animal. Conversely, while lineages W1, W2 and W3 produce less light, they do not have a trend of lower bacterial shedding, therefore suggesting a reduced total *C. rodentium* burden in these infected animals. One potential explanation for this observation is that the treatment group receiving antibiotic have reduced hurdles to overcome to initiate an infection, and therefore fewer bacteria are required to be shed from an animal in order to successfully infect its cage mate. This has parallels in human infections, whereby impairment of the commensal flora by antibiotic use is a major risk factor for acquiring a *Clostridium difficile* infection²⁰⁷.

It is important to note, however, that the slopes for these datasets have large standard errors indicating a high degree of variability in the data. Furthermore, the group with the highest levels of *C. rodentium* shedding was N1 from the treatment group receiving nalidixic acid, while W3, a lineage from the treatment group receiving untreated water, shed the lowest amount of *C. rodentium*, although the difference between the two lineages corresponds to only half the amount shed (median values of 5.377×10^9 for N1 and 2.179×10^9 for W3). Interestingly, the slopes for the AUC values from each animal throughout the *in vivo* evolution experiment for these lineages revealed no statistically significant downward or upward slope for these lineages, indicating that the high and low *C. rodentium* shed from these lineages was consistently higher or lower, but did not appear to be increasing or decreasing over time. This suggests that the observed high and low levels of *C. rodentium* shed from these lineages has

been a defining feature of these lineages throughout the five months, instead of a gained attribute over time. Increased shedding could be indicative of an increased overall bacterial burden, however this was not supported by the biophotonic imaging data.

4.4.5 Conclusions

Within the *in vivo* evolution experiment, there were no drastic differences between the two treatment groups (nalidixic acid vs. no antibiotic), as well as between the 10 different lineages. Interestingly, all lineages have a negative correlation of shedding over time even though in the majority this was not shown to be statistically significant. This may indicate adaptation of the bacteria to its host, but further studies are required to verify this.

To our knowledge, this is the longest and largest *in vivo* evolution experiment to date following natural transmission of a pathogen through its natural host. Due to time constraints, five months was the chosen time frame, however frozen stool samples mean that the experiment can be restarted and continued in the future. As samples from each mouse at the peak of infection have been stored, a complete “fossil record” is available indefinitely for future analysis. This resource is a gold mine for evolutionary biologists, available for further analysis which would provide knowledge beyond what is detailed in this thesis.

Chapter 5: *In vivo* assessment of *in vivo*-adapted *C. rodentium*

5.1 Introduction

One of the first logical steps to ascertain any changes in fitness following evolution to a particular condition is to place the evolved bacteria back into its original environment and perform experiments to determine if any changes have occurred. To do this, we wanted to measure the infection dynamics of the *in vivo*-adapted *C. rodentium* strains in comparison with the ancestral ICC180 strain, both in animals infected with each strain individually, as well as in competition with ICC180. As ICC180 and the *in vivo*-adapted strains are both bioluminescent and resistant to kanamycin and nalidixic acid, they are indistinguishable and therefore cannot be directly competed with each other. Therefore, the *in vivo*-adapted strains were competed against ICC169, the non-bioluminescent, kanamycin sensitive, nalidixic acid resistant parent of ICC180⁹³, and the results compared with the benchmark study in which ICC169 was competed against the ancestral ICC180 strain, as presented in Chapter 3¹⁵³.

As successful transmission from one mouse to another was a key step in the evolution experiment, and one of the easily identifiable aspects on which selection is likely to have occurred, mice co-infected with the competing strains were exposed to uninfected mice in order to determine infectiousness of the *in vivo*-adapted strains when competed with ICC169. Such a competition experiment is a way to determine if:

- 1) The *in vivo*-adapted strains are more readily excreted and shed from the animals, as measured by the ratio of *in vivo*-adapted *C. rodentium* and ICC169 recovered from stools
- 2) The *in vivo*-adapted strains are more proficient at transmitting to a new host, as determined by the numbers of animals successfully infected with the *in vivo*-adapted *C. rodentium* strains following exposure to contaminated cages.

The work reported in this chapter covers both single infections of mice with the *in vivo*-adapted *C. rodentium* strains as well as head-to-head competition of these strains with ICC169. The *in vivo*-adapted strains are named based on the environment to which they evolved, and the passage number they are taken from (passage 20). Hence, the strains used which were isolated from mice receiving no antibiotic are referred to as “W” (Water) strains (W₁_{P20}, W₂_{P20}, W₃_{P20}, W₄_{P20} and W₅_{P20}), while those which have been isolated from mice receiving nalidixic acid treatment in their drinking water are referred to as “N” (Nal) strains (N₁_{P20}, N₂_{P20}, N₃_{P20}, N₄_{P20} and N₅_{P20}).

5.2 Results

5.2.1 *In vivo* infection dynamics of the *in vivo*-adapted *C. rodentium* strains after oral gavage of mice

To evaluate the infection dynamics of the *in vivo*-adapted and the ancestral ICC180 *C. rodentium* strains in single infections, groups of six animals were housed in HEPA filtered cages and infected via oral gavage with either ICC180 or bacteria isolated from the 20th passage of each *in vivo*-adapted lineage. To recapitulate the environment in which each lineage had adapted, mice orally gavaged with bacteria from lineages N1 ($N1_{P20}$), N2 ($N2_{P20}$), N3 ($N3_{P20}$), N4 ($N4_{P20}$) or N5 ($N5_{P20}$) were administered nalidixic acid in their drinking water 24 hours prior to infection. Mice orally gavaged with bacteria from lineages W1 ($W1_{P20}$), W2 ($W2_{P20}$), W3 ($W3_{P20}$), W4 ($W4_{P20}$) and W5 ($W5_{P20}$) were given untreated drinking water. For comparison purposes, groups of nalidixic acid-treated and untreated mice were orally gavaged with the ancestral ICC180 strain. The animals were followed over the course of infection with their stools sampled and plated onto selective media (LB agar supplemented with kanamycin [$50 \mu\text{g ml}^{-1}$]) to determine the numbers of *C. rodentium* shed from each animal.

The levels of *C. rodentium* shed from the animals over the course of infection are shown in Figure 5.1, with data presented as median values for the six animals (Fig. 5.1A/B). Untreated animals infected with ICC180 or the “W” strains are shown in Figure 5.1A, and treated animals infected with ICC180 or the “N” strains are shown in Figure 5.1B. The overall pattern for the numbers of bacteria shed by mice (given as CFU g^{-1} stool) infected with each *in vivo*-adapted strain is similar to the ancestral ICC180 strain, with each strain reaching peak shedding at day seven post-infection and being cleared by the animals after two weeks of infection.

I calculated AUC values from the bacterial shedding data obtained throughout the infection; these are presented in Fig. 5.1C/D. Of the mice infected with the “W” *in vivo*-adapted strains, two groups ($W1_{P20}$ and $W5_{P20}$) shed significantly higher numbers of *C. rodentium*, compared to mice infected with the ancestral strain ($p=0.0047$ and $p=0.0295$, respectively [Kruskal-Wallis with Dunn’s post hoc multiple comparison test]) (Fig. 5.1C). In contrast, mice infected with one of the *in vivo*-adapted strains evolved in the presence of antibiotic ($N3_{P20}$) shed significantly less bacteria than mice infected with the ancestral ICC180 strain ($p=0.0373$, Kruskal-Wallis with Dunn’s post hoc multiple comparison test) (Fig. 5.1D).

Interestingly, the presence of nalidixic acid resulted in a significant increase in the number of bacteria shed from mice infected with ancestral *C. rodentium* ICC180 (median AUC values of 1.500×10^8 compared to 5.843×10^8 ; $p=0.0087$ [Mann Whitney test]). However when the *in vivo*-adapted strains were combined, there was no significant difference between AUC values of animals infected with “W” or “N” strains ($p=0.1091$ [Mann Whitney test]). This lack of difference is perhaps a reflection of the improvements in shedding for W1_{P20} and W5_{P20} in one treatment group, and the overall higher shedding from mice receiving nalidixic acid.

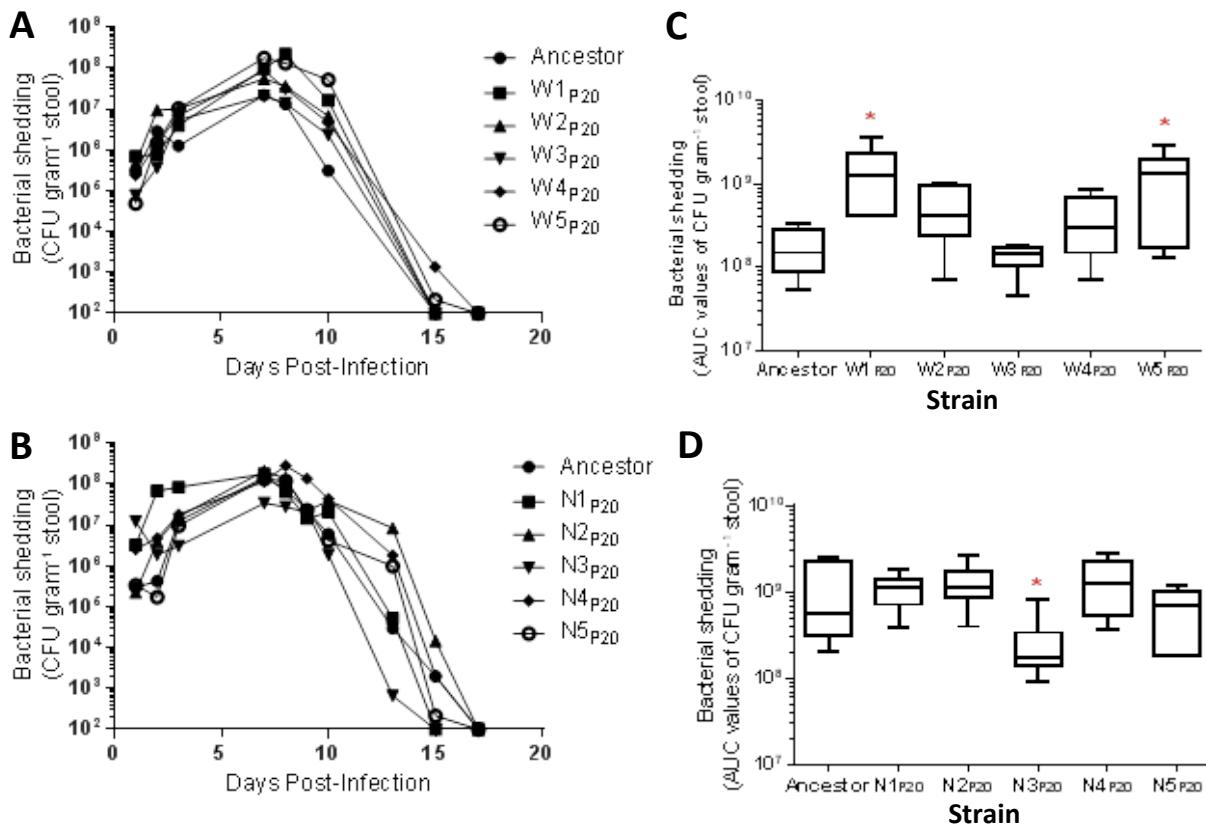


Figure 5.1. The *in vivo*-adapted *C. rodentium* strains W1 and W5 are shed at higher numbers compared with mice infected with the ancestral ICC180 strain. Groups of six animals were infected by oral gavage with either the strains evolved in the absence of antibiotics (W1_{P20}-W5_{P20}) (A/C) or in the presence of nalidixic acid (N1_{P20}-N5_{P20}) (B/D). Infection dynamics are presented as median bacterial numbers shed from the 6 animals at each time point (A/B) or as area under curve (AUC) values (C/D). A Kruskal-Wallis two-tailed test with Dunn's post hoc multiple comparison test revealed statistically significant differences between the ancestor and W1_{P20} ($p=0.0047$), the ancestor and W5_{P20} ($p=0.0295$), and between the ancestor and N3_{P20} ($p=0.0373$).

5.2.2 The *in vivo*-adapted *C. rodentium* strains are differentially shed from animals co-infected with ICC169

In order to assess the competitive advantage or disadvantage of the *in vivo*-adapted *C. rodentium* strains compared with the ancestral ICC180 strain, I designed a co-infection competition experiment (Fig. 5.2). Groups of six animals were housed in HEPA filtered cages and fed autoclaved food and water. I artificially infected the animals by oral gavage with a 1:1 mix of *C. rodentium* ICC169 and each of the *in vivo*-adapted strains (Fig. 5.2A). If the *in vivo*-adapted strain had evolved in the presence of low dose nalidixic acid, then the animals in the co-infection experiments were also exposed to the same antibiotic treatment for at least 24 hours prior to infection. I measured the total *C. rodentium* numbers and the ratio of each *C. rodentium* strain shed in the stools of infected animals by plating onto LB agar containing either kanamycin or nalidixic acid. At the peak of infection (6 days post infection), I placed the infected animals into a fresh cage (Fig. 5.2B). After 24 hours in the cage (7 days post infection), I moved these animals into another fresh cage for the remainder of the experiment (Fig. 5.2C), and placed six new naive animals into the cages contaminated by the infected animals (Fig. 5.2D). In this way, the new animals were exposed solely to environmental contaminants in the form of infectious faecal matter, and the potential variability of grooming practices of the animals was removed. I measured the quantity and ratio of *C. rodentium* species present in the stools of these animals to determine the effectiveness of *C. rodentium* transmission.

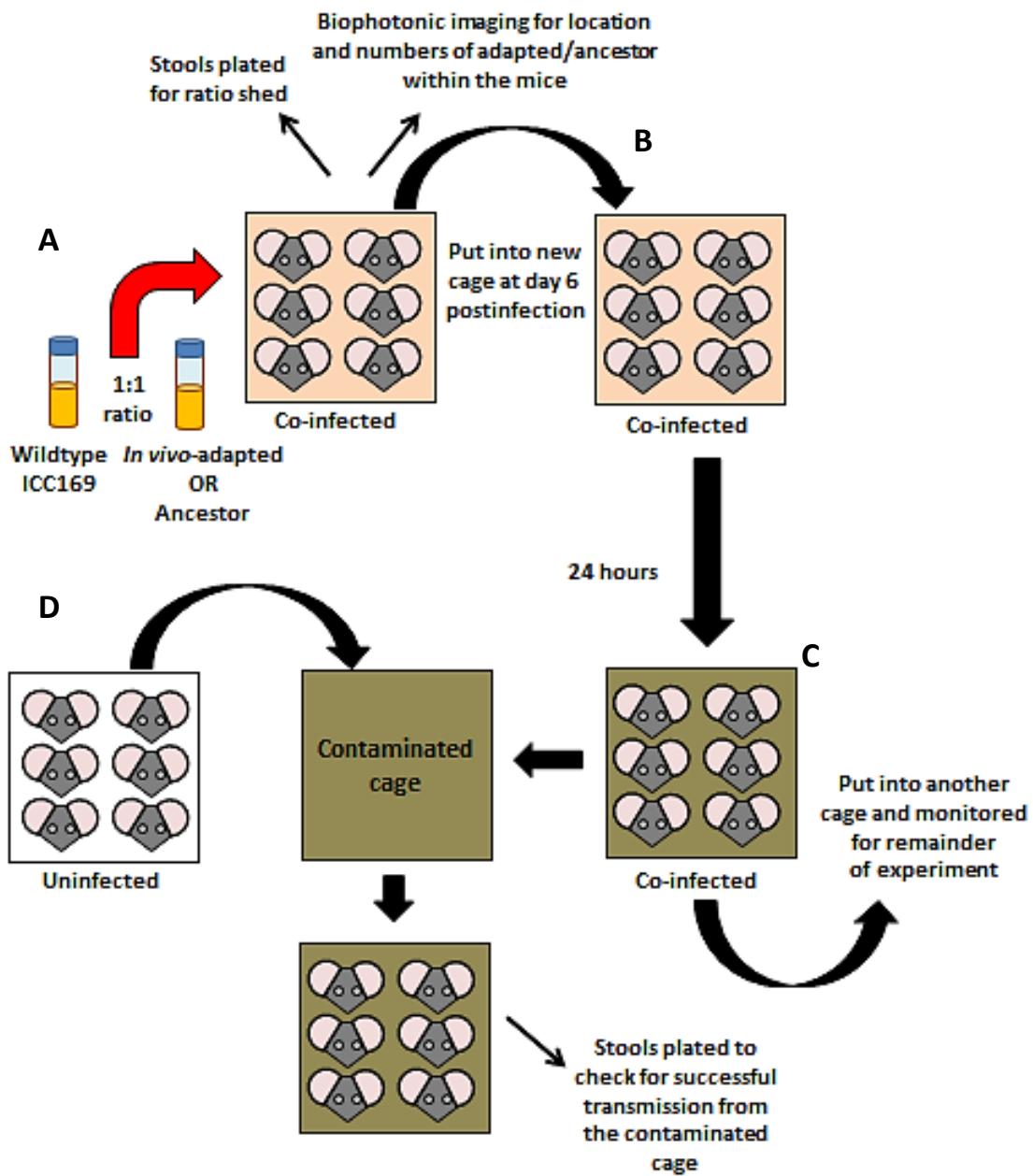


Figure 5.2. Schematic of the co-infection experiments.

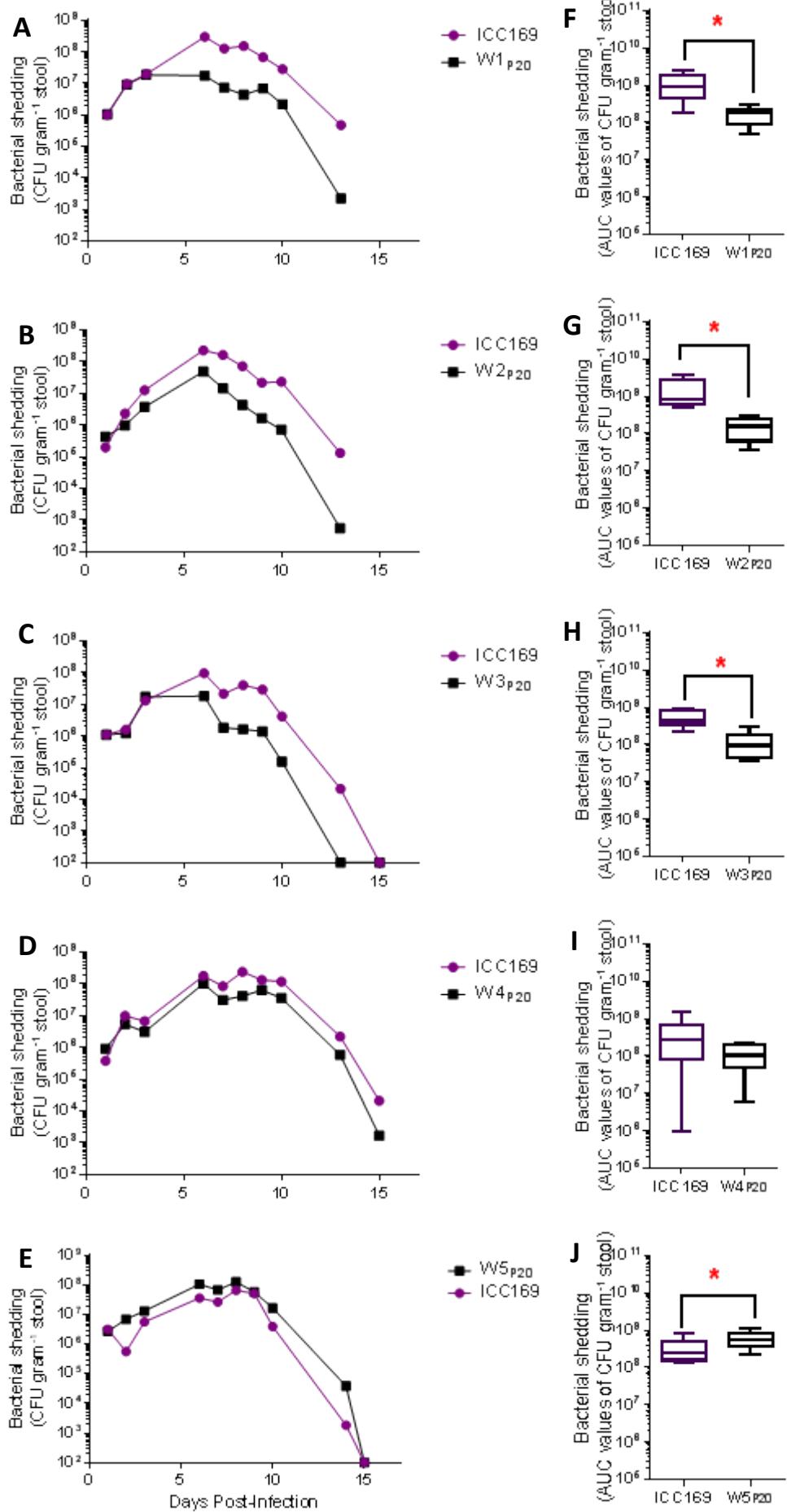
Representative data of the *in vivo*-adapted strains when competed against ICC169 are shown in Fig. 5.3 (“W” strains) and Fig 5.4 (“N” strains), presented as the median value for the viable counts of each of the competing strains shed (Fig. 5.3A-E and Fig. 5.4A-E) and calculated Area Under Curve (AUC) values from individual animals (Fig. 5.3F-I and Fig. 5.4F-I). To determine if there were any significant differences between the shedding ratios of each strain when compared with ICC169, I performed a two-tailed Wilcoxon Matched Paired Rank test. It is important to note that when the ancestral ICC180 strain is competed with ICC169, ICC180 is shed at significantly lower levels than ICC169 (Chapter 3,¹⁷²). Three of the *in vivo*-adapted strains showed an improvement in shedding compared with ICC180: two strains (W4_{P20} and N4_{P20}) are shed at equivalent levels to ICC169 (Fig 5.3I and Fig 5.4I), while strain W5_{P20} is shed at significantly higher levels than ICC169, (median AUC values of 5.72×10^8 for W5_{P20} compared to 2.45×10^8 for ICC169) (Fig 5.3J).

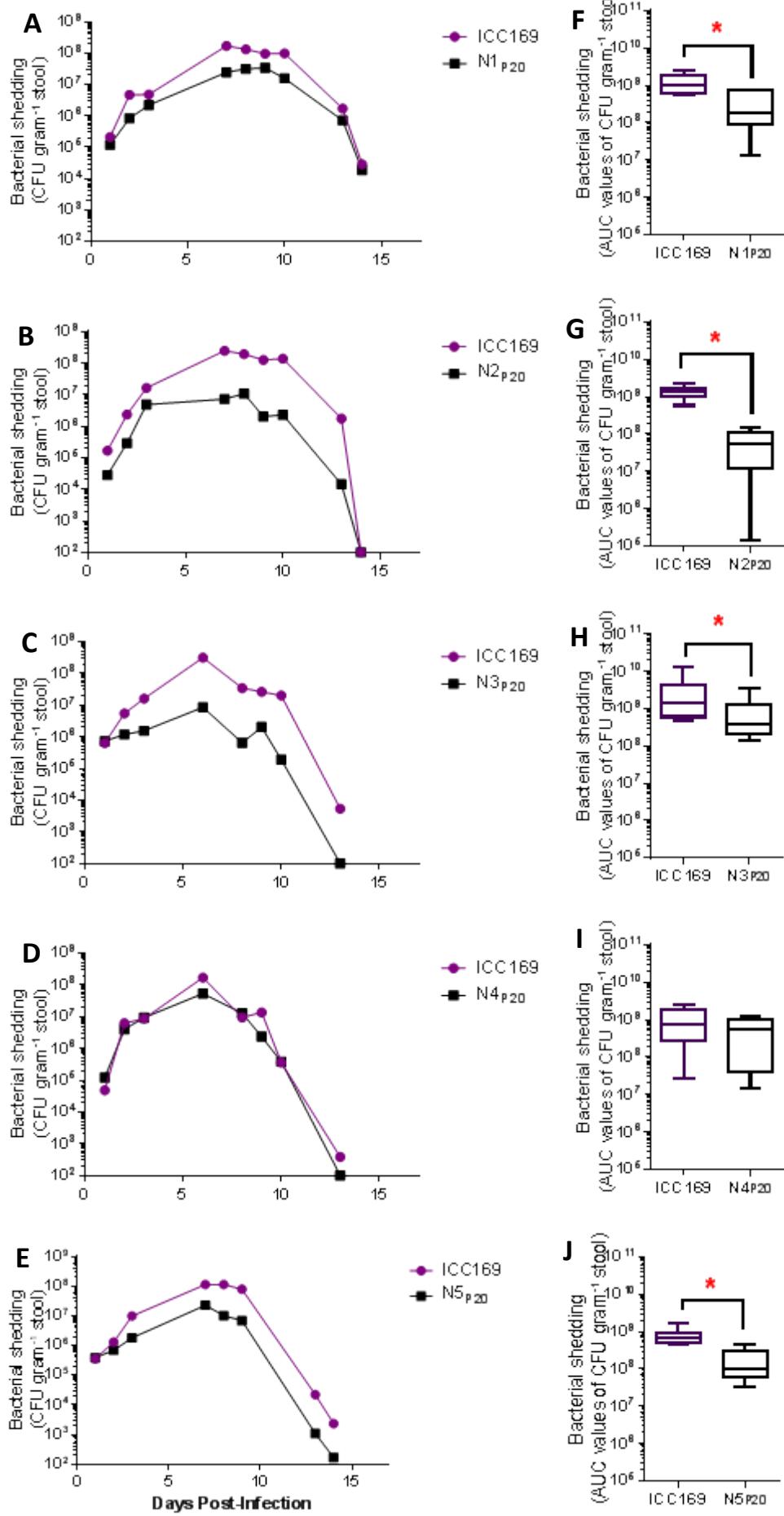
Competition indices for each of the *in vivo*-adapted strains are shown in Figure 5.5. The further a line representing one of the competing strains strays from 1 (the dashed line), the greater the competitive advantage or disadvantage, depending on if the line strays above or below the dashed line.

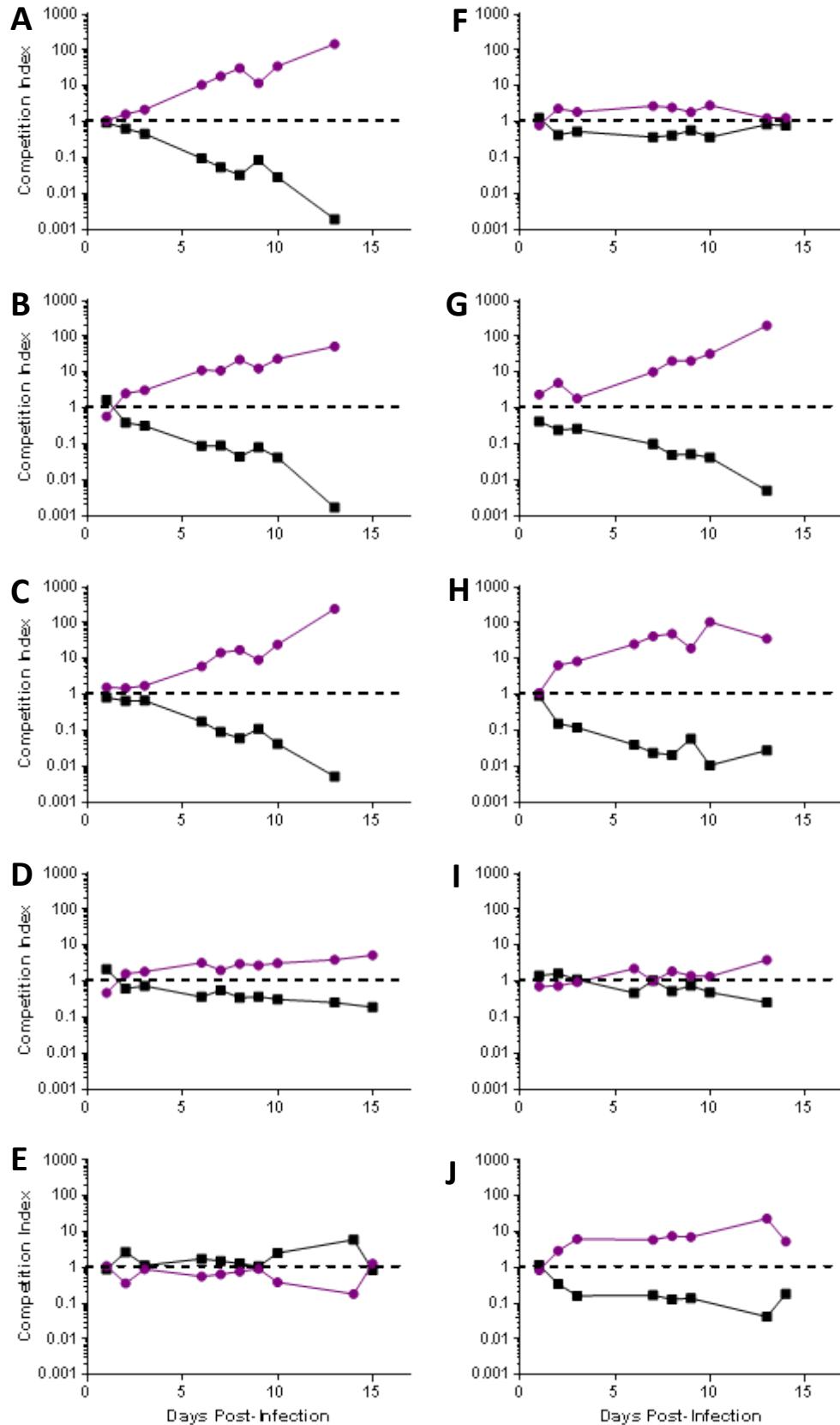
Figure 5.3. Competition of the *in vivo*-adapted *C. rodentium* strains evolved in the absence of antibiotics with ICC169 reveals fitness advantages in some lineages (overleaf). Groups of six animals were infected by oral gavage with a 1:1 ratio of ICC169 and each of the *in vivo*-adapted strains (W1_{P20}-W5_{P20}). Data is presented as median bacterial numbers shed in the stools of the animals over the course of the infection (A-E) and as calculated area under curve (AUC) values (F-J). A red star symbol denotes statistically significant differences ($p<0.05$, two-tailed Wilcoxon Matched Pairs signed rank test).

Figure 5.4. Competition of the *in vivo*-adapted *C. rodentium* strains evolved in the presence of nalidixic acid with ICC169 reveals fitness advantages in some lineages (overleaf). Groups of six animals were infected by oral gavage with a 1:1 ratio of ICC169 and each of the *in vivo*-adapted strains (N1_{P20}-N5_{P20}). Data is presented as median bacterial numbers shed in the stools of the animals over the course of the infection (A-E) and as calculated area under curve (AUC) values (F-J). A red star symbol denotes statistically significant differences ($p<0.05$, two-tailed Wilcoxon Matched Pairs signed rank test).

Figure 5.5. Competition indices reveal varying fitness of *in vivo*-adapted *C. rodentium* when competed head-to-head with ICC169 in mice (overleaf). Competition indices (CI) were calculated as follows: $CI = [\text{strain of interest output}/\text{competing strain output}]/[\text{strain of interest input}/\text{competing strain input}]$ ^{178,179}. Purple lines and filled circles indicate the CI of ICC169, while black lines and filled squares indicate the CI of the *in vivo*-adapted strain. A-E show the CI of *in vivo*-adapted strains which evolved in mice receiving untreated food and water (W1_{P20}-W5_{P20}); F-J show the CI of *in vivo*-adapted strains which evolved in mice receiving low dose nalidixic acid (N1_{P20}-N5_{P20}).







5.2.3 Differences in bacterial burden as measured by bioluminescence readings in co-infected animals

As the *in vivo*-adapted strains and ancestor ICC180 are bioluminescent, I performed IVIS imaging of the animals twice weekly to determine the quantity and location of the bacteria during the course of infection. The area under curve (AUC) values of the amount of light detected from each region of interest from each animal over the course of the competition experiments are depicted in Figure 5.6. Mice infected with strains W5_{P20} and N4_{P20} both have significantly higher levels of light measured from the abdominal ROI (Fig. 5.6A), while mice infected with W5_{P20} also have significantly higher levels of light measured from the rectal ROI (Fig. 5.6B).

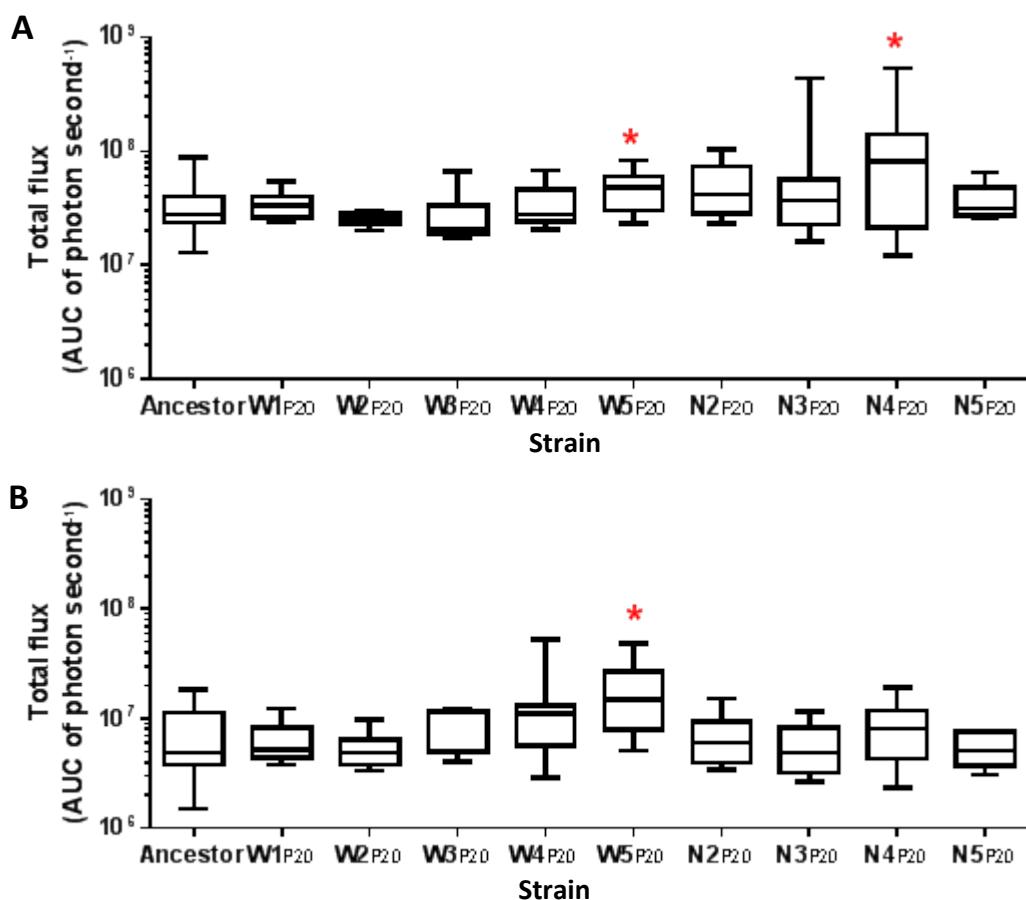


Figure 5.6. Higher levels of light detected from mice infected with two *in vitro*-adapted strains when competed with ICC169. Calculated area under curve (AUC) values of total flux (photons second⁻¹) for each infected animal for each region of interest, (A) abdomen and (B) rectum, when co-infected with either ICC180 or an *in vivo*-adapted strain and ICC169. A Kruskal-Wallis test revealed significant differences between the strains (abdomen: p=0.0215; rectum: p=0.0093). Dunn's post hoc multiple comparison test between the ancestral strain and W5_{P20} and the ancestral strain and N4_{P20} revealed differences for the abdomen ROI (W5_{P20}: p=0.0496; N4_{P20}: p=0.0075), and between the ancestral strain and W5_{P20} in the rectum ROI (p=0.0117). Combined data from all experiments is shown. The N1_{P20} strain was not tested due to technical difficulties.

5.2.4 The *in vivo*-adapted strains show changes in transmissibility compared to the ancestral ICC180 strain

As discussed previously in Chapter 3, the ancestral strain ICC180 has a competitive disadvantage when competed directly with ICC169, being consistently shed 1-2 logs lower in the stools of infected animals¹⁷². At the peak of infection, when the transmission experiments were performed, ICC180 is shed at approximately one log lower than ICC169, with the contaminated cage containing at least 90% ICC169 and less than 10% ICC180 (Table 5.1). Naïve mice exposed to this ratio of bacteria in their cage environment all become infected with ICC169 (6/6 animals), with none of these animals becoming infected with ICC180. This provides benchmark data for comparison of the *in vivo*-adapted strains; any strain which can transmit to the new animals in this model will have an improvement in transmissibility compared to the ancestral ICC180 strain.

Enumeration of *C. rodentium* from shed stools in the contaminated cage allows us to determine the total numbers and ratio of each strain (*in vivo*-adapted or ICC169) at the peak of infection and point of transmission (Table 5.1.). Although previous results have shown that some strains are shed at higher numbers (Fig. 5.3J/I and Fig. 5.4D), there are generally more ICC169 present in the contaminated cage environment than ICC180 or the *in vivo*-adapted strains at the point of transmission (Table 5.1). With the exception of W5_{P20} (which constitutes 72% of all detected *C. rodentium* at the peak of infection), all of the *in vivo*-adapted strains are below 50% of the total *C. rodentium* present in the stools within the cage. The next highest values are W3_{P20}, W4_{P20}, and N4_{P20} (19%, 23%, and 24%, respectively). N2_{P20} and N3_{P20} have the lowest detected percentage at 4% and 3%, respectively, which is a lower percentage than present in the ICC169/ICC180 competition model. W1_{P20} and W2_{P20} are detected at 8% and 10%, which are the closest values to those shed by the ancestral ICC180.

Table 5.1 also shows the ability of the *in vivo*-adapted strains to transmit to naïve animals when in competition with ICC169. Transmission is determined by the number of naïve animals which have a detectable *C. rodentium* infection following a week exposure to the contaminated cage. A colour code is used to determine the percentage of each strain; green indicating a high percentage (to 100%) and red indicating a low percentage (to 0%). In the presence of ICC169, the ancestral ICC180 strain is unable to transmit to the new animals (Table 5.1). This remains the case for two of the *in vivo*-adapted strains (W3_{P20} and N3_{P20}). All other *in vivo*-adapted strains show varying improvements in transmissibility, successfully infecting either some or all of the animals. All of the transmission animals are successfully infected with ICC169 in this model, with the notable exception of animals exposed to cages contaminated with stool from mice co-infected with ICC169 and N4_{P20}. The N4_{P20} strain successfully transmitted to all the naïve animals, however ICC169 did not.

Table 5.1. Changes in shedding at the peak of infection and varying transmission improvements in *in vivo*-adapted *C. rodentium* when competed with ICC169.

Competing Strain *	ICC169 **	Adapted **	% ICC169 ***	% Adapted ***	ICC169 Transmitted ****	Adapted Transmitted ****
Ancestor	1.40 x 10 ⁸	1.13 x 10 ⁷	93%	7%	100%	0%
W1 _{P20}	1.42 x 10 ⁸	1.20 x 10 ⁷	92%	8%	100%	84%
W2 _{P20}	1.83 x 10 ⁸	1.95 x 10 ⁷	90%	10%	100%	34%
W3 _{P20}	4.57 x 10 ⁷	1.05 x 10 ⁷	81%	19%	100%	0%
W4 _{P20}	1.50 x 10 ⁸	4.54 x 10 ⁷	77%	23%	100%	100%
W5 _{P20}	3.17 x 10 ⁷	8.00 x 10 ⁷	28%	72%	100%	100%
N1 _{P20}	1.58 x 10 ⁸	2.73 x 10 ⁷	85%	15%	100%	50%
N2 _{P20}	2.00 x 10 ⁸	7.42 x 10 ⁶	96%	4%	100%	100%
N3 _{P20}	3.17 x 10 ⁸	8.60 x 10 ⁶	97%	3%	100%	0%
N4 _{P20}	1.72 x 10 ⁸	5.35 x 10 ⁷	76%	24%	0%	100%
N5 _{P20}	1.16 x 10 ⁸	1.66 x 10 ⁷	87%	13%	100%	100%

* The strain competing in a 1:1 ratio with the non-bioluminescent ICC169

** The median CFU gram stool⁻¹ values of each strain shed at the point of transmission (days 6 and 7 post-infection; when the co-infected animals are housed in the new cage which will be used as the source of infection for the naïve animals)

*** The percentage of total shed which is ICC169 or the adapted strain. A colour code ranging from red (lowest) to green (highest) is used to indicate the percentage shed which is the *in vivo*-adapted strain.

**** The percentage of each strain which was present in the newly infected animals after 3 days exposure to the contaminated cage. A colour code shows the percentage of animals infected; green indicating a high percentage (100%) and red indicating a low percentage (0%).

5.3 Discussion

The data presented in this chapter provide strong evidence of adaptive evolution of the *in vivo*-adapted strains to fitness in the host and the improved ability to successfully establish new infections. First, I investigated the ability of the *in vivo*-adapted strains to infect and be shed from animals in single-infections, and next the ability of the *in vivo*-adapted strains to infect, be shed, and to transmit to a new host when being competed with ICC169. I have established that ICC169 is more fit in a co-infection model and better able to transmit than ICC180, the ancestor of all of the *in vivo*-adapted strains, and thus the game was already stacked for ICC169 to win. A summary of the observed changes of the *in vivo*-adapted *C. rodentium* when infecting mice, both in single infection and in direct competition with ICC169, can be found in Table 5.2. That many of the *in vivo*-adapted strains were able to successfully transmit to naïve animals even when present at lower levels than ICC169, and, in one case, able to transmit to all naive animals and prevent ICC169 from establishing an infection is evidence of observable adaptation to the mouse host. Only one strain (N3_{P20}) showed no evidence of improvement in terms of bacterial shedding from infected mice or the ability to transmit to naïve mice. Rather N3_{P20} appears to have a disadvantage, showing no improvements in competitive fitness or transmission, and being shed at significantly lower levels during single infection of mice.

5.3.1 Adaptation to mice results in changes in the numbers of *C. rodentium* shed in single infections in 3/10 lineages

Two out of five *in vivo*-adapted strains from mice receiving untreated water (W1_{P20} and W5_{P20}) were recovered from shed stools significantly more than the ancestral ICC180 strain was (Fig. 5.1). In contrast, one of the *in vivo*-adapted strains from mice receiving low-dose nalidixic acid in the drinking water (N3_{P20}) was recovered from shed stools at significantly lower levels than the ancestor (Fig. 5.1). This may be due to the effect of the nalidixic acid antibiotic on the presence of commensal bacteria; the strains which evolved in the absence of antibiotic might need higher infectious loads to successfully infect the next animal, due to the lack of antibiotic impacting the relative abundance of different commensal bacteria. Similarly, during the course of infection, competition from commensals might require the strains to increase their bacterial burden of the animal, thereby increasing the burden of *C. rodentium* in the animal leading to increased amounts of *C. rodentium* being shed in the stools: indeed, for one of these strains (W5_{P20}) a higher amount of light was detected from within the abdomen of mice co-infected with W5_{P20} and ICC169 compared with those infected with the ancestor and ICC169 (Fig. 5.6).

5.3.2 Shedding and transmissibility of *in vivo*-adapted *C. rodentium* when in competition with ICC169 are not necessarily linked

At the peak of infection, I observed that all of the *in vivo*-adapted strains are shed at below 50% of the total *C. rodentium* detected, with the lone exception of W5_{P20} which constitutes 72% of all detected *C. rodentium* (Table 5.1). Using competition of the ancestral ICC180 strain with ICC169 as a benchmark, where being shed at 7% results in a complete failure to transmit to naïve animals, one would expect that all strains except W5_{P20} would struggle to transmit successfully.

The data shows that two of the *in vivo*-adapted strains (N2_{P20} and N3_{P20}) were present at a lower percentage than the ancestral ICC180 strain at the point of transmission (4% and 3%, respectively). The N3_{P20} strain performed as expected, with no transmission to the naïve animals. In contrast, the N2_{P20} strain successfully infected 100% of the naïve animals despite being shed at a percentage far lower than the ancestor. This indicates that N2_{P20} has an improved ability to transmit, but that this is not due to an increase in shedding from infected animals.

Similarly, the W1_{P20} and W2_{P20} strains were present at similar levels to the ancestor, at 8% (1.20×10^7 CFU gram⁻¹ of stool) and 10% (1.95×10^7 CFU gram⁻¹ stool), respectively, and yet successfully transmitted to the naïve animals at modest levels (84% and 34%, respectively), evidence of an improvement in the ability to transmit in comparison to the ancestor. The N1_{P20} and N5_{P20} strains, present at a slightly higher percentage than the ancestor at 15% and 13%, respectively, were able to transmit to 50% and 100% of the naïve animals, indicating an improved ability to transmit that may or may not be linked to a higher overall infecting dose. W3_{P20} is shed at a higher percentage than the ancestral strain (19%), however is similarly unable to transmit to the naïve animals, indicating that shedding of higher bacterial numbers in the stool does not necessarily translate to an improved ability to transmit to naïve animals. W4_{P20} is also shed at a higher percentage than the ancestral strain (23%), however in this case the increased bacterial shedding does translate to an improvement in transmission, with W4_{P20} being able to transmit to 100% of naïve animals.

I observed that W5_{P20} is shed at considerably higher percentage than the ancestral strain, and indeed even higher than the competing ICC169 strain, at 72%. As W5_{P20} is shed at such high rates, this is reflected in the transmission data whereby W5_{P20} successfully infected 100% of the naïve animals. However, it is difficult to discern whether this improvement is due to an improvement in transmissibility and ability to colonise a new host, or if it is due simply to the overwhelmingly high numbers of W5 that the naïve animals were exposed to.

The N4_{P20} strain was present at 24% (5.35×10^7 CFU gram⁻¹ stool), a higher percentage than the ancestral strain (7% [1.13×10^7 CFU gram⁻¹ stool]), and was able to successfully infect 100% of the naïve animals. What is most interesting, and unique to N4_{P20}, is that in the presence of N4_{P20} the

competing ICC169 strain, which made up the majority of cultured *C. rodentium* from shed stools at 76% (1.72×10^8 CFU gram⁻¹ stool), was unable to transmit to the naïve animals. ICC169 has been shown to be capable of infecting naïve animals at percentages as low as 28% (3.17×10^7 CFU gram⁻¹ stool), as seen in the W5_{P20} competition, and yet in the presence of N4 the ICC169 strain was completely unable to infect any of the naïve animals. This could be due to a greatly improved ability for N4_{P20} to transmit and outcompete the wildtype ICC169 in the newly infected animals, or it could be that a more active competition is occurring and actively preventing ICC169 from successfully infecting the animals when N4_{P20} is present.

Overall, these data indicate that what may be necessary for improving the ability of *C. rodentium* to transmit to naïve mice (that is, colonisation of an animal and the ability to remain infectious in stools) is not the same as what is necessary to compete with the ancestor inside a host (ability to grow), and indicates that each of the *in vivo*-adapted strains which were successfully able to transmit in this model achieved this goal via different routes.

5.3.3 Treatment of mice with nalidixic acid does not appear to influence the trajectory of *in vivo* adaptation as measured by transmissibility and bacterial shedding

The data indicates that eight out of ten of the *in vivo*-adapted strains have a transmission advantage compared to the ancestral ICC180 strain, when in competition with the non-bioluminescent ICC169 strain, indicating a strong selection for improved transmissibility during the *in vivo* evolution experiment. 3/5 of the *in vivo*-adapted strains from each of the treatment groups (those adapted in mice being given nalidixic acid and those who were not) showed an increase in the percentage of the adapted strain shed at the peak of infection when competed with ICC169. 2/5 of the *in vivo*-adapted strains from the mice not receiving antibiotics were shed at percentages comparable to the ancestor, whereas the remaining 2/5 of the *in vivo*-adapted strains from the mice receiving antibiotics were shed at a lower percentage than the ancestor. Improvements in transmissibility occurred within the same number of lineages regardless of treatment groups (4/5), giving no suggestion of the presence of antibiotic influencing the evolution of this particular trait.

One strain from the antibiotic treated mice (N4_{P20}) appears to have undergone a major phenotypic change to being hypertransmissible, whereby the presence of this strain prevents the transmission of the wildtype ICC169 strain. This could be due to a greatly improved ability to transmit and outcompete ICC169 in the newly infected animals, or it could be that a more active competition is occurring and actively preventing ICC169 from successfully infecting the animals when N4_{P20} is present.

It is interesting to note that strains W1_{P20} and W5_{P20} show increased *C. rodentium* shedding in single infections, whereas W1_{P20} does not show increased shedding when competed with the wildtype ICC169 strain. As there is no indication of higher numbers of W1_{P20} within animals co-infected with W1_{P20} and ICC169, this could be an indication of fewer overall numbers of W1_{P20} when actively competing ICC169. After all, competition with commensal bacteria is not necessarily equivalent to competition with a similar strain (ICC169)¹⁹⁰. The W5_{P20} strain also has increased *C. rodentium* shedding in the single infections, and this is consistently reflected in the competition scenario as well. It is reassuring to see that W5_{P20} is both shed at higher numbers in single infections as well as when competed against ICC169.

5.3.4 Shedding and *in vivo* bacterial burden, as assessed by bioluminescence, are not necessarily correlated

It is interesting that some of the *in vivo*-adapted strains which exhibit improved transmissibility are not shed at greater numbers than the ancestral strain. This implies that the ability to colonise a new host is not just a numbers game, with better transmission requiring more bacteria to be shed by an infected host. Quantification of the amounts of *C. rodentium* shed in stools, while giving a measure of the dose of bacteria that animals are exposed to, is not a direct measure of the amount of *C. rodentium* present within the host animal's gastrointestinal tract. By monitoring infected animals using biophotonic imaging (BPI), the amount of *C. rodentium* present within an animal can be semi-quantified and therefore more directly inferred. The strain W5_{P20}, which was characterised by being shed at higher numbers in stool with a corresponding increase in transmission, also exhibited higher levels of light within the abdomen and rectum of infected mice (Fig. 5.6). In this case, greater numbers of bacteria present within the abdomen translated to greater numbers present in the rectum, which followed on to greater numbers being detected in the stool, and ultimately greater numbers able to be transmitted to naïve animals.

In contrast the N4_{P20} strain, which was shed at approximately three times the number of ICC180 from infected animals, also resulted in higher light levels being detectable from within the abdomen of co-infected animals but not higher amounts of light present in the rectum (Fig. 5.6), indicating a larger bacterial burden higher up in the gastrointestinal tract. The adaptation responsible for improved transmissibility in this strain, therefore, does not appear to be linked to an increase in the numbers of bacteria shed in stools, but rather an improved ability to colonise the host intestines.

While the N2_{P20} strain, which was shed at a lower percentage compared with the ancestor (4% vs. 7%), was capable of successfully transmitting to all naive animals, there was no evidence of greater numbers

present within the abdomen, and therefore improved colonisation does not appear to be associated with higher bacterial numbers *in vivo*.

The one caveat to note with these competition experiments is that, due to the inability to distinguish between the ancestral ICC180 and the *in vivo*-adapted strains, each of these strains was competed against ICC169 and the differences between the responses to this competition taken as absolute differences between ICC180 and the *in vivo*-adapted strains themselves. It is entirely possible that gaining an advantage in one environment (competing with ICC169) will not fully encapsulate the changes to another environment (competing with ICC180). As the *in vivo*-adapted strains would not have been exposed to or competed with ICC169 during their evolution period, I propose that choosing to compete with ICC169 instead of ICC180 would likely minimise, rather than emphasise, any adaptive changes, and therefore that such striking differences have been observed serves to highlight the adaptations that have occurred.

Table 5.2. Summary of observed changes of *in vivo*-adapted *C. rodentium* for *in vivo* assays

W strains		N strains	
Strain	Phenotype	Strain	Phenotype
W1_{P20}	No change in peak shedding Improved transmission (84%) Higher number shed in single infections	N1_{P20}	Slightly improved peak shedding (15%) Improved transmission (50%)
W2_{P20}	No change in peak shedding Slightly improved transmission (34%)	N2_{P20}	Reduced peak shedding (4%) Improved transmission (100%)
W3_{P20}	Slightly improved peak shedding (19%) No change in transmission	N3_{P20}	Reduced peak shedding (3%) No change in transmission (0%) Lower numbers shed in single infections
W4_{P20}	Slightly improved peak shedding (23%) Improved transmission (100%)	N4_{P20}	Slightly improved peak shedding (24%) Improved transmission (100%) Competitor ICC169 can not transmit More light detected from abdomen
W5_{P20}	Improved peak shedding (72%) Improved transmission (100%) More light detected from abdomen and rectum Higher numbers shed in single infections	N5_{P20}	Slightly improved peak shedding (13%) Improved transmission (100%)

5.3.5 Conclusions

In this chapter, I have investigated the *in vivo* fitness of each of the *in vivo*-adapted *C. rodentium* strains, as well as competed each of the strains head-to-head with the wild type ICC169 strain. This was used to determine the relative fitness, by comparing to benchmark data where the ancestral ICC180 strain is competed with the wild type ICC169 strain, as detailed in Chapter 3 and published in PeerJ¹⁵³. From these experiments I have shown that 9/10 *in vivo*-adapted *C. rodentium* strains have improved in comparison to their ancestor in at least one of the measured outcomes: numbers of bacteria shed during single infection of mice, numbers of bacteria shed when competed with ICC169 in mice, and the ability to transmit to a naïve host when in competition with ICC169.

There do not appear to be any trends with regards to impact of antibiotic treatment and the type or degree of improvements, and therefore based on this data I am still unable to disprove either of the two opposing hypotheses:

- 1) Low-dose nalidixic acid treatment will alter normal microflora, reducing the amount of genetic information available for recombination purposes and thus limiting adaptation; or
- 2) Low-dose nalidixic acid treatment will alter normal microflora, resulting in vacant space for *C. rodentium* and thus remove the barriers for colonisation, enabling efficient infection and faster adaptation.

Of the 9 strains which showed an improved ability to transmit between mice, one strain is capable of preventing the transmission of the competing ICC169 strain, either directly or indirectly, to naïve animals. Therefore, after only 5-months experimental adaptation to the mouse host, I have observed the emergence of new phenotypes in the *in vivo* infection model.

Chapter 6: *In vitro* assessment of *in vivo*-adapted *C. rodentium*

6.1 Introduction

This chapter will cover the results of assays I performed to further phenotypically investigate and characterise the adapted *C. rodentium* strains that emerged as a result of *in vivo* adaptation. In the first instance, I investigated the ability of the *in vivo*-adapted strains to grow and bioluminesce in rich and restricted laboratory media, and assessed them for morphological changes using light microscopy. I also quantified the mutation and recombination rates of the strains, hypothesising that the strains which had adapted in the absence of antibiotic would have an improved recombination rate. Finally, I investigated the improved ability of the *in vivo*-adapted strains to transmit between animals by assessing their ability to survive desiccation and to attach to mouse fibroblast cells *in vitro*.

6.2 Results

6.2.1 Emergence of an aggregating *C. rodentium* strain following *in vivo* adaptation

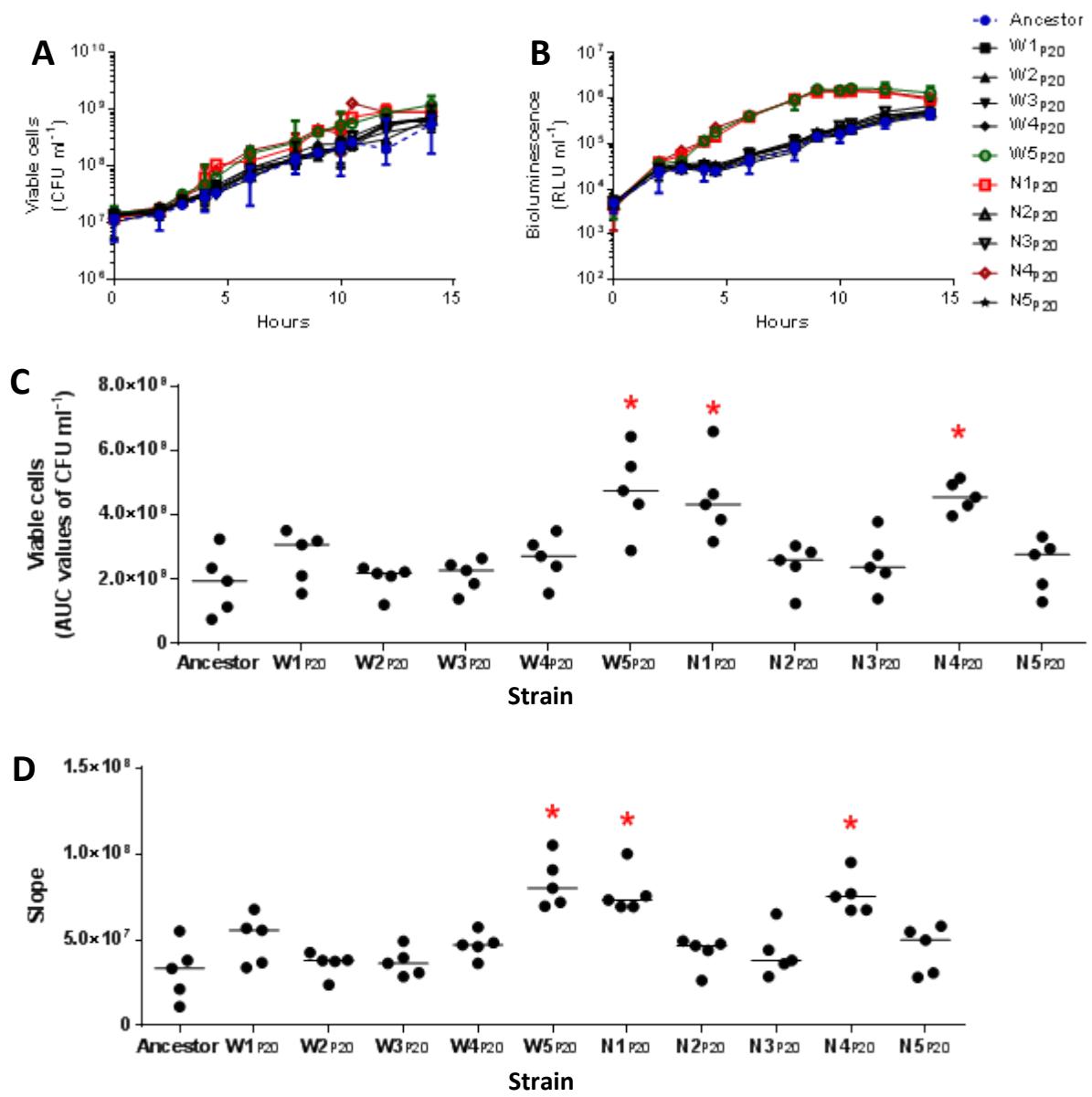
I investigated the ability of the *in vivo*-adapted strains to grow and produce light in rich and restricted laboratory media, by inoculating 50 ml tubes containing 10 ml of media with $\sim 1 \times 10^7$ CFU of either the *in vivo*-adapted strains or the ancestral ICC180 strain. I took samples at regular intervals to measure bacterial numbers (as CFU ml^{-1}) (Fig. 6.1A and Fig. 6.2A) and bioluminescence (as RLU ml^{-1}) (Fig. 6.1B and Fig. 6.2B).

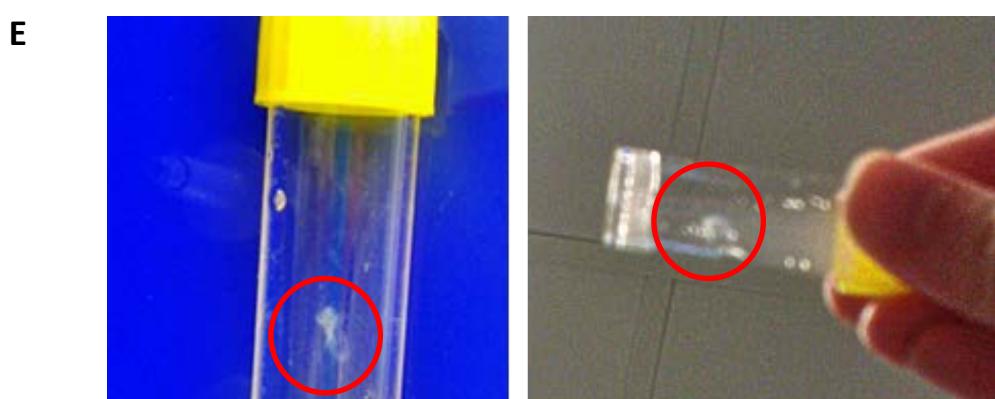
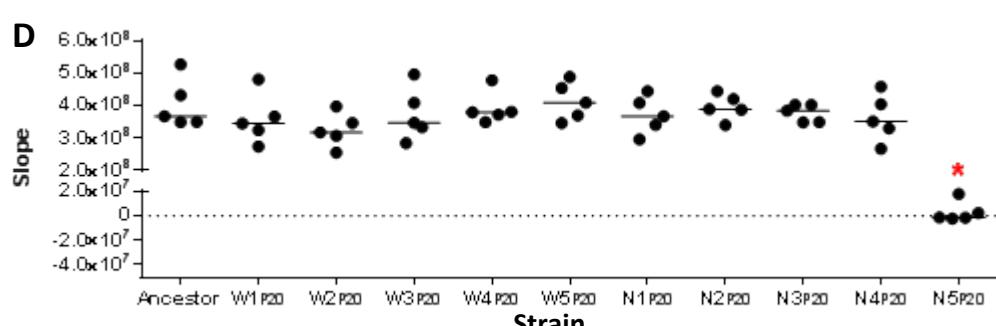
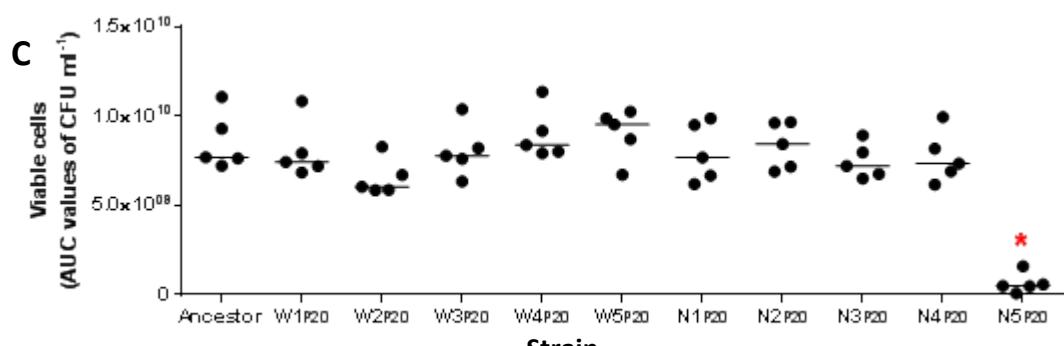
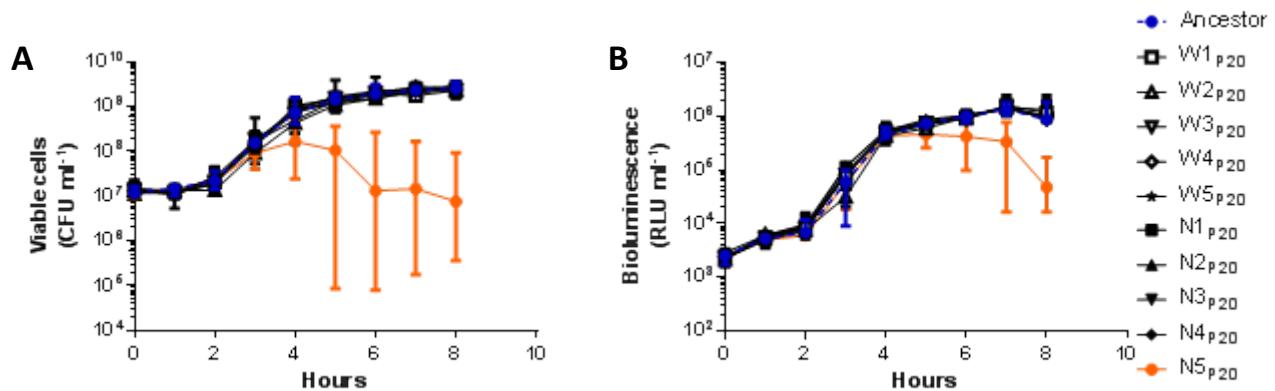
When grown in modified DM minimal media, supplemented with 1% glucose, I observed that three of the *in vivo*-adapted strains (W5_{P20} , N1_{P20} , N4_{P20}) showed improved growth in comparison to ICC180, as evidenced by higher bacterial counts (Fig. 6.1A) and calculated area under curve (AUC) values (Fig. 6.1C) ($p < 0.05$). The three strains also had significantly steeper slopes between 2 and 8 hours (Fig. 6.1D), indicating faster growth during the exponential growth phase. This faster growth was also reflected in higher light production by the three strains (Fig. 6.1B).

When grown in rich laboratory media (LB), I observed that nine of the ten *in vivo*-adapted strains were indistinguishable from the ancestral ICC180 strain, both by bacterial counts (Fig. 6.2A) and calculated AUC values (Fig. 6.2C), and light production (Fig. 6.2B). In contrast, I observed that there was a consistent drop in bacterial counts and bioluminescence for strain N5, beginning after approximately 4 hours growth (Fig. 6.2). This drop was reflected in significantly lower AUC values for bacterial growth when compared to ICC180 ($p < 0.05$) and coincided with the appearance of small, irregular shaped bacterial aggregates with poorly defined edges, which had a tendency to float to the air-liquid interface (Fig. 6.2E).

Figure 6.1. Some lineages of *in vivo*-adapted *Citrobacter rodentium* show improved growth in minimal media (overleaf). 10 ml of minimal media (modified DM media with 1% glucose) was inoculated with 20 μl (approximately 1×10^7 CFU) of *C. rodentium* and sampled at regular intervals for measuring bacterial numbers (A) and bioluminescence (B). Bacterial numbers were converted into area under the curve (AUC) values (C). The exponential-phase slopes (between 2 and 8 hours) for each strain was calculated (D). A non-parametric two-tailed Friedman test was applied to test for variation between all strains for AUC and slope data. A Wilcoxon signed-rank test was used to test for significant differences between the ancestor and each evolved strain; a red star marks statistically significant differences ($p < 0.05$).

Figure 6.2. Emergence of aggregates in the N5_{P20} lineage of *in vivo*-adapted *Citrobacter rodentium* in rich media (overleaf). 10 ml of rich media (LB) was inoculated with 20 μl ($\sim 1 \times 10^7$ CFU) of *C. rodentium* and sampled at regular intervals for measuring bacterial numbers (A) and bioluminescence (B). Bacterial numbers were converted into area under the curve (AUC) values (C). The exponential-phase slopes (between 2 and 4 hours) for each strain was calculated (D). A non-parametric two-tailed Friedman test was applied to test for variation between all strains for AUC and slope data. A Wilcoxon signed-rank test was used to test for significant differences between the ancestor and each evolved strain; a red star marks statistically significant differences ($p < 0.05$). Photographs of *C. rodentium* N5_{P20} following 8 hours of growth in rich media (LB) with aggregates highlighted within the red circle (E).





6.2.2 Aggregating *C. rodentium* N5_{P20} does not disperse in the presence of cellulase

To further investigate the nature of the aggregates produced by *C. rodentium* N5_{P20}, I grew the bacteria overnight in LB and stained them with calcofluor white. Calcofluor white is a fluorochrome which binds to β -1,3 and β -1,4 polysaccharides, such as cellulose and chitin^{146,147}. As a positive control, I also grew the uropathogenic *E. coli* (UPEC) strain 536 in RPMI; UPEC aggregates stained with calcofluor white fluoresce strongly under UV light due to the presence of cellulose¹⁴⁷ (Fig. 6.3). The N5_{P20} aggregates I stained with calcofluor white also appear to fluoresce when exposed to UV light (Fig. 6.3), while the non-aggregating ICC180 strain did not show any evidence of fluorescence. The UPEC aggregates have clearly defined edges, whereas the N5_{P20} aggregates appear to be formed by long ‘threads’ coiled together.

To investigate whether the aggregates produced by N5_{P20} are composed of cellulose, I exposed LB-grown cultures of N5_{P20} to a range of cellulase concentrations (10-100 $\mu\text{g ml}^{-1}$). Cellulase cleaves cellulose and disperses UPEC aggregates¹⁴⁷. The N5_{P20} strain remained able to form aggregates both with the pre-emptive and post-aggregate addition of cellulase at all concentrations tested (data not shown).

ICC180

UPEC 536

N5_{P20}

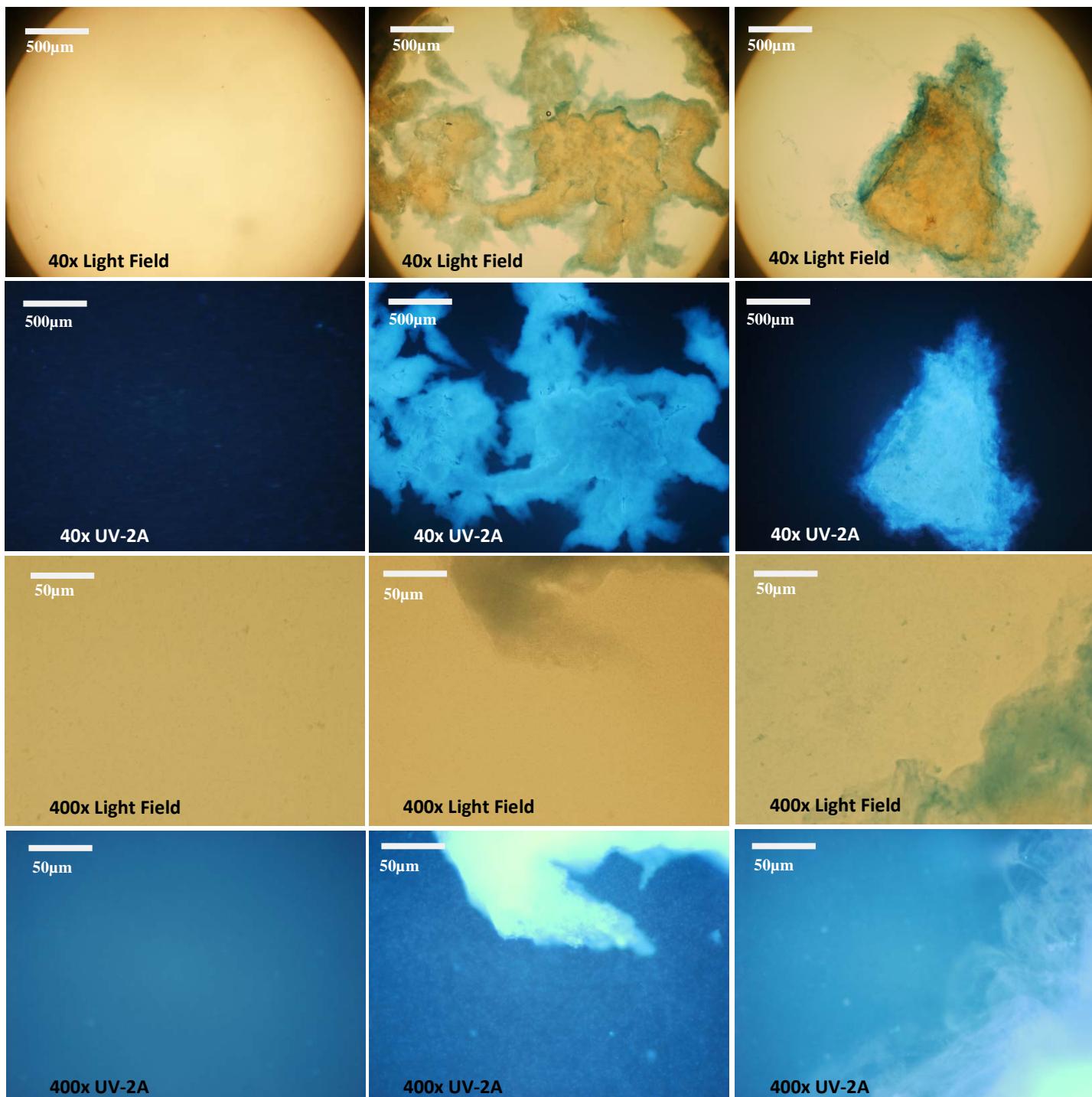


Figure 6.3. N5_{P20} aggregates fluoresce under UV light when stained with calcofluor white. *Citrobacter rodentium* N5_{P20} was grown overnight in LB. *C. rodentium* ICC180 was used as a negative control, and UPEC 536 grown in RPMI was used as a positive control. Samples were dried and heat-fixed on sterile glass slides and stained with calcofluor white by adding 20 μ l of stain and applying a coverslip. Slides were observed using light microscopy and fluorescence determined using UV-2A light. The scale bar is 500 μ m for 40x magnification, and 50 μ m for 400x magnification.

6.2.3 No obvious microscopic variation between evolved *C. rodentium* and the ancestor

To investigate any obvious morphologic differences between *in vivo*-adapted and ancestral *C. rodentium*, bacteria were stained and observed under light microscopy. Liquid cultures of each of the *C. rodentium* strains were grown in LB and stained using the Gram's method¹⁴⁵. No changes were observed for the *in vivo*-adapted strains, with the exception of N5P20 which aggregates when grown in LB (Fig. 6.4). A diffuse lightly staining substance was present around the cells, perhaps indicating the presence of a secreted compound (Fig. 6.4). The aggregates appear to be made of densely packed cells combined with the lightly staining substance, with individual *C. rodentium* cells visible at the edges of the aggregate.

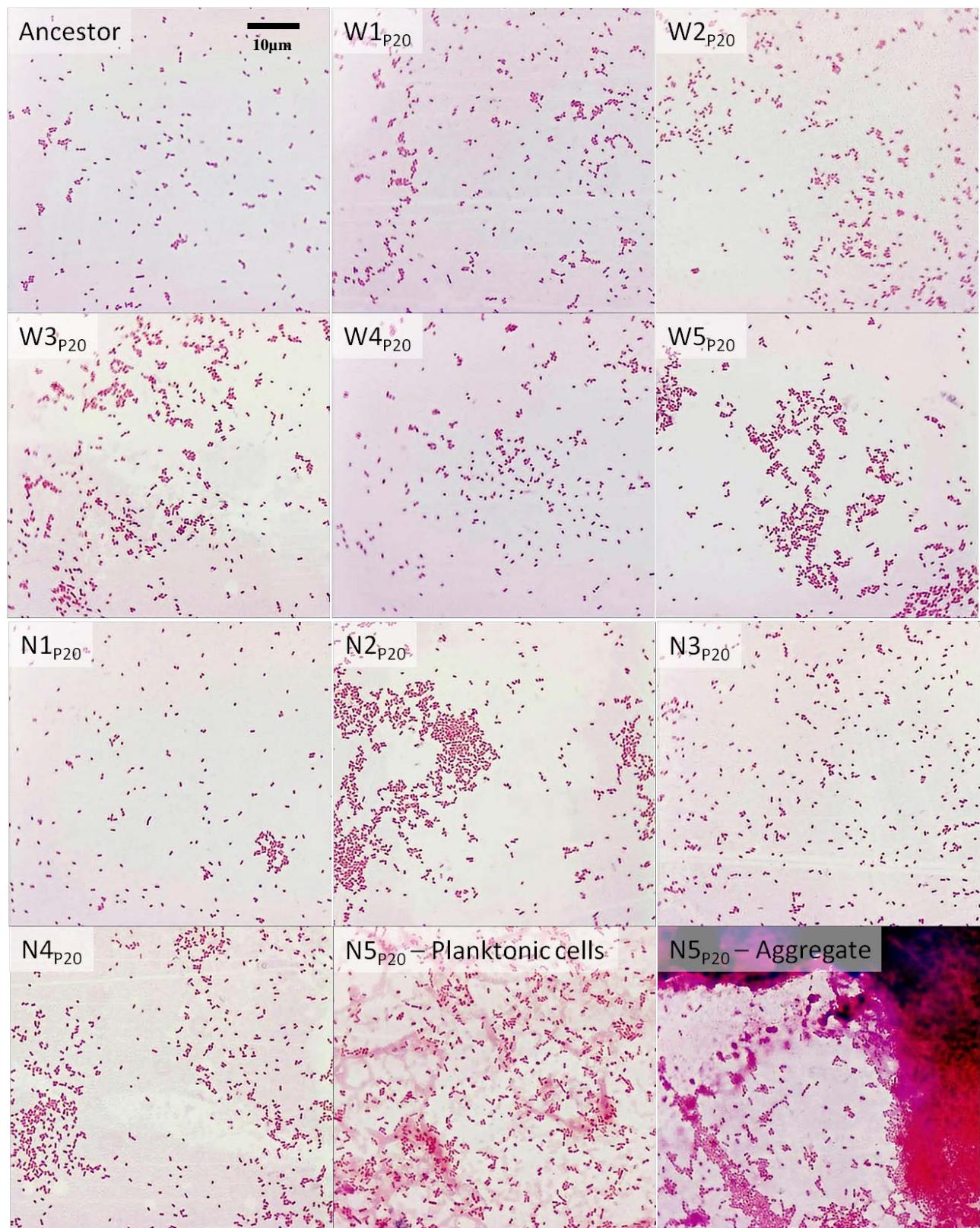


Figure 6.4. *In vivo*-adapted *Citrobacter rodentium* show no morphological changes under light microscopy. *Citrobacter rodentium* strains ICC180 (ancestor) and the *in vivo*-adapted *C. rodentium* (W1_{P20}-W5_{P20} and N1_{P20}-N5_{P20}) were dried and heat-fixed on sterile glass slides and then stained using the Gram's method¹⁶⁴, then photographed at 1000x magnification using a Leica microscope and the Leica AirLab application (Leica Microsystems). Representative images are shown. The scale bar is 10 μm.

6.2.4 No change in tolerance to desiccation in *in vivo*-adapted *C. rodentium*

As the source of infection for the naive animals in the competition experiments is solely from exposure of the mice to *C. rodentium* contaminating the cage environment, and not from direct mouse-to-mouse contact, and because the animals remain in the contaminated cage for several days before infection can be detected, one possible explanation for the improved transmission rates of the *in vivo*-adapted strains is an improved tolerance to desiccation and an increased ability to survive in the cage in dried faecal pellets. To investigate this, I aliquoted washed *C. rodentium* into sterile tubes, allowed the bacteria to air dry, and then revived the dry cultures at various time points, by resuspending in saline and plating onto solid media to determine the surviving number of *C. rodentium* (Fig. 6.5). I observed no evidence for a change in tolerance to desiccation (Friedman's test $p>0.7$). Starting with an inoculum of approximately 5×10^9 CFU, all strains had reduced by approximately 90% after 24 hours desiccation, further reducing to over 99% loss after seven days.

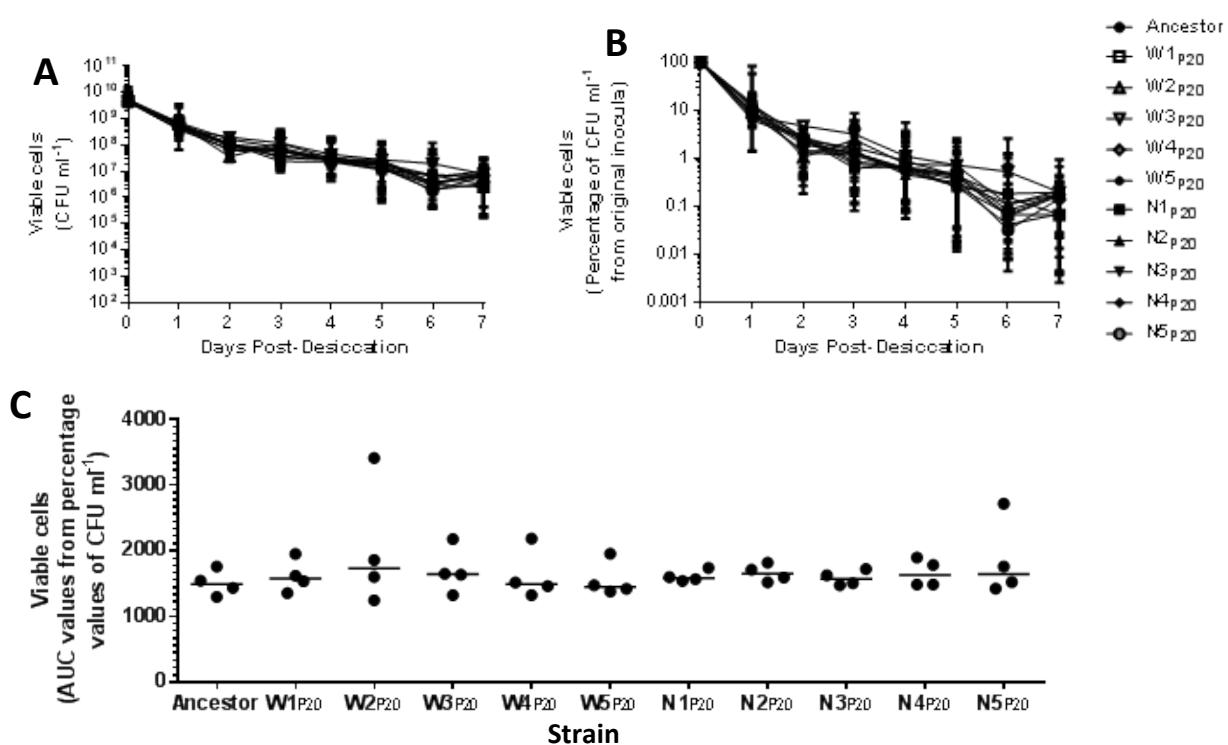


Figure 6.5. *In vivo*-adapted *Citrobacter rodentium* show no change in tolerance to desiccation. The ancestral ICC180 strain and the *in vivo*-adapted strains were washed in sterile phosphate buffered saline and 10 μl drops allowed to dry in tubes. Samples were recovered daily and plated to determine the surviving number of bacteria (A). This is represented as a percentage of initial population (B) and as area under the curve (AUC) of the percentage values (C). No statistically significant differences between the groups was observed (Friedman's test $p>0.7$). This experiment was performed on four independent occasions.

6.2.5 No change in the ability of *in vivo*-adapted *C. rodentium* strains to attach to mouse fibroblast cells

To assess the ability of the *in vivo*-adapted *C. rodentium* strains to attach to mouse cells, I grew the bacteria overnight in DMEM, supplemented with 10% foetal calf serum, then diluted the bacteria 1:100 in fresh DMEM. I added approximately 10^6 CFU of *C. rodentium* to a layer of confluent mouse fibroblast cells (L-929) growing in the wells of a 96-well clear bottomed microtitre plate (Perkin Elmer) for 48 hours to ensure a confluent monolayer of cells. I incubated the samples for 2 hours without shaking at 37 °C, and then measured light levels, as a surrogate for bacterial numbers, before and after gently washing the cells to remove any non-adherent bacteria. I observed no significant differences in the amount of light detected for each of the strains (Fig 6.6.A), indicating no observable change in the ability of *in vivo*-adapted *C. rodentium* to attach to mouse L-929 fibroblast cells. I also used light microscopy (100x magnification) to examine the fibroblasts after they had been incubated with the *C. rodentium* strains and observed no differences in the health of the fibroblasts; L-929 cells incubated with the *in vivo*-adapted strains appeared indistinguishable from cells exposed to the ancestral ICC180 strain and unexposed control cells (Fig 6.6.B).

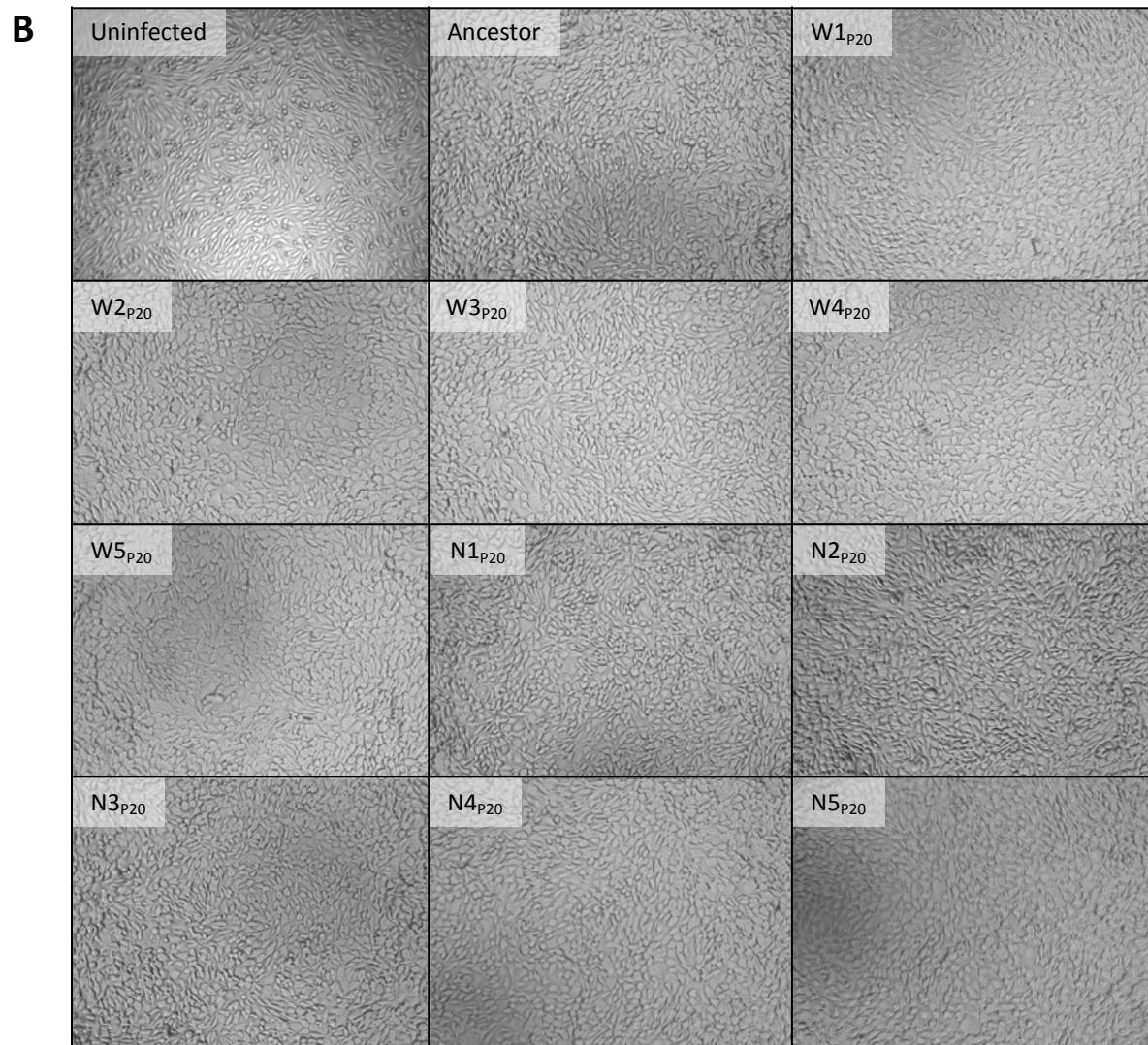
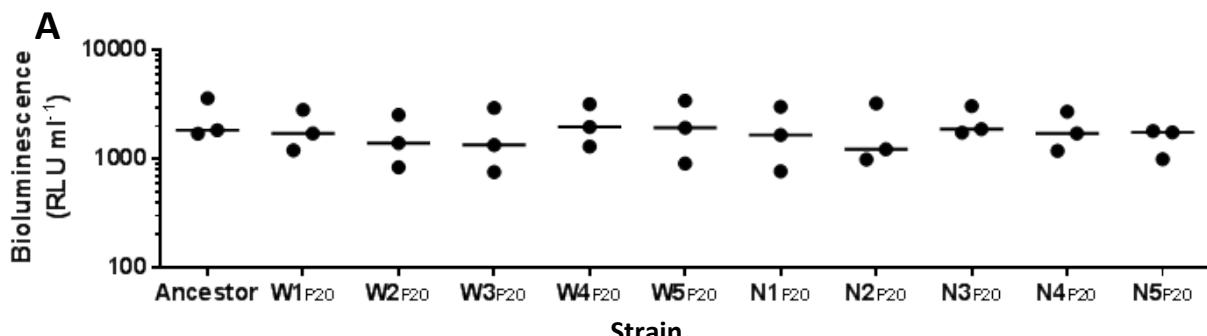


Figure 6.6. No change in ability of *in vivo*-adapted *Citrobacter rodentium* strains to attach to mouse L929 fibroblast cells. Confluent L-929 fibroblast cells were exposed to *C. rodentium* for 2 hours, and then gently washed to remove non-adherent bacteria three times with 200 µl of sterile phosphate buffered saline (PBS). (A) The amount of bioluminescence detected from wells minus the light from the media-only control wells. (B) Light microscopy images at 100x magnification of representative wells following washing with PBS.

6.2.6 Emergence of a hypermutable strain of *C. rodentium* following *in vivo* adaptation

I investigated whether *in vivo*-adaptation changed the mutation rate of any of the strains by measuring the rates of spontaneous resistance to two antibiotics with different modes of action: the aminoglycoside streptomycin²⁰⁸ and the ansamycin rifampicin²⁰⁹. Each of these antibiotics requires a single point mutation for resistance to occur. I used two different antibiotics to counteract the potential for a high rate of mutation in a particular gene. I grew the bacteria overnight in LB with no selection, and then plated onto LB agar plates containing either 20 µg ml⁻¹ streptomycin or 80 µg ml⁻¹ rifampicin. I chose these antibiotic concentrations as they are a bactericidal concentration at which detectable resistant colonies can be counted. The percentage of resistant colonies which emerged are shown in Figure 6.7. While nine of the ten *in vivo*-adapted strains showed no change in their mutation rates when compared with the ancestral ICC180 strain, N3_{P20} demonstrated a significant increase in the percentage of resistant colonies after exposure to both rifampicin (Fig. 6.7.C) and streptomycin (Fig. 6.7.D). This increase corresponded to a tenfold increase in resistance rates to rifampicin (9.146×10^{-7} vs. 7.900×10^{-6} median resistance rates) and to streptomycin (9.794×10^{-6} vs. 3.817×10^{-5} median resistance rates).

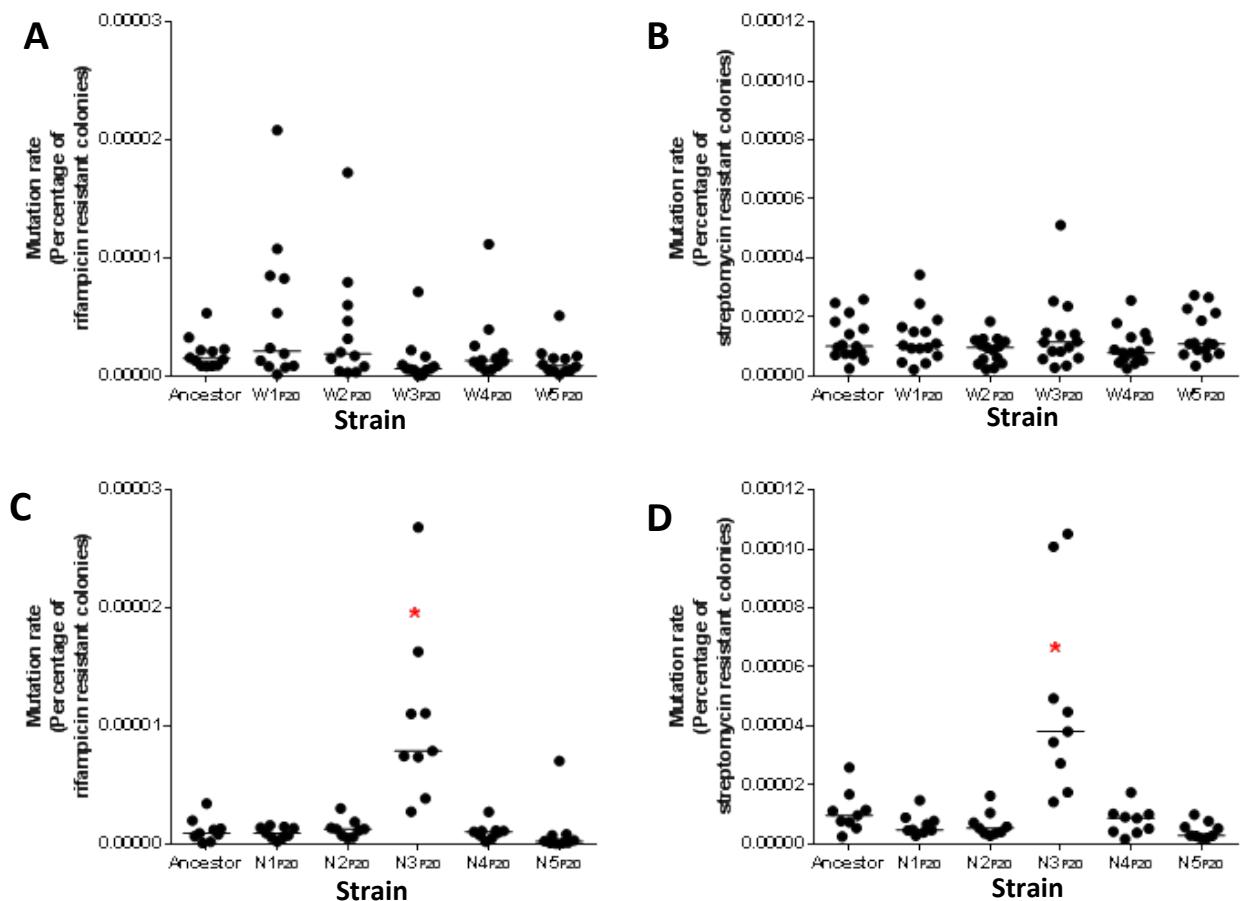


Figure 6.7. One lineage of *in vivo*-adapted *Citrobacter rodentium* has an increased mutation rate to two separate antibiotics. Overnight cultures of *C. rodentium* were plated onto LB agar plates containing either 80 µg ml⁻¹ rifampicin (A and C) or 20 µg ml⁻¹ streptomycin (B and D). The number of resistant colonies grown on the plates was taken as a proxy for mutation rate. A Kruskal-Wallis test was performed to determine statistically significant differences between groups ($p<0.0001$), and a post hoc Dunn's multiple comparisons test performed to determine *in vivo*-adapted strains which were significantly different from the ancestor; these are marked with a red star ($p<0.05$). Each data point represents an independent experiment.

6.2.7 No detectable change in recombination rate for *in vivo*-adapted *C. rodentium*

I investigated whether *in vivo*-adaptation changed the recombination rate of any of the strains using the pRhom plasmid and assay developed by Rodríguez-Beltrán and colleagues¹⁴⁴. The plasmid contains an ampicillin resistance gene (*bla*), and a gentamicin resistance gene (*aacCI*) flanked by two non-functional parts of a tetracycline resistance gene (*tetA* 5' and *tetA* 3') which share a 627 bp region of 100% homology. If recombination occurs, the gentamicin resistance gene is removed from the plasmid, and the tetracycline resistance gene restored. Without recombination the plasmid confers resistance to ampicillin and gentamicin. With recombination the plasmid confers resistance to ampicillin and tetracycline. Therefore, the ratio of colonies resistant to gentamicin (containing the non-recombined plasmid) and colonies resistant to tetracycline (containing the recombined plasmid) can be used to determine the recombination rate of each strain (Fig. 6.8). I observed no significant differences between the strains (Friedman's test; $p=0.1751$), indicating *in vivo*-adaptation did not alter the recombination rate of *C. rodentium* as assessed by this assay.

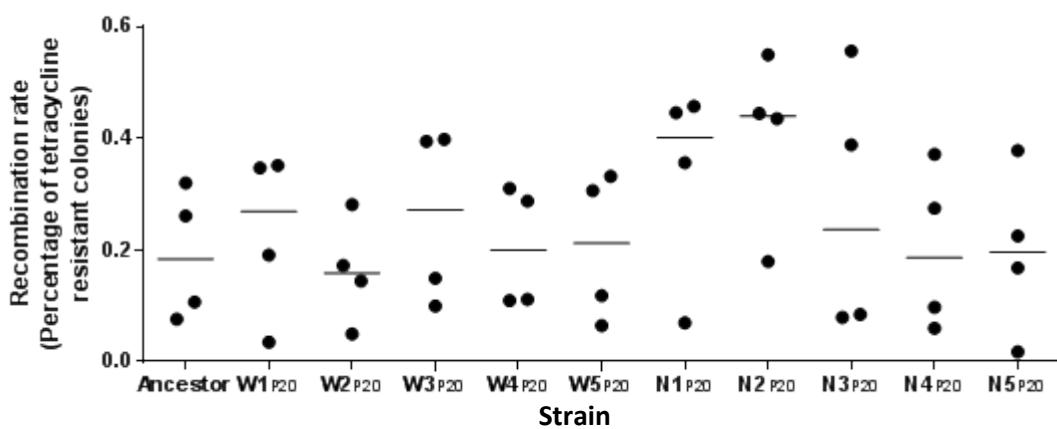


Figure 6.8. *in vivo*-adapted *Citrobacter rodentium* shows no apparent change in recombination rate. *C. rodentium* was transformed with pRhom plated onto LB containing $100 \mu\text{g ml}^{-1}$ ampicillin and $10 \mu\text{g ml}^{-1}$ gentamicin, and five successful transformants were grown overnight in 1 ml of LB containing $100 \mu\text{g ml}^{-1}$ ampicillin. The cultures were then serially diluted and plated onto LB agar plates containing $100 \mu\text{g ml}^{-1}$ ampicillin and either $10 \mu\text{g ml}^{-1}$ gentamicin or $10 \mu\text{g ml}^{-1}$ tetracycline. The ratio of gentamicin resistant mutants and tetracycline resistant mutants was used to determine the recombination rate. Each data point represents an independent experiment of five transformants.

6.3 Discussion

The data presented in this chapter sought to compare the *in vivo*-adapted *C. rodentium* strains to their ancestor in a variety of *in vitro* assays, assessing growth, visual differences, mutation rates and recombination rates. A summary of the observed changes of the *in vivo*-adapted *C. rodentium* can be found in Table 6.1. There were surprisingly few differences observed, with each *in vivo*-adapted strain performing at least as well as the ancestor in each of the assays.

6.3.1 *In vivo*-adapted *C. rodentium* show no changes in desiccation tolerance or attachment to mouse fibroblast cells

In order to functionally assess the reasons for improved transmissibility which was observed for many of the *in vivo*-adapted strains (Chapter 5), I investigated the ability of the bacteria to survive outside of the host and in the cage environment, and the ability to attach and therefore colonise mouse cells. Attachment of *C. rodentium* to the surface of the colon is an important initial step for colonisation of a new host. *C. rodentium* lacking the protein EspB, which is required for attachment of *C. rodentium* to epithelial cells, cannot colonise Swiss Webster mice and show a reduced ability to attach to HEp-2 cells *in vitro*¹²⁵. C57BL/6 mice receiving polyethylene glycol (an inhibitor of β_1 -integrin, which is the surface receptor in intestinal cells important for pathogen attachment) prior to infection with *C. rodentium* and during infection decreased the numbers of *C. rodentium* in the colon and reduced symptom severity²¹⁰. As attachment is an important part in the infection process, improvements in transmissibility might be explained by an improved ability to attach to cells. I observed no evidence of improvement in attachment of the *in vivo*-adapted strains to L-929 fibroblast cells in the experiments described in this chapter (Fig. 6.6). This may be due to the artificial nature of the *in vitro* experiment, with a layer of cells on the bottom of a microtitre plate not fully reflecting the mouse gut environment. Alternatively, priming the bacteria by incubation in DMEM may not reflect changes in gene expression which may occur following excretion from a mouse and subsequent ingestion by the next animal. Conversely, it may be that there is no improvement in attachment and that the improved ability of the majority of the *in vivo*-adapted strains to transmit are a result of improved ability of the evolved *C. rodentium* to physically reach the site of attachment, a step which is not assessed in this *in vitro* experiment.

Another important hurdle that pathogens must overcome in order to infect their host is to survive in the environment long enough to reach the next host. Desiccation has a major impact on bacterial survival, and we know that both experimentally, and in the field, enteric bacteria are recovered at lower levels in stools with lower moisture content^{211,212}. As the transmission animals in the experiments detailed in Chapter 5 were exposed to the cage containing infectious faecal matter for 1 week, I wondered whether

the improved transmission was an artefact of improved tolerance to desiccation. However, I observed no differences in the ability of the ancestral or *in vivo*-adapted strains to survive desiccation (Fig. 6.5). This experiment was limited, again, due to its artificial nature: the bacteria were grown in a nutrient rich environment (LB) and dried in a nutrient poor environment (PBS), which will not fully mimic growth in the mouse gut and subsequent drying on faecal matter and any associated gene expression. However, based on these results there is no indication that improved tolerance to a dry environment is a likely candidate, in the case of these *in vivo*-adapted strains, for improved transmission.

6.3.2 Changes in growth in rich and restricted media

The *in vivo*-adapted strains grow equally as well as the ancestor in rich media (LB), with no major growth changes, with the exception of N5_{P20} which forms aggregates reliably after approximately 4 hours into the growth curve (cell concentration of approximately 1×10^8 CFU ml⁻¹). This is reflected in a rapid drop in detected viable bacteria and light production due to unreliable sampling of increasingly aggregated cultures. Aggregation is not complete as in the case of UPEC grown in RPMI, where the non-aggregated media is as clear as uninoculated media. Instead, clumps of variable sized aggregates were present in turbid media with many measurable non-aggregated cells present (approximately 5×10^6 CFU ml⁻¹). Also, unlike UPEC clumps which are long, stringy, and readily fall to the bottom of the tube, the N5_{P20} aggregates float throughout the media and often rise to the liquid/air interface. When shaken, air bubbles form easily and are stable for hours, indicating possible protein or surfactant molecules secreted into the media. I did not observe any evidence of clumping in the restricted DM media.

All *in vivo*-adapted strains performed at least as well as the ancestral ICC180 strain when grown in a more restricted environment (DM media with 1% glucose supplementation). Three of the *in vivo*-adapted strains grew significantly better than the ancestor, with steeper exponential phase slopes and greater calculated AUC values of CFU ml⁻¹ over time (Fig. 6.1). Two of the strains were from the treatment group which evolved in the presence of nalidixic acid (N1_{P20} and N4_{P20}), and one strain from the group which received no antibiotic (W5_{P20}), giving no indication that the antibiotic treatment influenced ability to grow in a restricted environment. Instead, these results offer assurance that adaptation to the mouse environment did not result in any growth defect when grown in either rich or restricted laboratory media.

6.3.3 One strain has evidence of increased mutation rate, no evidence of recombination rate change

Genetic variation is crucial to provide a pool of information on which adaptation can occur and select for beneficial genome compositions. The inherent mutation rate can determine the amount of genetic variation available to a population, and ‘mutator strains’ of bacteria have been shown to occur in evolving populations^{36,213–215}. I assessed the mutation rate of all *in vivo*-adapted strains and the ancestor using two separate antibiotics which each only require a single point mutation for resistance to occur: the aminoglycoside streptomycin²⁰⁸ and the ansamycin rifampicin²⁰⁹. Streptomycin acts on protein synthesis and resistance typically arises following point mutations in the *rpsL* gene^{208,216–218}. Rifampicin acts on DNA transcription and resistance typically arises following point mutations in the *rpoB* gene^{209,219–221}. By assessing the spontaneous mutation rate to two separate antibiotics with separate mechanisms of resistance, I propose that the effect of increased propensity to develop mutations in certain parts of the genome is negated²²² and offers further weight to any changes in mutation rate. One out of the ten *in vivo*-adapted strains (N3_{P20}) exhibited a tenfold increase in resistance rates to rifampicin (9.146×10^{-7} vs. 7.900×10^{-6} median resistance rates) and to streptomycin (9.794×10^{-6} vs. 3.817×10^{-5} median resistance rates), compared to ICC180 (Fig. 6.7). I did not observe any significant change in mutation rate for any of the other strains.

In *E. coli*, it is estimated that recombination is more frequent than point mutations²²³, with the suggestion that bacterial species in complex mixed environments may favour gaining genetic variability through acquiring new genetic information from plasmids, phages, and homologous recombination instead of through point mutations²²⁴. Rodríguez-Beltrán and colleagues have recently developed the pRhomo plasmid for the purpose of estimating relative recombination rates in commensal and extraintestinal pathogenic *E. coli* strains¹⁴⁴. The *in vivo*-adapted strains did not appear to have an increase or decrease in recombination rate when grown in LB (Fig. 6.8). There are many potential technical explanations for this: for example, by selecting five colonies grown on a plate following each transformation to inoculate liquid broth, a different amount of bacteria was sampled each time due to the non-standard sampling of taking a colony instead of a known quantity from sample grown in liquid broth. As the number of generations has been shown to impact the rate of evolution and would likely influence the number of recombination events occurring³⁵, it is crucial to have a standardised and known starting point. Similarly, allowing for a greater number of generations to occur by starting with a smaller inoculum would provide a greater chance for smaller differences in recombination rate to be highlighted. As Rodríguez-Beltrán and colleagues explored in their paper, the type of growth media also has an impact on recombination rate¹⁴⁴. Investigating the recombination rate of the *in vivo*-adapted strains in a more restrictive media than LB, or a media better able to mimic the mouse gut environment would help to fully explore any changes in recombination rate.

6.3.4 Conclusions

A summary of observed phenotypic changes from the *in vitro* assays is presented in Table 6.1. Interestingly, the majority of distinct phenotypes were found in the “N” strains, with only one “W” strain (**W5_{P20}**) presenting with increased growth in the restricted media environment. It is important to note that the assays are limited in the degree of phenotypes that can be detected, and so further phenotypic changes may be present which were not assessed.

Table 6.1. Summary of observed changes of *in vivo*-adapted *C. rodentium* for *in vitro* assays

W strains		N strains	
Strain	Phenotype	Strain	Phenotype
W1_{P20}	None identified	N1_{P20}	Improved growth and bioluminescence in restricted DM media
W2_{P20}	None identified	N2_{P20}	None identified
W3_{P20}	None identified	N3_{P20}	Tenfold increase in mutation rate to streptomycin and rifampicin
W4_{P20}	None identified	N4_{P20}	Improved growth and bioluminescence in restricted DM media
W5_{P20}	Improved growth and bioluminescence in restricted DM media	N5_{P20}	Formation of aggregates in rich media

Chapter 7: Genomic changes in *in vivo*-adapted *C. rodentium*

7.1 Introduction

After identifying the phenotypic changes which have occurred following *Citrobacter rodentium* adaptation to the mouse environment, I sought to determine any genetic changes present in these *C. rodentium* populations. The *C. rodentium* genome has already been shown to be unstable, prone to rearrangements, and with the suggestion that it is still in the process of changing to its environment¹³⁴. I used a population based method in order to gain insight into the types and frequency of change, instead of sequencing a single clone and potentially losing a large degree of genetic diversity. In Chapter 3, the genome of the ancestor ICC180 was sequenced¹⁵³. In this chapter, I will outline the findings from whole genome sequencing of each of the *in vivo*-adapted strains.

I tested and compared a variety of pipelines (data not shown), as differences in downstream analysis of sequencing data can influence the changes observed. The BreSeq software from the Barrick laboratory was ultimately chosen for its ability to handle varied populations and for performing statistical tests on each identified change to determine its validity^{171,225}. This pipeline is capable of determining “missing coverage” (deletion events), “read alignment” changes (SNPs and small insertions/deletions), and “junction changes” (possible rearrangement events), as well as assessing the frequency of these changes.

7.2 Results

C. rodentium in frozen stool from each lineage at the end of the evolution experiment (passage 20) was revived and grown in rich media (LB) with 50 µg ml⁻¹ kanamycin to enrich for *C. rodentium* populations. I extracted genomic DNA using a Qiagen Genomic-tip 20-G (Qiagen) system as per the manufacturer's instructions. I then assessed the genomic DNA for quality by running on a 0.8% agarose TAE gel (data not shown) and by NanoDrop™ assessment, with recorded 260/280 ratios of between 1.79 and 1.95. The DNA was then sequenced by New Zealand Genomics Limited on an Illumina MiSeq. I mapped the resulting .fastq files to the reference genome ICC168 (Genbank accession number FN543502.1¹⁰²), and to the known plasmids present in *C. rodentium*: pCROD1 (Genbank accession number FN543503.1), pCROD2 (Genbank accession number FN543504.1), pCROD3 (Genbank accession number FN543505.1) and pCRP3 (Genbank accession number NC_003114). Mapping and identification of statistically valid genomic changes were assessed using BreSeq version 0.24rc6^{171,225}, with an average read coverage of 53.5.

7.2.1 Single nucleotide polymorphisms and other genetic changes

I chose a cut-off genetic change frequency of 25% to distinguish between true genetic changes and sequencing artefacts. A filtered list of polymorphisms, including single nucleotide changes and small insertions or deletions, is shown in Table 7.1. Beginning from the left, the first column indicates the location of each particular genetic change in the genome, then the genetic change itself and, if located within a gene (intragenic), the predicted change to the amino acid. Presence in each strain and its frequency in the population is shown on the right hand side of the table, with a colour code reflecting the proportion of the reads containing each genetic change, ranging from red (a low percentage) to green (a high percentage). I observed a wide range of frequencies for the single nucleotide polymorphisms (SNPs), with some present in 100% of the reads analysed, and others present at 50% or below. This suggests the presence of subpopulations within the assessed *in vivo*-adapted strains, with no one single clone encompassing the niche.

The changes that I observed are, for the most part, conserved between *in vivo*-adapted strains. The two regions highlighted in orange are regions where different lineages have different genetic changes in the same location on the genome; the first of these is a change from an adenine to either a guanine (strain W5_{P20}) or a thymine (strain N1_{P20}) at location 297,042. The second change, which is at locations 4,288,267 and 4,288,269 and impacts on the same residue, is a change from a cytosine to a thymine (strains N1_{P20} and N5_{P20}) and a change from a guanine to a thymine (strains W2_{P20}, W3_{P20}, W4_{P20}, W5_{P20}, N2_{P20}, and N4_{P20}), each of which is a synonymous mutation with no predicted effect on the lysine. Outside of these two examples, all other changes are either unique to each strain or shared between lineages, with no other observations of the same site being changed in different ways by different lineages.

The total number of recorded genetic changes (SNPs and insertions/deletions) over all of the *in vivo*-adapted genomes is 210, of which 9 were insertions or deletions. The total number of non-synonymous SNPs recorded over all of the *in vivo*-adapted genomes is 145, and the total number of synonymous SNPs is 34 (total intragenic SNPs: 179; total intergenic SNPs: 22). I calculated a d_N/d_s ratio of 4.294, with ratios greater than 1 commonly suggesting positive selection of the mutations^{226,227}, which suggests that a number of these mutations will provide a fitness advantage in the mouse environment. The majority of the mutations are unique to a single lineage (135 SNPs or small insertions/deletions), while 74 mutations are present in two or more of the lineages. I observed a similar number of SNPs present in the “W” strains and the “N” strains, with 132 SNPs found in the “W” strains and 127 found in the “N” strains.

Table 7.1. Intergenic and intragenic genetic changes present at ≥25% frequency

Location	Genetic change	Gene affected	Function	W1 _{P20}	W2 _{P20}	W3 _{P20}	W4 _{P20}	W5 _{P20}	N1 _{P20}	N2 _{P20}	N3 _{P20}	N4 _{P20}	N5 _{P20}
17,805	G→T (intergenic)	ROD_00141 ← / → ROD_00151	hypothetical protein/polysaccharide degrading enzyme								100		
89,119	T→C (E320G)	thiP ←	thiamine ABC transporter permease					29.8					
154,437	C→T (F234F)	ROD_01271 →	2-keto-3-deoxygluconate permease							28			
154,495	T→C (S254P)	ROD_01271 →	"	38.1		30.5	41.3					37.8	
178,734	T→G (V217G)	hrpB →	ATP-dependent helicase									33.8	
193,139	T→G (T167P)	ROD_01611 ←	hypothetical protein		100			100					
193,141	T→A (Y166F)	ROD_01611 ←	"					32.9					
193,151	T→G (T163P)	ROD_01611 ←	"		41.3	47.3	44.1		42			46.7	
193,156	T→G (H161P)	ROD_01611 ←	"				44.9					34.3	
193,163	T→C (S159G)	ROD_01611 ←	"				44.1						
297,042	A→G (intergenic)	ROD_02491 ← / ← ROD_02501	hypothetical protein/hypothetical protein					25.9					
297,042	A→T (intergenic)	ROD_02491 ← / ← ROD_02501	"					26.8					
327,122	T→A (intergenic)	cI → I → IfiJ	phage repressor protein/flagellar export/assembly protein							37.1			
353,482	A→G (M252V)	lafU →	flagellar motor protein B									32.8	
353,498	T→G (V257G)	lafU →	"		32.4				27.4			28.8	
353,516	T→G (M263R)	lafU →	"						28.6		29.7		
492,511	C→A (A227A)	sbcC ←	exonuclease SbcC			100							
613,718	T→G (V316G)	ROD_05511 →	permease									27.6	
628,648	A→T (intergenic)	ROD_05651 ← / ← ROD_05671	hypothetical								26.5		
650,402	C→G (G270G)	ROD_05911 →	transketolase					31.6					
650,479	A→C (H296P)	ROD_05911 →	"		31.5								
650,532	T→A (*314K)	ROD_05911 →	"								26.7		
672,367	C→G (A129G)	cstA →	carbon starvation protein A			28.8							
672,399	A→C (T140P)	cstA →	"						27.4				
755,854	T→C (E188G)	ROD_06891 ←	hypothetical protein						35.1		34.7		
787,459	T→G (intergenic)	ROD_07151 → / ← abrB	hypothetical protein/ammonia monooxygenase								28.6		
787,954	T→G (T341P)	abrB ←	ammonia monooxygenase	40.8			29.2	29.9					
793,586	C→T (F543F)	sdhA →	succinate dehydrogenase flavoprotein subunit		100								
798,523	A→C (T323P)	sucB →	dihydrolipoamide succinyltransferase component (E2)			25.8							
798,562	A→C (T336P)	sucB →	"	38.1	41.4			37.8			30.8	100	
853,846	A→C (W62G)	ROD_07741 ←	phosphatidylethanolamine-binding protein			30.3							
865,320	A→G (E47G)	moaD →	molybdopterin converting factor, subunit 1	30.5		27.1							
1,027,890	T→G (intergenic)	ROD_09461 → / ← ROD_09471	prophage transcriptional regulator/LuxR family transcriptional regulator					30.2					
1,030,348	A→G (intergenic)	ROD_09461 → / ← ROD_09471	"			25.3							
1,032,441	T→C (intergenic)	ROD_09461 → / ← ROD_09471	"					25.6					
1,039,375	G→T (S219R)	cydC ←	transport ATP-binding protein CydC								100		
1,062,039	A→C (V330G)	ROD_09671 ←	hypothetical protein								100		
1,176,000	A→C (V663G)	putA ←	bifunctional protein PutA				35.2	37.7					
1,236,860	A→G (E267G)	ROD_11451 →	LysR family transcriptional regulator	30.8									
1,236,931	A→C (T291P)	ROD_11451 →	"	37.4			53.3	49.3			56.4		
1,285,412	G→C (N534K)	ROD_11911 ←	adhesin autotransporter	100									
1,329,007	T→G (N154T)	rliE ←	ribosomal large subunit pseudouridine synthase E (rRNA pseudouridylylate synthase E)							100			
1,345,856	A→C (T84P)	ROD_12451 →	aconitate hydratase	31.1									
1,391,177	C→G (G155G)	selD →	Selenide,water dikinase				30.1						
1,406,096	A→C (T281P)	astC →	succinylornithine transaminase			26.3							
1,429,605	T→G (T82P)	pfkB ←	6-phosphofructokinase	35.4			36.3	29.6	27	38.6			
1,429,649	T→C (E67G)	pfkB ←	"							28.1			
1,429,656	T→G (T65P)	pfkB ←	"					46.7	34.3				
1,541,548	A→G (E106G)	ROD_14451 →	oxidoreductase						30.6				
1,599,772	A→G (E28G)	ROD_15031 →	ABC transporter ATP-binding protein							27			
1,599,775	T→G (V29G)	ROD_15031 →	"							31.3	48.2		
1,599,802	T→G (V38G)	ROD_15031 →	"			32.6	33.8	39.2	34.1	34.3	41.4		

Location	Genetic change	Gene affected	Function	W1 _{P20}	W2 _{P20}	W3 _{P20}	W4 _{P20}	W5 _{P20}	N1 _{P20}	N2 _{P20}	N3 _{P20}	N4 _{P20}	N5 _{P20}	
1,645,594	T→C (L184P)	<i>dos</i> →	heme-regulated cyclic AMP phosphodiesterase (direct oxygen sensor protein)	100		100								
1,700,546	T→G (T343P)	<i>ROD_15951</i> ←	GntR family transcriptional regulator	31.7										
1,700,605	A→C (V323G)	<i>ROD_15951</i> ←	"	26.4										
1,746,363	A→C (V665G)	<i>ROD_16371</i> ←	hypothetical protein		33.9			37.7			34.3			
1,746,394	T→C (T655A)	<i>ROD_16371</i> ←	"								32.8			
1,746,396	A→C (V654G)	<i>ROD_16371</i> ←	"	36.1			42.8							
1,746,409	T→C (T650A)	<i>ROD_16371</i> ←	"							52.2				
1,746,421	A→C (W646G)	<i>ROD_16371</i> ←	"					49.2						
1,749,372	T→C (intergenic)	<i>ROD_16381</i> → / → <i>ROD_16391</i>	hypothetical protein/outer membrane autotransporter							100				
1,762,318	9 bp deletion (intergenic)	<i>uspF</i> → / ← <i>nleC</i>	universal stress protein F/T3SS effector protein NleC							71.8				
1,845,380	T→C (E257G)	<i>rluB</i> ←	ribosomal large subunit pseudouridine synthase B							100				
1,892,282	T→G (Y78S)	<i>narG</i> ←	respiratory nitrate reductase 1 subunit alpha									100		
1,896,011	C→T (A440V)	<i>narX</i> →	nitrate/nitrite two-component sensor kinase							100				
2,019,363	+C insertion	<i>znuA</i> ←	high-affinity zinc uptake system protein			88.6			91.6					
2,141,190	G→T (L50F)	<i>ROD_20381</i> →	hypothetical protein					27.3						
2,141,225	T→C (L62P)	<i>ROD_20381</i> →	"	31.4										
2,168,671	G→C (R239P)	<i>ROD_20671</i> →	integrase		38.2									
2,168,700	A→G (S249G)	<i>ROD_20671</i> →	"		29.4						43.7			
2,168,710	A→T (D252V)	<i>ROD_20671</i> →	"							29.9				
2,168,720	C→A (N255K)	<i>ROD_20671</i> →	"		42.4			44.7	44.5		36.8			
2,169,843	A→G (E46G)	<i>ROD_20681</i> →	IS102 transposase			33.4								
2,169,868	G→A (S54S)	<i>ROD_20681</i> →	"			36.4								
2,169,880	A→T (S58S)	<i>ROD_20681</i> →	"			41.2								
2,169,889	A→C (G61G)	<i>ROD_20681</i> →	"			47.6								
2,169,892	G→A (R62R)	<i>ROD_20681</i> →	"			50.9								
2,214,415	A→G (I106T)	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase							100				
2,214,427	T→C (E102G)	<i>cbiH</i> ←	"			32.2	35.3	51.7	44.6		44.6	39.9		
2,214,430	A→C (V101G)	<i>cbiH</i> ←	"			48.7	46	46.3	37.4		44.9			
2,214,433	T→C (D100G)	<i>cbiH</i> ←	"								29.2			
2,214,457	T→C (E92G)	<i>cbiH</i> ←	"	35.4	37.9	41.5	46			43.3	50			
2,214,461	G→C (L91V)	<i>cbiH</i> ←	"		28.1									
2,214,463	A→C (V90G)	<i>cbiH</i> ←	"				41.5							
2,214,467	G→C (L89V)	<i>cbiH</i> ←	"			37.1					27.5			
2,286,893	G→C (R269G)	<i>fcl</i> ←	GDP-L-fucose synthetase				30.7							
2,286,907	A→C (V264G)	<i>fcl</i> ←	"		34.9	34.4	33.8	45.2		30.9				
2,286,910	A→C (V263G)	<i>fcl</i> ←	"	39.7	31.2						32			
2,286,931	A→C (L256R)	<i>fcl</i> ←	"						26.4					
2,286,934	T→C (E255G)	<i>fcl</i> ←	"						33.6					
2,286,940	A→G (I253T)	<i>fcl</i> ←	"					45.4						
2,294,346	T→C (Q611R)	<i>wzc</i> ←	tyrosine-protein kinase							28.8				
2,329,361	A→G (R332R)	<i>ROD_22201</i> ←	ABC transporter substrate-binding protein								26.4			
2,359,115	A→C (intergenic)	<i>ROD_22461</i> ← / ← <i>ROD_22491</i>	major facilitator superfamily transporter/hypothetical protein	25.3										
2,379,722	T→G (T112P)	<i>ROD_22651</i> ←	FAA-hydrolase-family protein					37.2						
2,406,607	T→A (I96F)	<i>ROD_22911</i> ←	nucleoside transporter	25.2	29.4	29.8				25.1	27.1	30.2		
2,406,664	T→A (I77F)	<i>ROD_22911</i> ←	"											
2,406,707	G→C (G62G)	<i>ROD_22911</i> ←	"		25.8									
2,406,723	T→C (E57G)	<i>ROD_22911</i> ←	"				38.7				44.2			
2,406,726	A→C (V56G)	<i>ROD_22911</i> ←	"								33.7			
2,406,743	T→G (P50P)	<i>ROD_22911</i> ←	"								36.4			
2,438,536	C→T (P172S)	<i>ROD_23201</i> →	sulfatase			100				100		100		
2,458,789	T→C (S454G)	<i>mgtE</i> ←	magnesium transporter						25.9					

Location	Genetic change	Gene affected	Function	W1 _{p20}	W2 _{p20}	W3 _{p20}	W4 _{p20}	W5 _{p20}	N1 _{p20}	N2 _{p20}	N3 _{p20}	N4 _{p20}	N5 _{p20}
2,517,503	A→C (V46G)	<i>eutK</i> ←	ethanolamine utilization protein	28.4									
2,522,793	A→G (S149P)	<i>eutH</i> ←	ethanolamine transporter							35.4			
2,524,686	A→C (V203G)	<i>eutJ</i> ←	ethanolamine utilization protein		25.2	34.2							
2,524,689	A→C (V202G)	<i>eutJ</i> ←	"			34.1							
2,524,731	G→C (A188G)	<i>eutJ</i> ←	"						25				
2,525,977	A→C (V244G)	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase						33.5				31.4
2,525,980	A→C (V243G)	<i>eutE</i> ←	"		33.7				47.2				
2,525,983	A→C (V242G)	<i>eutE</i> ←	"				27.2						
2,526,033	A→T (R225R)	<i>eutE</i> ←	"					26.3					
2,554,865	A→C (intergenic)	<i>dapA</i> ← / → <i>gcvR</i>	dihydrodipicolinate synthase/glycine cleavage system transcriptional repressor				29.3		47.7			37.2	
2,560,330	A→C (T189P)	<i>ROD_24351</i> →	peptidase						31.9				
2,625,748	A→C (T90P)	<i>ROD_24921</i> →	ROK-family regulatory protein	37.7									
2,680,561	G→C (A131G)	<i>cipB</i> ←	chaperone (heat-shock protein F84.1)		34.8	26							
2,835,878	C→T (R216H)	<i>nuoG</i> ←	NADH-quinone oxidoreductase subunit G								100		
2,851,098	A→C (intergenic)	<i>ackA</i> → / → <i>pta</i>	acetate kinase/phosphate acetyltransferase								100		
2,913,821	T→C (K293R)	<i>ROD_27631</i> ←	outer membrane autotransporter					26.9					
2,913,832	T→G (A289A)	<i>ROD_27631</i> ←	"										33.5
2,913,847	A→C (G284G)	<i>ROD_27631</i> ←	"		100				100			36.1	
2,932,432	G→C (A336G)	<i>ROD_27781</i> ←	fimbrial usher protein	25.4									
2,932,435	T→C (E335G)	<i>ROD_27781</i> ←	"							25.6			36.7
2,932,441	A→C (V333G)	<i>ROD_27781</i> ←	"	29.8						29.4			
2,932,447	A→C (V331G)	<i>ROD_27781</i> ←	"					30.4	29.4		52.4		
3,086,114	T→G (V843G)	<i>fimD</i> →	fimbrial usher protein FimD						26.4	26.4			
3,086,120	T→G (V845G)	<i>fimD</i> →	"						36.8				
3,133,163	A→G (K646K)	<i>ROD_29581</i> →	large repetitive protein										33.1
3,147,079	A→G (P236P)	<i>espF</i> ←	T3SS effector protein EspF	26.7									
3,147,094	A→G (S231S)	<i>espF</i> ←	"								25.9		
3,147,187	A→G (A200A)	<i>espF</i> ←	"				42.7			26.2			
3,147,408	A→G (S127P)	<i>espF</i> ←	"					25.9	27				
3,155,158	A→C (intergenic)	<i>escD</i> → / ← <i>eae</i>	T3SS structural protein EscD/intimin	100									
3,180,461	1 bp deletion (intergenic)	<i>ler</i> ← / ← <i>ROD_30161</i>	Ler transcriptional regulator/ISCr01 transposase C			100							
3,229,848	A→C (T428P)	<i>cysN</i> →	Sulfate adenyllyltransferase subunit 1		31.9				31.4				
3,256,765	A→G (S461G)	<i>hycC</i> →	formate hydrogenlyase subunit 3	27.5		30.3							
3,403,644	112 bp x 2 duplication	[<i>ROD_t54</i>]	[<i>ROD_t54</i>]			33.1				40.6		34	
3,403,833	+111 or 224 bp insertion (intergenic)	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	32.9		65.2							
3,403,834	T→A (intergenic)	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	"		35.3			31.4	37.3			26.6	
3,403,837	+223 bp insertion (intergenic)	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	"		37.6			28.4	40.9			33.3	
3,526,665	G→A (P63P)	<i>ROD_33401</i> ←	IS102 transposase	53.3		55.8		56.2		61.1			
3,526,668	C→T (R62R)	<i>ROD_33401</i> ←	"	50						57.9		100	
3,526,668	C→T (R62R)	<i>ROD_33401</i> ←	"					46.3					
3,526,671	T→G (G61G)	<i>ROD_33401</i> ←	"	50.3				47.6		57.9			
3,526,680	T→A (S58S)	<i>ROD_33401</i> ←	"	44.5				39.9	45.3	52.4			
3,526,692	C→T (S54S)	<i>ROD_33401</i> ←	"	40				34.2	33.8	48.1			
3,526,717	T→C (E46G)	<i>ROD_33401</i> ←	"	30.7				31.2		40.8			
3,526,789	C→T (R22K)	<i>ROD_33401</i> ←	"						25				
3,598,416	G→A (E33K)	<i>cts2S</i> →	T6SS protein Cts2S							28.9			
3,598,429	T→G (V37G)	<i>cts2S</i> →	"	31						37.9	32.2		
3,598,432	T→G (V38G)	<i>cts2S</i> →	"						26.8	35.5	34.8	27.6	
3,629,017	A→T (I33F)	<i>ROD_34401</i> →	dimethyl sulfoxide reductase							29.1			
3,649,770	C→T (intergenic)	<i>acs</i> ← / → <i>nrfA</i>	acetyl-coenzyme A synthetase/cytochrome c552							25.4			

Location	Genetic change	Gene affected	Function	W1 _{p20}	W2 _{p20}	W3 _{p20}	W4 _{p20}	W5 _{p20}	N1 _{p20}	N2 _{p20}	N3 _{p20}	N4 _{p20}	N5 _{p20}	
3,671,444	C→G (A119G)	<i>ROD_34781</i> →	hypothetical protein		28.1									
3,671,446	T→G (W120G)	<i>ROD_34781</i> →	"			28.5		36.8						
3,671,467	A→C (T127P)	<i>ROD_34781</i> →	"	31.6		25.6			30					31.1
3,671,492	A→T (Y135F)	<i>ROD_34781</i> →	"						25.3					
3,671,499	C→T (F137F)	<i>ROD_34781</i> →	"			29								
3,750,685	A→C (V676G)	<i>ROD_35621</i> ←	radical SAM superfamily protein			26.9								
3,777,356	C→T (L21L)	<i>zupT</i> →	zinc transporter									27.3		
3,777,402	C→T (S36L)	<i>zupT</i> →	"	27.5										
3,809,617	A→C (V76G)	<i>gcp</i> ←	O-sialoglycoprotein endopeptidase (glycoprotease)							27.5				
3,952,015	A→G (E253G)	<i>thiG</i> →	thiazole biosynthesis protein			28		25.6						
3,996,563	G→A (R148C)	<i>pflC</i> ←	pyruvate formate-lyase 2 activating enzyme			100								
4,034,122	T→A (L430F)	<i>ROD_38291</i> ←	hypothetical protein						26.7					
4,034,138	T→A (Y425F)	<i>ROD_38291</i> ←	"						26.5					
4,091,214	T→G (V110G)	<i>ROD_38821</i> →	aldose 1-epimerase			26.9								
4,099,272	A→C (T66P)	<i>ROD_38891</i> →	hypothetical protein			29.3								
4,099,344	A→T (I90F)	<i>ROD_38891</i> →	"	30.1										
4,173,823	T→C (E217G)	<i>ROD_39511</i> ←	amino acid permease						33.6					
4,173,826	A→C (V216G)	<i>ROD_39511</i> ←	"		25.3									
4,177,633	C→T (A72T)	<i>rffT</i> ←	4-alpha-L-fucosyltransferase								27.6			
4,196,514	T→G (V89G)	<i>ROD_39711</i> →	permease							25.5				
4,288,254	A→G (G543G)	<i>ROD_40471</i> →	sodium:solute symporter	28.6	42.4	35.6						33.1		
4,288,265	A→T (Y547F)	<i>ROD_40471</i> →	"			31.8	31.6	37.8						
4,288,267	C→T (L548L)	<i>ROD_40471</i> →	"						32.7			27.6		
4,288,269	G→T (L548L)	<i>ROD_40471</i> →	"		46.6	26.9	33.1	43.3		34		37.4		
4,288,282	T→G (W553G)	<i>ROD_40471</i> →	"						36.9					
4,297,770	A→C (V262G)	<i>ROD_40571</i> ←	hypothetical protein			40.6								
4,376,649	T→G (V8G)	<i>ROD_41482</i> →	hypothetical protein		26.8									
4,465,504	A→T (G186G)	<i>ROD_42331</i> →	AraC family transcriptional regulator	66.6										
4,465,506	A→T (E187V)	<i>ROD_42331</i> →	"	66.6										
4,520,106	T→G (V1066G)	<i>bcsC</i> →	cellulose synthase operon protein C (TPR-repeat-containing protein)									32.8		
4,567,745	A→C (V167G)	<i>nikE</i> ←	nickel ABC transporter ATP-binding protein		25.2									
4,604,024	T→G (V136G)	<i>ftsY</i> →	cell division protein						27.2		34.1	30.3		
4,622,322	1 bp deletion	<i>ugpA</i> →	glycerol-3-phosphate ABC transporter permease							100				
4,627,516	T→G (Y581D)	<i>ggt</i> →	gamma-glutamyltranspeptidase			100								
4,671,835	A→C (V91G)	<i>greB</i> ←	transcription elongation factor	28.1	31.8	28.4			49					
4,671,839	A→C (W90G)	<i>greB</i> ←	"		41.1				34.8		30.3			
4,703,780	T→G (V73G)	<i>pabA</i> →	para-aminobenzoate synthase glutamine amidotransferase component I									25.6		
4,754,701	A→G (P55P)	<i>gspC</i> ←	T2SS protein C							27.7				
4,880,554	C→G (G63G)	<i>ROD_46251</i> →	organic solvent tolerance ABC-transporter ATP-binding component		27.7									
4,955,826	T→G (G114G)	<i>ROD_47081</i> →	hypothetical protein					25						
4,987,380	G→T (G240G)	<i>gatC</i> ←	component IIC of galactitol-specific phosphotransferase system							100				
5,012,789	T→G (V340G)	<i>ROD_47661</i> →	hypothetical protein								100			
+TTATTCC / or +14bp repeat insertion (intergenic)		<i>espN2-2</i> ← / ← <i>ROD_47851</i>	T3SS effector protein EspN2-2/hypothetical protein	96.9							97.6			
5,035,744	7 bp deletion (intergenic)	<i>espN2-2</i> ← / ← <i>ROD_47851</i>	"		100									
5,068,776	A→G (intergenic)	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS			32.8	53.3							
5,068,782	A→C (intergenic)	<i>ROD_48151</i> → / → <i>ROD_48181</i>	"	34.8		31.9	49.3							
5,082,494	T→G (intergenic)	<i>ROD_48151</i> → / → <i>ROD_48181</i>	"	29.4					48.7	40.3		36.7	35.1	
5,082,495	T→G (intergenic)	<i>ROD_48151</i> → / → <i>ROD_48181</i>	"										25	
5,086,717	A→G (I162V)	<i>ROD_48191</i> →	membrane fusion protein of T1SS	100										
5,108,248	T→A (I140F)	<i>ROD_48371</i> ←	carbon starvation protein			28.9								
5,108,256	A→C (V137G)	<i>ROD_48371</i> ←	"							25.8				
5,108,271	A→C (V132G)	<i>ROD_48371</i> ←	"			34.1								

Location	Genetic change	Gene affected	Function	W1 _{P20}	W2 _{P20}	W3 _{P20}	W4 _{P20}	W5 _{P20}	N1 _{P20}	N2 _{P20}	N3 _{P20}	N4 _{P20}	N5 _{P20}
5,175,929	T→G (V409G)	<i>radA</i> →	DNA repair protein								39.8		
5,175,934	A→G (S411G)	<i>radA</i> →	"		39.1	35.6				41.8			
5,175,944	A→G (E414G)	<i>radA</i> →	"								42.1		
5,318,302	T→G (intergenic)	<i>ROD_50641 ← / → epd</i>	IS102 transposase/d-erythrose-4-phosphate dehydrogenase				34.2					33	
5,318,308	T→A (intergenic)	<i>ROD_50641 ← / → epd</i>	"								34.9		

7.2.2 Genes with multiple genetic changes present

To investigate any common themes present in the types of genes commonly changed in the *in vivo*-adapted strains, I narrowed the list of SNPs present at least 25% frequency (Table 7.1) to show genes with multiple SNPs identified; that is, genes with more than one genetic change that may or may not be present in multiple lineages. Out of the *C. rodentium* genes which have evidence of multiple genetic events (Table 7.2), a large number of these encode for incompletely characterised proteins: *ROD_01611*, *ROD_20381*, *ROD_34781*, *ROD_38291*, and *ROD_38891*, which each encode for putative membrane proteins with as yet unknown functions, and when assessed using the BLAST tool^{228–230} against the UniProtKB database²³¹ only show similarity to uncharacterised proteins and predicted membrane proteins found in many Gram negative bacterial species. Other incompletely characterised proteins with multiple genetic changes identified are *ROD_16371*, a putative exported protein with similarity to many uncharacterised proteins found in many Gram negative bacterial species, and *ROD_27631*, a putative outer membrane autotransporter with some similarity to other autotransporters (the closest matches are 86.0% similarity to an outer membrane autotransporter found in *Enterobacter cancerogenus*, 73.4% similarity to a membrane protein in *Enterobacter* sp. GN02174, and 61.3% similarity to an outer membrane autotransporter in *Serratia plymuthica* PRI-2C^{228–231}).

I also identified a number of genes with multiple genetic changes which have suggested functions based on sequence similarity, including *ROD_11451* (a putative LysR-family transcriptional regulator; 2 changes), *ROD_15031* (a putative ABC transport ATP-binding subunit; 3 changes), *ROD_15951* (a putative GntR-family transcriptional regulator; 2 changes), *ROD_22911* (a probable nucleoside permease; 6 changes), *ROD_39511* (probable amino acid permease; 2 changes), *ROD_40471* (a putative sodium:solute symporter; 5 changes), *ROD_42331* (a AraC-family transcriptional regulator; 2 changes), and *ROD_48371* (a putative carbon starvation protein; 3 changes).

There are also genes present with more well characterised functions. *ROD_01271*, which encodes for 2-keto-3-deoxygluconate permease, is known to be involved in transportation of pectin products into the cell in other species (2 changes)^{232,233}; *lafU*, which encodes for lateral flagellar motor protein B (3 changes)²³⁴; *ROD_05911*, which encodes for a transketolase, known in other species to be important for linking the pentose phosphate pathway and glycolysis (3 changes)²³⁵; *cstA*, which encodes for carbon starvation protein A and is important in the stress response to carbon starvation (2 changes)²³⁶; *sucB*, which encodes for the dihydrolipoamide succinyltransferase component of 2-oxoglutarate dehydrogenase complex, which 2-oxoglutarate dehydrogenase complex catalyzes conversion of 2-oxoglutarate to succinyl-CoA and CO₂ (2 changes)²³⁷; *pfkB*, which encodes for phosphofructokinase, which is an important enzyme for glycolysis (3 changes)²³⁸; *ROD_20671*, which encodes for an integrase (4 changes); *ROD_20681* and *ROD_33401*, which each encode for IS102 transposases (5 and 8 changes, respectively); *cbiH*, which encodes for

cobalt-precorrin-3B C(17)-methyltransferase, and likely catalyzes the methylation of C17 in cobalt-precorrin-3B (8 changes)²³⁹; *fcl*, which encodes for GDP-L-fucose synthase, which has been shown in other species to be involved in the NADP-dependent conversion of GDP-4-dehydro-6-deoxy-D-mannose to GDP-fucose and NADPH, however this GDP mannose biosynthesis pathway in *C. rodentium* appears incomplete, with no recorded GDP-mannose 4,6-dehydratase enzyme present, which is required for the initial step of converting GDP-D-mannose to GDP-4-dehydro-6-deoxy-D-mannose (6 changes)²⁴⁰; *eutJ* and *eutE*, encoding for an ethanolamine utilization protein and ethanolamine utilization aldehyde dehydrogenase, respectively, each which have roles in ethanolamine degradation (3 and 4 changes, respectively)²⁴¹; *ROD_2778I*, which encodes for a fimbrial usher protein (4 changes); *fimD*, which encodes for the fimbrial usher protein FimD, which is part of the *fim* operon and is necessary in *E. coli* for localisation of type 1 fimbriae on the surface of bacteria (2 changes)^{242,243}; *espF*, encoding for the T3SS effector protein EspF which is found on the locus of enterocyte effacement pathogenicity island, on the LEE4 operon, is involved in disruption of tight junctions of enterocytes and induction of apoptosis, but is not required for colonisation of the mouse colon (4 changes)^{128,129,244}; *cts2S*, which encodes for a T6SS protein with an unknown function – *cts1S*, however, has been shown to play a role in interbacterial competition (3 changes)²⁴⁵; *zupT*, which encodes for the zinc transporter ZupT and mediates zinc uptake (2 changes)²⁴⁶; *greB*, which encodes for the transcription elongation factor GreB, and has been shown in other species to be necessary for efficient RNA polymerase transcription elongation (2 changes)²⁴⁷, and *radA*, which encodes for the DNA repair protein RadA (3 changes)²⁴⁸.

Table. 7.2. List of genes with multiple genetic events

Gene	Product	Changes*	Present in strains**									
			W1 P20	W2 P20	W3 P20	W4 P20	W5 P20	N1 P20	N2 P20	N3 P20	N4 P20	N5 P20
<i>cbiH</i>	Cobalt-precorrin-3B C(17)-methyltransferase	8	■									
<i>ROD_40471</i>	Sodium:solute symporter	5	■	■								■
<i>ROD_22911</i>	Nucleoside transporter	6		■								
<i>ROD_01611</i>	Hypothetical protein	5										■
<i>ROD_33401</i>	IS102 transposase	8	■		■						■	
<i>fcl</i>	GDP-L-fucose synthetase	6		■	■						■	
<i>ROD_34781</i>	Hypothetical protein	5		■	■				■			■
<i>ROD_16371</i>	Hypothetical protein	5	■		■					■	■	
<i>ROD_27781</i>	Fimbrial usher protein	4					■			■		
<i>pfkB</i>	6-phosphofructokinase	3	■				■	■	■	■		
<i>ROD_15031</i>	ABC transporter ATP-binding protein	3				■						■
<i>sucB</i>	Dihydrolipoamide succinyltransferase component (E2)	2		■	■			■				■
<i>ROD_20671</i>	Integrase	4				■			■	■		
<i>eutE</i>	Ethanolamine utilization aldehyde dehydrogenase	4			■	■	■				■	
<i>espF</i>	T3SS effector protein EspF	4			■		■			■		
<i>lafU</i>	Flagellar motor protein B	3	■						■	■		
<i>cts2S</i>	T6SS protein Cts2S	3	■									■
<i>ROD_27631</i>	Outer membrane autotransporter	3		■			■		■			■
<i>ROD_01271</i>	2-keto-3-deoxygluconate permease	2	■				■		■	■		■
<i>greB</i>	Transcriptional elongation factor	2		■	■	■			■	■		
<i>radA</i>	DNA repair protein	3	■	■					■	■		
<i>ROD_11451</i>	LysR family transcriptional regulator	2	■					■	■			■
<i>eutJ</i>	Ethanolamine utilization protein	3			■	■			■			
<i>ROD_05911</i>	Transketolase	3		■	■							■
<i>ROD_48371</i>	Carbon starvation protein	3				■				■		
<i>cstA</i>	Carbon starvation protein A	2				■				■		
<i>ROD_20381</i>	Hypothetical protein	2	■					■				
<i>ROD_39511</i>	Amino acid permease	2			■	■			■			
<i>ROD_38891</i>	Hypothetical protein	2	■		■							
<i>fimD</i>	Fimbrial usher protein FimD	2							■	■		
<i>zupT</i>	Zinc transporter	2	■							■		
<i>ROD_42331</i>	AraC family transporter regulator	2	■									
<i>ROD_15951</i>	GntR family transcriptional regulator	2	■									
<i>ROD_20681</i>	IS102 transposase	5					■					
<i>ROD_38291</i>	Hypothetical protein	2						■				

* Number of types of SNPs identified in a given gene.

** Presence of at least one type of SNP in a given gene in the listed strains is marked by a green filled box.

7.2.3 Intragenic SNPs constitute the majority of SNPs and are found throughout the genome

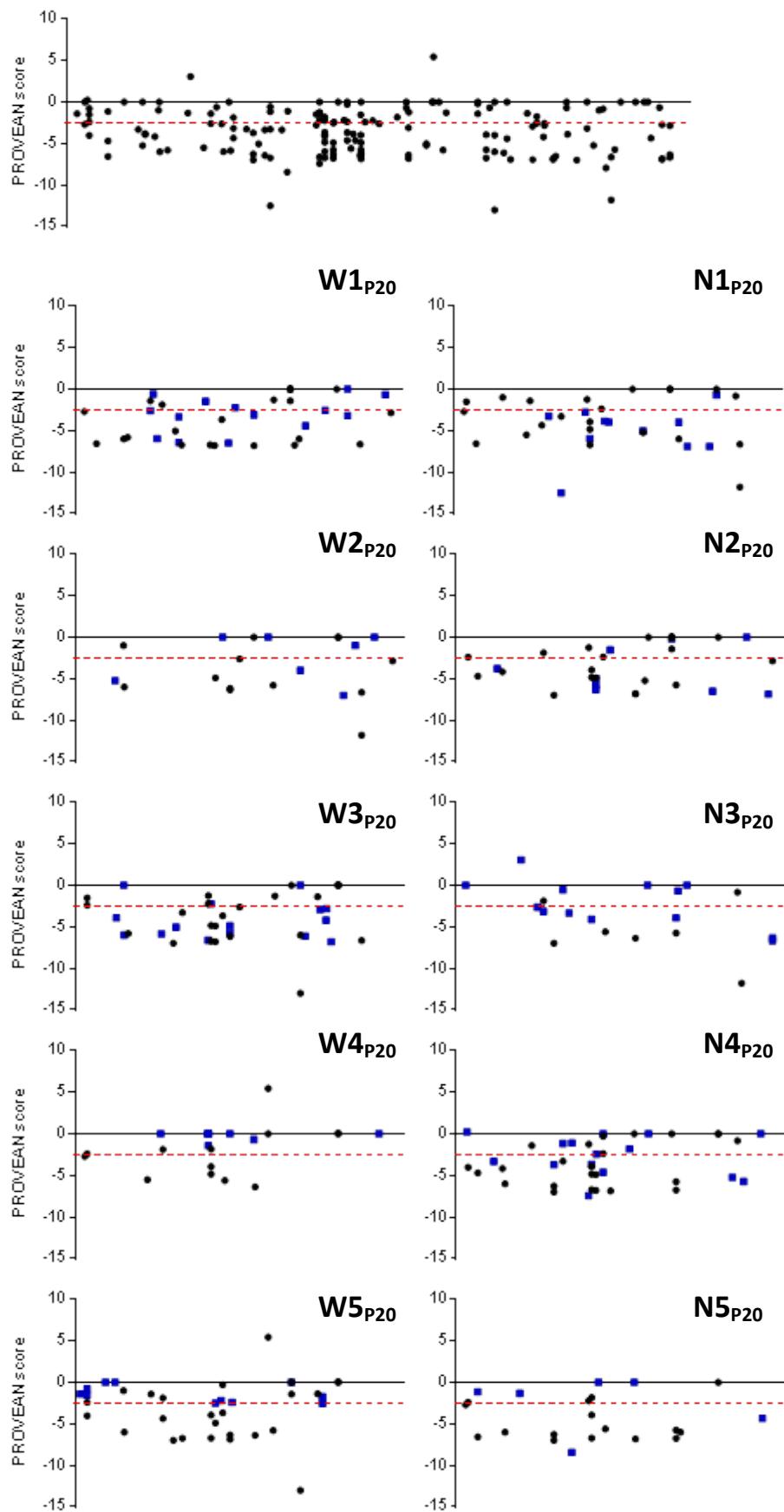
I observed a similar total number of intragenic SNPs for each of the strains, which also constitutes the majority of the type of SNPs for each strain (Table 7.3). Of these intragenic SNPs, a similar percentage is unique to each strain, with the sole exception of N3_{P20} which has more unique SNPs (65%) than shared SNPs. PROVEAN scores for each of the intragenic SNPs were compared to the location of each SNP along the *C. rodentium* genome (Fig. 7.2). Scores of less than -2.5, which were predicted to have a deleterious or non-neutral effect on the protein, are indicated as data points below the dashed red line. The location of intragenic SNPs is spread out throughout the genome, with no large conserved areas which do not have genetic changes present (Fig. 7.2). For each of the individual strains, the present intragenic SNPs and their associated PROVEAN score is shown, with SNPs which are shared with at least one other lineage shown in black and the blue squares indicating intragenic SNPs unique to that particular strain.

Table 7.3. Percentage of intragenic and unique intragenic SNPs.

Lineage	Total SNPs	Intragenic SNPs (% of total)	Unique intragenic SNPs (% of intragenic)	dN/dS value
W1 _{P20}	43	38 (88%)	14 (37%)	4.429
W2 _{P20}	24	20 (83%)	7 (35%)	2.332
W3 _{P20}	50	38 (76%)	15 (39%)	7.600
W4 _{P20}	25	22 (88%)	9 (41%)	1.443
W5 _{P20}	44	38 (86%)	11 (29%)	4.571
N1 _{P20}	38	33 (87%)	11 (33%)	7.250
N2 _{P20}	32	32 (100%)	9 (28%)	3.000
N3 _{P20}	29	20 (69%)	13 (65%)	6.000
N4 _{P20}	45	39 (87%)	15 (38%)	4.571
N5 _{P20}	26	22 (85%)	6 (27%)	6.667

Figure 7.1. PROVEAN scores for all SNPs and unique SNPs (overleaf). SNPs were assessed for predicted changes to protein function using the Protein Variation Effect Analyzer (PROVEAN) tool²⁴⁹⁻²⁵¹. Any SNP which has a score of -2.5 or lower (indicated by the dashed red line) was predicted to result in a change to the protein function. The x axis represents location on the genome (from 0 to 5,347,979). SNPs present in each strain is also shown (W1_{P20}-W5_{P20} and N1_{P20}-N5_{P20}), with SNPs unique for each strain shown as blue squares and SNPs shared with at least one other strain shown at black circles.

Intragenic SNPs from all *in vivo*-adapted strains



7.2.4 Genetic changes in *C. rodentium* adapted to the mouse environment

From a filtered list, I identified a total of 133 SNPs present in the *C. rodentium* which adapted in mice in the absence of antibiotic (W1_{P20}-W5_{P20}); 83 of these are only present in W1_{P20}-W5_{P20} *C. rodentium*, and 67 of these are unique to a single lineage. The W1_{P20} strain has a total of 43 SNPs (Table 7.4), with 16 unique and 27 SNPs shared with at least one other lineage. 38 of the SNPs are intragenic and 5 intergenic; of the intragenic SNPs, 31 are non-synonymous and 7 synonymous. The resulting d_N/d_S ratio for W1_{P20} is 4.429, indicating positive selection. In addition, there is a 111-224 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein).

Intragenic SNPs were investigated for predicted functional changes to their coding gene using the Protein Variation Effect Analyzer (PROVEAN) web based tool ^{249–251}. Scores of -2.5 or below are predicted to have a deleterious or non-neutral effect on the protein, however whether this change would result in a non-functional protein or a protein with an altered function was not determined.

Table 7.4. Mutations in W1_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→C (S254P)	-2.700	<i>ROD_01271</i> →	2-keto-3-deoxygluconate permease	
T→G (V257G)	-6.542	<i>lafU</i> →	flagellar motor protein B	
A→C (T336P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	
A→G (E47G)	-5.808	<i>moaD</i> →	molybdopterin converting factor, subunit 1	
A→G (E267G)	-2.599	<i>ROD_11451</i> →	LysR family transcriptional regulator	✓
A→C (T291P)	-1.400	<i>ROD_11451</i> →	“	
G→C (N534K)	-0.624	<i>ROD_11911</i> ←	adhesin autotransporter	✓
A→C (T84P)	-5.967	<i>ROD_12451</i> →	aconitate hydratase	✓
T→G (T82P)	-1.868	<i>pfkB</i> ←	6-phosphofructokinase	
T→C (L184P)	-5.032	<i>dos</i> →	heme-regulated cyclic AMP phosphodiesterase (direct oxygen sensor protein)	
T→G T343P)	-3.354	<i>ROD_15951</i> ←	GntR family transcriptional regulator	✓
A→C (V323G)	-6.427	<i>ROD_15951</i> ←	“	✓
A→C (V654G)	-6.727	<i>ROD_16371</i> ←	hypothetical protein	
T→C (L62P)	-1.467	<i>ROD_20381</i> →	hypothetical protein	✓
T→C (E92G)	-6.708	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
A→C (V263G)	-6.772	<i>fcl</i> ←	GDP-L-fucose synthetase	
A→C (intergenic)	N/A	<i>ROD_22461</i> ← / ← <i>ROD_22491</i>	major facilitator superfamily transporter/hypothetical protein	✓
T→A (I96F)	-3.655	<i>ROD_22911</i> ←	nucleoside transporter	
A→C (V46G)	-6.493	<i>eutK</i> ←	ethanolamine utilization protein	✓
A→C (T90P)	-2.237	<i>ROD_24921</i> →	ROK-family regulatory protein	✓
G→C (A336G)	-3.088	<i>ROD_27781</i> ←	fimbrial usher protein	✓
A→C (V333G)	-6.797	<i>ROD_27781</i> ←	“	
A→C (intergenic)	N/A	<i>escD</i> → / ← <i>eae</i>	T3SS structural protein EscD/intimin	✓

A→G (S461G)	-1.272	<i>hycC</i> →	formate hydrogenlyase subunit 3	
111/224 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
G→A (P63P)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
C→T (R62R)	0.000	<i>ROD_33401</i> ←	“	
T→G (G61G)	0.000	<i>ROD_33401</i> ←	“	
T→A (S58S)	0.000	<i>ROD_33401</i> ←	“	
C→T (S54S)	0.000	<i>ROD_33401</i> ←	“	
T→C (E46G)	-1.400	<i>ROD_33401</i> ←	“	
T→G (V37G)	-6.750	<i>cts2S</i> →	T6SS protein Cts2S	
A→C (T127P)	-6.000	<i>ROD_34781</i> →	hypothetical protein	
C→T (S36L)	-4.410	<i>zupT</i> →	zinc transporter	✓
A→T (I90F)	-2.562	<i>ROD_38891</i> →	hypothetical protein	✓
A→G (G543G)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
A→T (G186G)	0.000	<i>ROD_42331</i> →	AraC family transcriptional regulator	✓
A→T (E187V)	-3.186	<i>ROD_42331</i> →	“	✓
A→C (V91G)	-6.633	<i>greB</i> ←	transcription elongation factor	
A→C (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	“	
A→G (I162V)	-0.663	<i>ROD_48191</i> →	membrane fusion protein of T1SS	✓
A→G (S411G)	-2.842	<i>radA</i> →	DNA repair protein	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The W2_{P20} strain has a total of 24 SNPs (Table 7.5), with 7 unique and 17 SNPs shared with at least one other lineage. 20 of the SNPs are intragenic and 4 intergenic; of the intragenic SNPs, 14 are non-synonymous and 6 synonymous. The resulting d_N/d_S ratio for W2_{P20} is 2.332, indicating positive selection. In addition, there is a 223 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein) present as well as a 7-14 bp TTATTCC repeat inserted in the intergenic space between *espN2-2* and *ROD_47851* (type 3 secretion system effector protein EspN2-2 and a hypothetical protein).

Table 7.5. Mutations in W2_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
A→C (H296P)	-5.247	<i>ROD_05911</i> →	transketolase	✓
T→G (T341P)	-0.982	<i>abrB</i> ←	ammonia monooxygenase	
A→C (T336P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	
A→C (V264G)	-4.892	<i>fcl</i> ←	GDP-L-fucose synthetase	
G→C (G62G)	0.000	<i>ROD_22911</i> ←	nucleoside transporter	✓
A→C (V203G)	-6.150	<i>eutJ</i> ←	ethanolamine utilization protein	
A→C (V243G)	-6.333	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase	
G→C (A131G)	-2.611	<i>clpB</i> ←	chaperone (heat-shock protein F84.1)	
A→C (G284G)	0.000	<i>ROD_27631</i> ←	outer membrane autotransporter	
A→G (P236P)	0.000	<i>espF</i> ←	T3SS effector protein EspF	✓
A→C (T428P)	-5.791	<i>cysN</i> →	Sulfate adenylyltransferase subunit 1	
T→A (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
223 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	"	
C→G (A119G)	-4.000	<i>ROD_34781</i> →	hypothetical protein	✓
A→G (G543G)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
G→T (L548L)	0.000	<i>ROD_40471</i> →	"	
T→G (V8G)	-7.000	<i>ROD_41482</i> →	hypothetical protein	✓
A→C (V167G)	-0.985	<i>nikE</i> ←	nickel ABC transporter ATP-binding protein	✓
A→C (V91G)	-6.633	<i>greB</i> ←	transcription elongation factor	
A→C (W90G)	-11.777	<i>greB</i> ←	"	
C→G (G63G)	0.000	<i>ROD_46251</i> →	organic solvent tolerance ABC-transporter ATP-binding component	✓
TTATTCC/14 bp repeat insertion (intergenic)	N/A	<i>espN2-2</i> ← / ← <i>ROD_47851</i>	T3SS effector protein EspN2-2/hypothetical protein	
A→G (S411G)	-2.842	<i>radA</i> →	DNA repair protein	
T→G (intergenic)	N/A	<i>ROD_50641</i> ← / → <i>epd</i>	IS102 transposase/d-erythrose-4-phosphate dehydrogenase	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The W3_{P20} strain has a total of 50 SNPs (Table 7.6), with 22 unique and 28 SNPs shared with at least one other lineage. 43 of the SNPs are intragenic and 7 intergenic; of the intragenic SNPs, 38 are non-synonymous and 5 synonymous. The resulting d_N/d_S ratio for W3_{P20} is 7.600, indicating positive selection. In addition, a 111-224 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein) is present as well as a 1 bp deletion in the intergenic space between *ler* and *ROD_30161* (Ler transcriptional regulator and ISCr01 transposase C), and a 112 bp duplication of part of *ROD_t54* (tRNA-Gly).

Table 7.6. Mutations in W3_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→G (T167P)	-1.512	<i>ROD_01611</i> ←	hypothetical protein	
T→G (T163P)	-2.383	<i>ROD_01611</i> ←	"	
C→G (A129G)	-3.934	<i>cstA</i> →	carbon starvation protein A	✓
C→T (F543F)	0.000	<i>sdhA</i> →	succinate dehydrogenase flavoprotein subunit	✓
A→C (T323P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	✓
A→C (W62G)	N/A	<i>ROD_07741</i> ←	phosphatidylethanolamine-binding protein	✓
A→G (E47G)	-5.808	<i>moaD</i> →	molybdopterin converting factor, subunit 1	
A→G (intergenic)	N/A	<i>ROD_09461</i> → / ← <i>ROD_09471</i>	prophage transcriptional regulator/LuxR family transcriptional regulator	✓
A→C (T281P)	-5.834	<i>astC</i> →	succinylornithine transaminase	✓
T→G (V38G)	-6.961	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	
T→C (L184P)	-5.032	<i>dos</i> →	heme-regulated cyclic AMP phosphodiesterase (direct oxygen sensor protein)	
A→C (V665G)	-3.274	<i>ROD_16371</i> ←	hypothetical protein	
G→C (R239P)	-6.615	<i>ROD_20671</i> →	integrase	✓
A→G (S249G)	-2.218	<i>ROD_20671</i> →	"	
C→A (N255K)	-1.232	<i>ROD_20671</i> →	"	
A→C (V101G)	-4.818	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
T→C (E92G)	-6.708	<i>cbiH</i> ←	"	
G→C (L91V)	-2.247	<i>cbiH</i> ←	"	✓
A→C (V264G)	-4.892	<i>fcl</i> ←	GDP-L-fucose synthetase	
A→C (V263G)	-6.772	<i>fcl</i> ←	"	
T→A (I96F)	-3.655	<i>ROD_22911</i> ←	nucleoside transporter	
A→C (V203G)	-6.150	<i>eutJ</i> ←	ethanolamine utilization protein	
A→C (V202G)	-5.747	<i>eutJ</i> ←	"	✓
A→C (V242G)	-4.885	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase	✓
A→C (intergenic)	N/A	<i>dapA</i> ← / → <i>gcvR</i>	dihydrodipicolinate synthase/glycine cleavage system transcriptional repressor	
G→C (A131G)	-2.611	<i>clpB</i> ←	chaperone (heat-shock protein F84.1)	
1 bp deletion (intergenic)	N/A	<i>ler</i> ← / ← <i>ROD_30161</i>	Ler transcriptional regulator/ISCr01 transposase C	✓
A→G (S461G)	-1.272	<i>hycC</i> →	formate hydrogenlyase subunit 3	
112 bp x 2 duplication	N/A	[<i>ROD_t54</i>]	[<i>ROD_t54</i>]	
111/224 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
G→A (P63P)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
T→G	-13.000	<i>ROD_34781</i> →	hypothetical protein	

(W120G)				
A→C (T127P)	-6.000	<i>ROD_34781</i> →	“	
C→T (F137F)	0.000	<i>ROD_34781</i> →	“	✓
A→C (V676G)	-6.112	<i>ROD_35621</i> ←	radical SAM superfamily protein	✓
A→G (E253G)	-1.364	<i>thiG</i> →	thiazole biosynthesis protein	
G→A (R148C)	-2.926	<i>pflC</i> ←	pyruvate formate-lyase 2 activating enzyme	✓
T→G (V110G)	-4.208	<i>ROD_38821</i> →	aldose 1-epimerase	✓
A→C (T66P)	-2.823	<i>ROD_38891</i> →	hypothetical protein	✓
A→C (V216G)	-6.811	<i>ROD_39511</i> ←	amino acid permease	✓
A→G (G543G)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
A→T (Y547F)	0.074	<i>ROD_40471</i> →	“	
G→T (L548L)	0.000	<i>ROD_40471</i> →	“	
A→C (V262G)	-3.887	<i>ROD_40571</i> ←	hypothetical protein	✓
T→G (Y581D)	-7.926	<i>ggt</i> →	gamma-glutamyltranspeptidase	✓
A→C (V91G)	-6.633	<i>greB</i> ←	transcription elongation factor	
7 bp deletion (intergenic)	N/A	<i>espN2-2</i> ← / ← <i>ROD_47851</i>	T3SS effector protein EspN2-2/hypothetical protein	✓
A→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	
A→C (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	“	
T→A (I140F)	-2.722	<i>ROD_48371</i> ←	carbon starvation protein	✓
A→C (V132G)	-6.882	<i>ROD_48371</i> ←	“	✓

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The W4_{P20} strain has a total of 25 SNPs (Table 7.7), with 10 unique and 15 SNPs shared with at least one other lineage. 22 of the SNPs are intragenic and 3 intergenic; of the intragenic SNPs, 13 are non-synonymous and 9 synonymous. The resulting d_N/d_S ratio for W4_{P20} is 1.443, indicating positive selection. In addition, there is a cytosine insertion in *znuA* (high-affinity zinc uptake system protein) resulting in a reading frame-shift and possible disruption of the protein product.

Table 7.7. Mutations in W4_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→C (S254P)	-2.700	<i>ROD_01271</i> →	2-keto-3-deoxygluconate permease	
T→G (T163P)	-2.383	<i>ROD_01611</i> ←	hypothetical protein	
T→G (intergenic)	N/A	<i>ROD_09461</i> → / ← <i>ROD_09471</i>	prophage transcriptional regulator/LuxR family transcriptional regulator	✓
A→C (V663G)	-5.496	<i>putA</i> ←	bifunctional protein PutA	
C→G (G155G)	0.000	<i>selD</i> →	Selenide,water dikinase	✓
T→G (T82P)	-1.868	<i>pfkB</i> ←	6-phosphofructokinase	
1 bp (C) insertion	Frameshift	<i>znuA</i> ←	high-affinity zinc uptake system protein	
A→G (E46G)	-1.400	<i>ROD_20681</i> →	IS102 transposase	✓
G→A (S54S)	0.000	<i>ROD_20681</i> →	"	✓
A→T (S58S)	0.000	<i>ROD_20681</i> →	"	✓
A→C (G61G)	0.000	<i>ROD_20681</i> →	"	✓
G→A (R62R)	0.000	<i>ROD_20681</i> →	"	✓
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
A→C (V101G)	-4.818	<i>cbiH</i> ←	"	
G→C (L89V)	-1.831	<i>cbiH</i> ←	"	
C→T (P172S)	-5.590	<i>ROD_23201</i> →	sulfatase	
A→T (R225R)	0.000	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase	✓
T→C (K293R)	-0.698	<i>ROD_27631</i> ←	outer membrane autotransporter	✓
A→C (V331G)	-6.364	<i>ROD_27781</i> ←	fimbrial usher protein	
A→G (A200A)	0.000	<i>espF</i> ←	T3SS effector protein EspF	
A→G (S127P)	5.442	<i>espF</i> ←	"	
A→T (Y547F)	0.074	<i>ROD_40471</i> →	sodium:solute symporter	
G→T (L548L)	0.000	<i>ROD_40471</i> →	"	
T→G (G114G)	0.000	<i>ROD_47081</i> →	hypothetical protein	✓
A→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	✓
A→C (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	"	✓

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The W5_{P20} strain has a total of 44 SNPs (Table 7.8), with 12 unique and 32 SNPs shared with at least one other lineage. 39 of the SNPs are intragenic and 5 intergenic; of the intragenic SNPs, 32 are non-synonymous and 7 synonymous. The resulting d_N/d_S ratio for W5_{P20} is 4.571, indicating positive selection. In addition, there is a 223 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein).

Table 7.8. Mutations in W5_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→C (E320G)	-1.400	<i>thiP</i> ←	thiamine ABC transporter permease	✓
T→A (Y166F)	-0.777	<i>ROD_01611</i> ←	hypothetical protein	
T→G (T163P)	-2.383	<i>ROD_01611</i> ←	"	
T→G (H161P)	-4.012	<i>ROD_01611</i> ←	"	
T→C (S159G)	-1.508	<i>ROD_01611</i> ←	"	✓
A→G (intergenic)	N/A	<i>ROD_02491</i> ← / ← <i>ROD_02501</i>	hypothetical protein/hypothetical protein	✓
C→A (A227A)	0.000	<i>sbcC</i> ←	exonuclease SbcC	✓
C→G (G270G)	0.000	<i>ROD_05911</i> →	transketolase	✓
T→G (T341P)	-0.982	<i>abrB</i> ←	ammonia monooxygenase	
A→C (T336P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	
A→C (T291P)	-1.400	<i>ROD_11451</i> →	LysR family transcriptional regulator	
T→G (T82P)	-1.868	<i>pfkB</i> ←	6-phosphofructokinase	
T→G (T65P)	-4.339	<i>pfkB</i> ←	"	
T→G (V38G)	-6.961	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	
A→C (V654G)	-6.727	<i>ROD_16371</i> ←	hypothetical protein	
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
T→C (E92G)	-6.708	<i>cbiH</i> ←	"	
G→C (R269G)	-2.508	<i>fcl</i> ←	GDP-L-fucose synthetase	✓
A→C (V264G)	-4.892	<i>fcl</i> ←	"	
T→G (T112P)	-2.168	<i>ROD_22651</i> ←	FAA-hydrolase-family protein	✓
T→A (I96F)	-3.655	<i>ROD_22911</i> ←	nucleoside transporter	
T→C (E57G)	-0.286	<i>ROD_22911</i> ←	"	
A→C (V244G)	-6.847	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase	
A→C (V243G)	-6.333	<i>eutE</i> ←	"	
A→C (intergenic)	N/A	<i>dapA</i> ← / → <i>gcvR</i>	dihydridopicolinate synthase/glycine cleavage system transcriptional repressor	
A→C (T189P)	-2.410	<i>ROD_24351</i> →	peptidase	✓
A→C (V331G)	-6.364	<i>ROD_27781</i> ←	fimbrial usher protein	
A→G (S127P)	5.442	<i>espF</i> ←	T3SS effector protein EspF	
A→C (T428P)	-5.791	<i>cysN</i> →	Sulfate adenylyltransferase subunit 1	
T→A (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
+223 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	"	
G→A (P63P)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
C→T (R62R)	0.000	<i>ROD_33401</i> ←	"	✓
T→G (G61G)	0.000	<i>ROD_33401</i> ←	"	
T→A (S58S)	0.000	<i>ROD_33401</i> ←	"	
C→T (S54S)	0.000	<i>ROD_33401</i> ←	"	
T→C (E46G)	-1.400	<i>ROD_33401</i> ←	"	
T→G (W120G)	-13.000	<i>ROD_34781</i> →	hypothetical protein	
A→G (E253G)	-1.364	<i>thiG</i> →	thiazole biosynthesis protein	
T→A (L430F)	-2.526	<i>ROD_38291</i> ←	hypothetical protein	✓
T→A (Y425F)	-1.731	<i>ROD_38291</i> ←	"	✓
A→T (Y547F)	0.074	<i>ROD_40471</i> →	sodium:solute symporter	
G→T (L548L)	0.000	<i>ROD_40471</i> →	"	
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

7.2.5 Genetic changes in *C. rodentium* adapted to the mouse environment in the presence of nalidixic acid

From a filtered list, I identified a total of 127 SNPs present in the *C. rodentium* which adapted in mice in the presence of antibiotic (N1_{P20}-N5_{P20}); 78 of these are only present in N1_{P20}-N5_{P20} *C. rodentium*, and 69 of these are unique to a single lineage. The N1_{P20} strain has a total of 38 SNPs (Table 7.9), with 13 unique and 25 SNPs shared with at least one other lineage. 33 of the SNPs are intragenic and 5 intergenic; of the intragenic SNPs, 29 are non-synonymous and 4 synonymous. The resulting d_N/d_S ratio for N1 is 7.250, indicating positive selection. In addition, a 223 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein) is present.

Intragenic SNPs were investigated for predicted functional changes to their coding gene using the Protein Variation Effect Analyzer (PROVEAN) web based tool ^{249–251}. Scores of -2.5 or below are predicted to have a deleterious or non-neutral effect on the protein, however whether this change would result in a non-functional protein or a protein with an altered function was not determined.

Table 7.9. Mutations in N1_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→C (S254P)	-2.700	<i>ROD_01271</i> →	2-keto-3-deoxygluconate permease	
T→G (T167P)	-1.512	<i>ROD_01611</i> ←	hypothetical protein	
A→T (intergenic)	N/A	<i>ROD_02491</i> ← / ← <i>ROD_02501</i>	hypothetical protein/hypothetical protein	✓
T→G (V257G)	-6.542	<i>lafU</i> →	flagellar motor protein B	
T→G (T341P)	-0.982	<i>abrB</i> ←	ammonia monooxygenase	
T→C (intergenic)	N/A	<i>ROD_09461</i> → / ← <i>ROD_09471</i>	prophage transcriptional regulator/LuxR family transcriptional regulator	✓
A→C (V663G)	-5.496	<i>putA</i> ←	bifunctional protein PutA	
A→C (T291P)	-1.400	<i>ROD_11451</i> →	LysR family transcriptional regulator	
T→G (T65P)	-4.339	<i>pfkB</i> ←	6-phosphofructokinase	
A→G (E106G)	-3.265	<i>ROD_14451</i> →	oxidoreductase	✓
A→C (V665G)	-3.274	<i>ROD_16371</i> ←	hypothetical protein	
A→C (W646G)	-12.497	<i>ROD_16371</i> ←	"	✓
G→T (L50F)	-2.770	<i>ROD_20381</i> →	hypothetical protein	✓
C→A (N255K)	-1.232	<i>ROD_20671</i> →	integrase	
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
A→C (V101G)	-4.818	<i>cbiH</i> ←	"	
T→C (E92G)	-6.708	<i>cbiH</i> ←	"	
A→C (V90G)	-5.958	<i>cbiH</i> ←	"	✓
T→A (I77F)	-2.359	<i>ROD_22911</i> ←	nucleoside transporter	
T→C (S454G)	-3.835	<i>mgtE</i> ←	magnesium transporter	✓
G→C (A188G)	-3.974	<i>eutJ</i> ←	ethanolamine utilization protein	✓
A→C (G284G)	0.000	<i>ROD_27631</i> ←	outer membrane autotransporter	
T→G (V843G)	-5.224	<i>fimD</i> →	fimbrial usher protein FimD	
T→G (V845G)	-5.012	<i>fimD</i> →	"	✓
T→A (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
+223 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	"	

T→A (S58S)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
C→T (S54S)	0.000	<i>ROD_33401</i> ←	“	
A→C (T127P)	-6.000	<i>ROD_34781</i> →	hypothetical protein	
A→T (Y135F)	-4.000	<i>ROD_34781</i> →	“	✓
A→C (V76G)	-6.904	<i>gcp</i> ←	O-sialoglycoprotein endopeptidase (glycoprotease)	✓
T→C (E217G)	-6.911	<i>ROD_39511</i> ←	amino acid permease	✓
C→T (L548L)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
T→G (W553G)	-0.706	<i>ROD_40471</i> →	“	✓
T→G (V136G)	-0.828	<i>ftsY</i> →	cell division protein	
A→C (V91G)	-6.633	<i>greB</i> ←	transcription elongation factor	
A→C (W90G)	-11.777	<i>greB</i> ←	“	
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The N2_{P20} strain has a total of 32 SNPs (Table 7.10), with 9 unique and 23 SNPs shared with at least one other lineage. All 32 of the SNPs are intragenic; 24 are non-synonymous and 8 synonymous. The resulting d_N/d_S ratio for N2_{P20} is 3.000, indicating positive selection. In addition, a cytosine insertion in *znuA* (high-affinity zinc uptake system protein) resulting in a reading frame-shift and possible disruption of the protein product. A 112 bp duplication of part of *ROD_t54* (tRNA-Gly) is also present.

Table 7.10. Mutations in N2_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→G (T163P)	-2.383	<i>ROD_01611</i> ←	hypothetical protein	
T→G (M263R)	-4.684	<i>lafU</i> →	flagellar motor protein B	
A→C (T140P)	-3.775	<i>cstA</i> →	carbon starvation protein A	✓
T→C (E188G)	-4.157	<i>ROD_06891</i> ←	hypothetical protein	
T→G (T82P)	-1.868	<i>pfkB</i> ←	6-phosphofructokinase	
T→G (V38G)	-6.961	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	
1 bp (C) insertion	Frameshift	<i>znuA</i> ←	high-affinity zinc uptake system protein	
C→A (N255K)	-1.232	<i>ROD_20671</i> →	integrase	
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
A→C (V101G)	-4.818	<i>cbiH</i> ←	"	
A→C (V264G)	-4.892	<i>fcl</i> ←	GDP-L-fucose synthetase	
A→C (L256R)	-5.736	<i>fcl</i> ←	"	✓
T→C (E255G)	-6.296	<i>fcl</i> ←	"	✓
A→G (I253T)	-4.948	<i>fcl</i> ←	"	✓
T→A (I77F)	-2.359	<i>ROD_22911</i> ←	nucleoside transporter	
A→G (S149P)	-1.529	<i>eutH</i> ←	ethanolamine transporter	✓
T→C (E335G)	-6.797	<i>ROD_27781</i> ←	fimbrial usher protein	
A→C (V333G)	-6.797	<i>ROD_27781</i> ←	"	
T→G (V843G)	-5.224	<i>fimD</i> →	fimbrial usher protein FimD	
A→G (A200A)	0.000	<i>espF</i> ←	T3SS effector protein EspF	
112 bp x 2 duplication	112 bp duplication	[<i>ROD_t54</i>]	[<i>ROD_t54</i>]	
G→A (P63P)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
C→T (R62R)	0.000	<i>ROD_33401</i> ←	"	
T→G (G61G)	0.000	<i>ROD_33401</i> ←	"	
T→A (S58S)	0.000	<i>ROD_33401</i> ←	"	
C→T (S54S)	0.000	<i>ROD_33401</i> ←	"	
T→C (E46G)	-1.400	<i>ROD_33401</i> ←	"	
C→T (R22K)	-0.190	<i>ROD_33401</i> ←	"	✓
T→G (V38G)	-5.744	<i>ctsS</i> →	T6SS protein CtsS	
T→G (V89G)	-6.506	<i>ROD_39711</i> →	permease	✓
G→T (L548L)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
A→G (P55P)	0.000	<i>gspC</i> ←	T2SS protein C	✓
A→C (V137G)	-6.847	<i>ROD_48371</i> ←	carbon starvation protein	✓
A→G (S411G)	-2.842	<i>radA</i> →	DNA repair protein	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The N3_{P20} strain has a total of 29 SNPs (Table 7.11), with 22 unique and 7 SNPs shared with at least one other lineage. 21 of the SNPs are intragenic and 8 intergenic; of the intragenic SNPs, 18 are non-synonymous and 3 synonymous. The resulting d_N/d_S ratio for N3_{P20} is 6.000, indicating positive selection. In addition, there is a 9 bp deletion in the intergenic space between *uspF* and *nleC* (universal stress protein F and type 3 secretion system effector protein NleC), as well as a 1 bp deletion in *ugpA* (glycerol 3 phosphate ABC transporter permease) resulting in a reading frame-shift and possible disruption of the protein product.

Table 7.11. Mutations in N3_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
G→T (intergenic)	N/A	<i>ROD_00141</i> ← / → <i>ROD_00151</i>	hypothetical protein/polysaccharide degrading enzyme	✓
C→T (F234F)	0.000	<i>ROD_01271</i> →	2-keto-3-deoxygluconate permease	✓
T→A (intergenic)	N/A	<i>cI</i> → / → <i>lfjJ</i>	phage repressor protein/flagellar export/assembly protein	✓
A→T (intergenic)	N/A	<i>ROD_05651</i> ← / ← <i>ROD_05671</i>	hypothetical protein/mechanosensitive ion channel protein	✓
A→C (V330G)	3.083	<i>ROD_09671</i> ←	hypothetical protein	✓
T→G (N154T)	-2.638	<i>rluE</i> ←	ribosomal large subunit pseudouridine synthase E (rRNA pseudouridylate synthase E)	✓
T→G (T82P)	-1.868	<i>pfkB</i> ←	6-phosphofructokinase	
T→C (E67G)	-3.149	<i>pfkB</i> ←	"	✓
T→G (V38G)	-6.961	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	
T→C (T650A)	-0.567	<i>ROD_16371</i> ←	hypothetical protein	✓
T→C (intergenic)	N/A	<i>ROD_16381</i> → / → <i>ROD_16391</i>	hypothetical protein/outer membrane autotransporter	✓
9 bp deletion (intergenic)	N/A	<i>uspF</i> → / ← <i>nleC</i>	universal stress protein F/T3SS effector protein NleC	✓
T→C (E257G)	-3.350	<i>rluB</i> ←	ribosomal large subunit pseudouridine synthase B	✓
A→G (I106T)	-4.104	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	✓
C→T (P172S)	-5.590	<i>ROD_23201</i> →	sulfatase	
A→C (V331G)	-6.364	<i>ROD_27781</i> ←	fimbrial usher protein	
A→G (K646K)	0.000	<i>ROD_29581</i> →	large repetitive protein	✓
G→A (E33K)	-3.933	<i>cts2S</i> →	T6SS protein Cts2S	✓
T→G (V38G)	-5.744	<i>cts2S</i> →	"	
A→T (I33F)	-0.681	<i>ROD_34401</i> →	dimethyl sulfoxide reductase	✓
C→T (intergenic)	N/A	<i>acs</i> ← / → <i>nrfA</i>	acetyl-coenzyme A synthetase/cytochrome c552	✓
C→T (L21L)	0.000	<i>zupT</i> →	zinc transporter	✓
C→T (A72T)		<i>rffT</i> ←	4-alpha-L-fucosyltransferase	✓
T→G (V136G)	-0.828	<i>ftsY</i> →	cell division protein	
1 bp deletion	Frameshift	<i>ugpA</i> →	glycerol-3-phosphate ABC transporter permease	✓
A→C (W90G)	-11.777	<i>greB</i> ←	transcription elongation factor	
T→G (V409G)	-6.646	<i>radA</i> →	DNA repair protein	✓
A→G (E414G)	-6.318	<i>radA</i> →	"	✓
T→A (intergenic)	N/A	<i>ROD_50641</i> ← / → <i>epd</i>	IS102 transposase/d-erythrose-4-phosphate dehydrogenase	✓

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The N4_{P20} strain has a total of 45 SNPs (Table 7.12), with 17 unique and 28 SNPs shared with at least one other lineage. 39 of the SNPs are intragenic and 6 intergenic; of the intragenic SNPs, 32 are non-synonymous and 7 synonymous. The resulting d_N/d_S ratio for N4 is 4.571, indicating positive selection. In addition, a 112 bp duplication of part of *ROD_t54* (tRNA-Gly) is present, as well as a 7-14 bp TTATTCC repeat inserted in the intergenic space between *espN2-2* and *ROD_47851* (type 3 secretion system effector protein EspN2-2 and a hypothetical protein).

Table 7.12. Mutations in N4_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→G (V217G)	0.245	<i>hrpB</i> →	ATP-dependent helicase	✓
T→G (H161P)	-4.012	<i>ROD_01611</i> ←	hypothetical protein	
T→G (M263R)	-4.684	<i>lafU</i> →	flagellar motor protein B	
T→G (V316G)	-3.309	<i>ROD_05511</i> →	permease	✓
T→C (E188G)	-4.157	<i>ROD_06891</i> ←	hypothetical protein	
T→G (intergenic)	N/A	<i>ROD_07151</i> → / ← <i>abrB</i>	hypothetical protein/ammonia monooxygenase	✓
A→C (T336P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	
A→C (T291P)	-1.400	<i>ROD_11451</i> →	LysR family transcriptional regulator	
A→G (E28G)	-3.709	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	✓
T→G (V29G)	-6.266	<i>ROD_15031</i> →	"	
T→G (V38G)	-6.961	<i>ROD_15031</i> →	"	
A→C (V665G)	-3.274	<i>ROD_16371</i> ←	hypothetical protein	
T→C (T655A)	-1.158	<i>ROD_16371</i> ←	"	✓
C→T (A440V)	-1.102	<i>narX</i> →	nitrate/nitrite two-component sensor kinase	✓
A→T (D252V)	-7.406	<i>ROD_20671</i> →	integrase	✓
C→A (N255K)	-1.232	<i>ROD_20671</i> →	"	
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
A→C (V101G)	-4.818	<i>cbiH</i> ←	"	
T→C (D100G)	-3.637	<i>cbiH</i> ←	"	✓
T→C (E92G)	-6.708	<i>cbiH</i> ←	"	
A→C (V264G)	-4.892	<i>fcl</i> ←	GDP-L-fucose synthetase	
A→C (V263G)	-6.772	<i>fcl</i> ←	"	
T→C (Q611R)	-2.420	<i>wzc</i> ←	tyrosine-protein kinase	✓
T→A (I77F)	-2.359	<i>ROD_22911</i> ←	nucleoside transporter	
T→C (E57G)	-0.286	<i>ROD_22911</i> ←	"	
A→C (V56G)	-4.618	<i>ROD_22911</i> ←	"	✓
T→G (P50P)	0.000	<i>ROD_22911</i> ←	"	✓
A→C (V244G)	-6.847	<i>eutE</i> ←	ethanolamine utilization aldehyde dehydrogenase	
C→T (R216H)	-1.819	<i>nuoG</i> ←	NADH-quinone oxidoreductase subunit G	✓
A→C (intergenic)	N/A	<i>ackA</i> → / → <i>pta</i>	acetate kinase/phosphate acetyltransferase	✓
A→C (G284G)	0.000	<i>ROD_27631</i> ←	outer membrane autotransporter	
A→G (S231S)	0.000	<i>espF</i> ←	T3SS effector protein EspF	✓
112 bp x 2 duplication	112 bp duplication	[<i>ROD_t54</i>]	[<i>ROD_t54</i>]	
T→A (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
C→T (R62R)	0.000	<i>ROD_33401</i> ←	IS102 transposase	
T→G (V37G)	-6.750	<i>cts2S</i> →	T6SS protein Cts2S	

T→G (V38G)	-5.744	<i>ctsS</i> →	“	
A→G (G543G)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
G→T (L548L)	0.000	<i>ROD_40471</i> →	“	
T→G (V1066G)	-5.225	<i>bcsC</i> →	cellulose synthase operon protein C (TPR-repeat-containing protein)	✓
T→G (V136G)	-0.828	<i>ftsY</i> →	cell division protein	
T→G (V73G)	-5.721	<i>pabA</i> →	para-aminobenzoate synthase glutamine amidotransferase component I	✓
G→T (G240G)	0.000	<i>gatC</i> ←	component IIC of galactitol-specific phosphotransferase system	✓
+TTATTCC / or +14bp repeat insertion (intergenic)	N/A	<i>espN2-2</i> ← / ← <i>ROD_47851</i>	T3SS effector protein EspN2-2/hypothetical protein	
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	
T→G (intergenic)	N/A	<i>ROD_50641</i> ← / → <i>epd</i>	IS102 transposase/d-erythrose-4-phosphate dehydrogenase	

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

The N5_{P20} strain has a total of 26 SNPs (Table 7.13), with 8 unique and 18 SNPs shared with at least one other lineage. 23 of the SNPs are intragenic and 3 intergenic; of the intragenic SNPs, 20 are non-synonymous and 3 synonymous. The resulting d_N/d_S ratio for N5_{P20} is 6.667, indicating positive selection. In addition, there is a 223 bp insertion in the intergenic space between *ROD_t55* and *ROD_32251* (tRNA-Gly and a hypothetical protein).

Table 7.13. Mutations in N5_{P20} genome

Mutation	PROVEAN score*	Gene affected	Function	Unique
T→C (S254P)	-2.700	<i>ROD_01271</i> →	2-keto-3-deoxygluconate permease	
T→G (T163P)	-2.383	<i>ROD_01611</i> ←	hypothetical protein	
A→G (M252V)	-1.145	<i>lafU</i> →	flagellar motor protein B	✓
T→G (V257G)	-6.542	<i>lafU</i> →	"	
T→A (*314K)	N/A	<i>ROD_05911</i> →	transketolase	✓
A→C (T336P)	-5.986	<i>sucB</i> →	dihydrolipoamide succinyltransferase component (E2)	
G→T (S219R)	-1.328	<i>cydC</i> ←	transport ATP-binding protein CydC	✓
T→G (V29G)	-6.266	<i>ROD_15031</i> →	ABC transporter ATP-binding protein	
T→G (V38G)	-6.961	<i>ROD_15031</i> →	"	
T→G (Y78S)	-8.421	<i>narG</i> ←	respiratory nitrate reductase 1 subunit alpha	✓
A→G (S249G)	-2.218	<i>ROD_20671</i> →	integrase	
T→C (E102G)	-3.936	<i>cbiH</i> ←	cobalt-precorrin-3B C(17)-methyltransferase	
T→C (E92G)	-6.708	<i>cbiH</i> ←	"	
G→C (L89V)	-1.831	<i>cbiH</i> ←	"	
A→G (R332R)	0.000	<i>ROD_22201</i> ←	ABC transporter substrate-binding protein	✓
C→T (P172S)	-5.590	<i>ROD_23201</i> →	sulfatase	
T→G (A289A)	0.000	<i>ROD_27631</i> ←	outer membrane autotransporter	✓
T→C (E335G)	-6.797	<i>ROD_27781</i> ←	fimbrial usher protein	
+223 bp insertion (intergenic)	N/A	<i>ROD_t55</i> → / ← <i>ROD_32251</i>	tRNA-Gly/hypothetical protein	
T→G (V37G)	-6.750	<i>cts2S</i> →	T6SS protein Cts2S	
T→G (V38G)	-5.744	<i>cts2S</i> →	"	
A→C (T127P)	-6.000	<i>ROD_34781</i> →	hypothetical protein	
C→T (L548L)	0.000	<i>ROD_40471</i> →	sodium:solute symporter	
T→G (V340G)	-4.334	<i>ROD_47661</i> →	hypothetical protein	✓
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	outer membrane efflux protein of T1SS/ATP-binding protein of T1SS	
T→G (intergenic)	N/A	<i>ROD_48151</i> → / → <i>ROD_48181</i>	"	✓

* Provean scores which are -2.5 or below, and therefore predicted to have an impact on the protein, are highlighted in bold.

7.2.6 Categorisation of the genes mutated in the *in vivo*-adapted *C. rodentium* strains

I further categorised the mutations present in each of the *in vivo*-adapted strains based on the function of the gene, or nearby gene, which was altered (Fig. 7.2). The categories used included:

- Transporters and permeases
- Genes known to be involved in *C. rodentium* virulence
- Genes involved in metabolism or catabolism
- Transcriptional regulators
- Genes which encode for hypothetical proteins with unknown functions
- Transposases and integrases
- Genes involved in translation, transcription, and DNA repair
- Genes involved in the type 6 secretion system.

It is important to note that the categories are broad and designed to highlight overall trends, but that a single gene may be involved with more than one category.

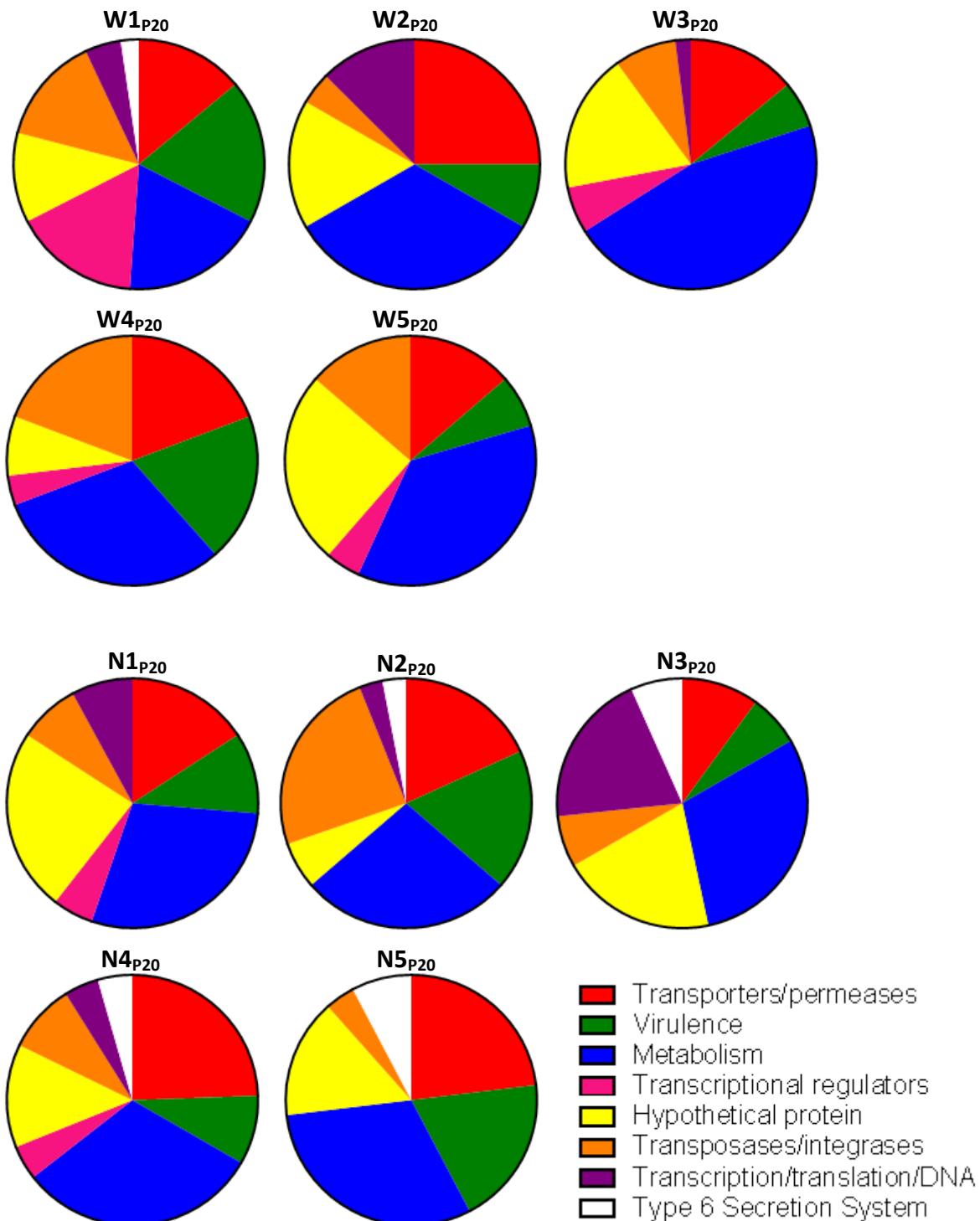


Figure 7.2. Proportions SNPs in genes with estimated functions. Genes with SNPs present for each *in vivo*-adapted *C. rodentium* strain were categorised based on estimated function, with broad categories including transporters/permeases (red); genes known to be involved in *C. rodentium* virulence (green); genes involved in metabolism or catabolism (blue); transcriptional regulators (pink); hypothetical proteins with unknown functions (yellow); transposases and integrases (orange); genes involved in transcription, translation, or DNA repair pathways (purple); and SNPs present in the type 6 secretion system (white). Proportions given are relative to the total number of SNPs present in each given strain.

7.2.7 Little evidence of foreign DNA uptake in *in vivo*-adapted *C. rodentium*

I performed *de novo* assembly using MIRA 4 on reads which did not map to the ICC168 reference genome¹⁵². I then searched the contigs for similarity to known sequences using the BLAST® program^{229,230,252}. I confirmed the presence of the known *C. rodentium* plasmids, pCROD1, pCROD2, pCROD3, and pCRP3, as well as the presence of the *lux* operon insertion. In addition, there were several contigs which had sequence similarity with other enterobacterial phages and plasmids, with no strong similarity to any particular known phage/plasmid (not shown).

Many of the reads which did not show any sequence similarity in BLAST® searches were highly repetitive sequences (Appendix 2). I found non-repetitive contigs in N4_{P20} and N5_{P20}, translated the contigs into protein sequences using ExPASy²⁵³, and then searched open reading frames for similarity to known protein sequences using BLASTP (version 2.2.31+)²²⁸. For the N4_{P20} strain, the majority of unmapped contigs present with no sequence similarity were shown to encode for *Candida* spp. proteins, and subsequently disregarded as contamination (not shown). The other two contigs encoded for proteins with no discernible protein similarity to known species or were a repetitive sequence. The N5_{P20} strain had one non-repetitive sequence with no discernible protein similarity to known species, and two repetitive sequences.

7.2.8 Minimal changes to plasmids present in *C. rodentium*

To determine any changes to plasmid DNA, I mapped the sequences against the known plasmids present in *C. rodentium*: pCROD1 (Genbank accession number FN543503.1), pCROD2 (Genbank accession number FN543504.1), pCROD3 (Genbank accession number FN543505.1) and pCRP3 (Genbank accession number NC_003114). The *in vivo*-adapted strains all had the same pCROD1 plasmid which is present in the ancestor ICC180 (Chapter 3¹⁵³); that is, the plasmid, while present, is missing 41 out of 60 of its genes. I did not observe any further loss of the pCROD1 plasmid in the *in vivo*-adapted strains, which is in contrast to previous reports that the pCROD1 plasmid is frequently lost¹³⁴. pCROD2 remains largely unchanged in the *in vivo*-adapted strains, with the exception of N1_{P20} pCROD2 with a 6 bp insertion in *ROD_RS25435*, the conjugation transfer protein TrbI, and W4_{P20} showing evidence of two junction change events (Table 7.15). The two remaining plasmids, pCROD3 and pCRP3, had no evidence of any genetic changes in any of the *in vivo*-adapted strains.

Table 7.15. Genetic changes to pCROD2 in *in vivo*-adapted *C. rodentium*

Location	Change	Gene	Gene function	Strain
26583	+CCGCCG insertion	<i>ROD_RS25435</i> ←	conjugation transfer protein TrbI	N1 _{P20}
31057 / 25579	Junction change	<i>18 bp inverted repeat of IS102 / ROD_RS25430</i>	18 bp inverted repeat of IS102 / P-type DNA transfer ATPase VirB11	W4 _{P29}
25579 / 32112	Junction change	<i>ROD_RS25430 / noncoding (2/18 nt)</i>	P-type DNA transfer ATPase VirB11 / 18 bp inverted repeat of IS102	W4 _{P20}

7.3 Discussion

I have taken a population-based approach to sequencing the *in vivo*-adapted *C. rodentium* strains isolated from the 20th passage through mice. Sequencing of populations has in recent years become an attractive option for gaining insight into genetic changes in experimental evolution systems^{254–257}. In this way, I have identified mutations and SNPs and their frequencies within the population, but it is important to note that this does not reveal the complete genomes of each population and subpopulation; that is, whether two SNPs at a 50% frequency on separate reads are present in the same genome, or whether each of the SNPs is present in two subpopulations, or a combination thereof. As the mutations present are identified at a wide range of frequencies, I propose that it would be safe to assume that a number of subpopulations are present within the total populations, and therefore by using a population-based approach I have identified mutations which would have been missed if a single clone was sequenced.

7.3.1 Genes known to be important for *C. rodentium* virulence are conserved

There are three main options for interpreting genes which are mutated following adaptation:

- 1) Changes in the gene indicate an adaptive change to the function of the gene, and highlight that the altered gene plays an important role in the experimental environment.
- 2) Changes in the gene are neutral and have a negligible effect.
- 3) Changes in the gene result in a non-functional pseudogene, indicating the gene and its product are not important in the experimental environment.

Following the third interpretation, genes which are important in the experimental environment will remain unchanged. Out of the list of genes known to be important for *C. rodentium* virulence (Chapter 1, Table 1.4), there is a 1 bp deletion present in the intergenic space between *ler* and *ROD_30161* in W3_{P20} (frequency 100%); an adenine to cytosine SNP present in the intergenic space between *escD* and *eae* in W1_{P20} (100% frequency); and a number of SNPs present in *espF*: three synonymous mutations, and one non-synonymous SNP of an adenine change to a guanine, resulting in a serine to proline change at residue 127, which is present in W4_{P20} and W5_{P20} (25.9% and 27% frequencies, respectively). That there are so few known virulence genes altered in the *in vivo*-adapted strains, I propose that these genes are highly conserved, and that it is in the bacterium's best interests to leave these genes unchanged. Indeed, the only known virulence gene with a non-synonymous mutation present (*espF*) has been shown experimentally to not be essential for colonisation and virulence *in vivo*

¹²².

7.3.2 Mutations present in genes encoding for hypothetical proteins

The number of pseudogenes present in a genome is reflective of adaptation to a host and the removal of the unnecessary products, which is referred to as reductive evolution²⁵⁸. Examples of genomes which have a large number of pseudogenes include *M. leprae* and *Rickettsia prowazekii*²⁵⁸. I observed that mutations affecting genes which encode for a hypothetical protein accounted for a sizeable proportion of mutations present in each strain, present at a proportion of 12%, 17%, 18%, 8%, and 25% for W1_{P20}, W2_{P20}, W3_{P20}, W4_{P20}, and W5_{P20}, respectively, and a proportion of 24%, 6%, 20%, 13%, and 15% for N1_{P20}, N2_{P20}, N3_{P20}, N4_{P20}, and N5_{P20}, respectively (Fig. 7.1). Given the wide range of genes which have SNPs, one would not assume that all of these genetic changes have led to non-functional genes, and therefore at least some of the SNPs may alter the function of the respective genes. Further investigation and modelling of the genes and the altered protein would be required to determine whether the mutations present in the affected genes result in a non-functional pseudogene, or a product with an altered function.

7.3.3 The Type 6 Secretion System *cts2S* gene is mutated in the majority of “N” strains

W1_{P20}, N2_{P20}, N3_{P20}, N4_{P20}, and N5_{P20} all have mutations in the *cts2S* gene, part of the *cts2* gene cluster which encodes for a type 6 secretion system unique to *C. rodentium*. The mutations present are all close together in the gene, with the first guanine to adenosine change resulting in a glutamate to lysine change at residue 33, the second a thymine to guanine change resulting in a valine to glycine change at residue 37, and the third a thymine to guanine change resulting in a valine to glycine change at residue 38. N3_{P20} has the first type of change at a frequency of 28.9%; W1_{P20} has the second type of change at a frequency of 31%; and the third type of change is present in N2_{P20}, N3_{P20}, N4_{P20}, and N5_{P20} at frequencies of 26.8%, 35.5%, 34.8%, and 27.6%, respectively. That this change is present in strains from so many lineages suggests that it provides some benefit for *C. rodentium* *in vivo*. That the change is present predominantly in the “N” strains perhaps signifies that it is of greater importance in animals whose microbiota are altered by the antibiotic treatment.

7.3.4 Genes involved in ethanolamine metabolism are frequently altered

I observed two genes involved in ethanolamine metabolism with multiple genetic changes present: *eutJ* and *eutE*, encoding for an ethanolamine utilization protein and an ethanolamine utilization aldehyde dehydrogenase enzyme (Table 7.2). Other genes involved in ethanolamine metabolism were also modified, with 8 out of 10 strains possessing at least one altered ethanolamine metabolism gene. Ethanolamine, which is present in the colon, has been shown to play a role in allowing EHEC to recognise the gut and promote the expression of genes involved in virulence²⁵⁹.

7.3.5 Similar genetic changes present in strains from distinct lineages

I observed a single instance of a particular location in the genome containing a different genetic change in different lineages: a change from an adenine to either a guanine (strain W5_{P20}) or a thymine (strain N1_{P20}) in the intergenic space between *ROD_02491* and *ROD_02501*, both of which encode for hypothetical proteins. It is interesting to note in all other cases the observed mutation is identical in the different lineages, implying one or more of two things: that the mutations for each location is genetically more likely to occur, or that the mutation at each seemingly random location plays an important part and has been selected for in the separate lineages, indicating that each change is adaptive and not simply there by random chance. It has been shown that transitions from cytosine → thiamine, and guanine → adenine, are more likely in bacteria²⁶⁰, however when I compared the filtered list of mutations for transition (cytosine ↔ thymine and guanine ↔ adenine) and transversion (all other combinations) mutations I found 66 transitions (of which only 21 involved the cytosine → thymine, and guanine → adenine transition) and 134 transversions. Therefore, I propose that the observation of the same genetic change present in the same location in multiple strains is a result of the change being advantageous in the experimental environment.

7.3.6 The plasmids present in *C. rodentium* appear to be stable

Changes in plasmids and other mobile genetic elements have been recognised as a potential avenue for large changes in genomes, and large transfers of genetic information between species²⁶¹. I mapped the reads to each of the known *C. rodentium* plasmids, as well as assembling and running database searches on reads which did not map to *C. rodentium* ICC168 in order to address any non-genome genetic changes. While I identified very few changes in the known *C. rodentium* plasmids (Table 7.15), there was some evidence of new genetic information with similarity to bacteriophages and plasmids. Further analysis and confirmation would be required to determine if the DNA is present in the *in vivo*-adapted strains, or if they are simply a result of the sampling method used, which involved stools with antibiotic enrichment and therefore did not completely remove the chance of contamination.

7.3.7 Conclusions

While the work detailed in this chapter provides a list of genetic changes present in the *in vivo*-adapted *C. rodentium* strains, knowing the changes present does not necessarily point towards function of the changes, and it may be that it is the mutations in combination that lead towards the fitness advantages and phenotypes I have observed. This concept is referred to as epistasis – that the sum of genetic changes has a greater impact than the impact of its parts²⁶². Fully elucidating the effect of each mutation will involve the construction of numerous combinations of mutants.

Chapter 8: Experimental *in vitro* Evolution of *C. rodentium*

8.1 Introduction

To determine the differences in evolutionary trajectories between the widely used *in vitro* experimental paradigm and the more complex environment proposed in these *in vivo* experiments, I also designed and ran a complementary *in vitro* experiment to run alongside the novel *in vivo* experiment. I used six replicates of *C. rodentium* ICC180, reflecting the Lenski “gold standard”. I hypothesise that evolution in laboratory media will result in different changes and observable phenotypes when compared to evolution in the more complex mouse host.

8.2 Methods

I inoculated six 50 ml tubes, each containing 15 ml of modified Davis & Mingoli (DM) media, with *C. rodentium* ICC180 and incubated them for 24 hours at 37°C with shaking at 200 revolution per minute (RPM). After 24 hours of incubation, I subcultured 50 µl from each tube into fresh media and incubated for a further 24 hours, then repeated this process for a 5 month period. These lineages were designated L1, L2, L3, L4, L5 and L6, and bacteria taken from each lineage were given the same designation. Based on the number of *C. rodentium* used to inoculate each successive tube, and the final population size following 24-hour incubation, I estimate that the bacteria underwent 8 generations per day during the *in vitro* evolution experiment.

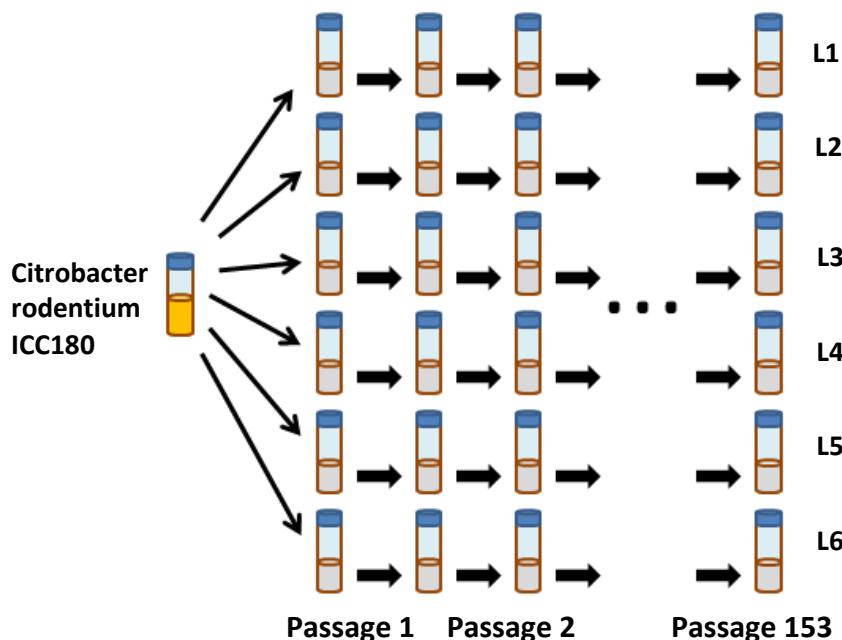


Figure 8.1. Schematic of the *Citrobacter rodentium* *in vitro* evolution experiment. A schematic of the *in vitro* media evolution experiment. Six 50 ml tubes containing 15 ml of modified Davis & Mingoli media were inoculated *C. rodentium* ICC180 and incubated for 24 hours with 200 revolutions per minute (RPM) shaking at 37°C. 50 µl of culture (1×10^8 CFU) was used to inoculate tubes with fresh 15 ml of modified Davis & Mingoli media, which were then incubated for 24 hours as before. This subculturing process was repeated over a period of five months.

8.3 Results

8.3.1 Reproducible loss of bioluminescence following evolution of ICC180 in laboratory media

I observed that five of the six ICC180 lineages evolved in laboratory media lost the ability to produce light after approximately 500 generations. These lineages also regained sensitivity to kanamycin, suggesting complete disruption or removal of the *lux* operon and kanamycin resistance gene insertion. I initially thought that lineages which had lost light had become contaminated, so restarted the experiment from frozen stocks of a previous passage number. However, the loss of light consistently occurred after a similar number of generations. Lineage 1 (L1) retained its ability to glow throughout the 5 months of passaging. Lineage 2 (L2) lost light between passages 57 and 64 (generations 450 and 550); lineage 3 (L3) lost light between passages 86 and 93 (generations 690 and 750); lineage 4 (L4) lost light between passages 57 and 64 (generations 450 and 550); lineage 5 (L5) lost light between passages 64 and 78 (generations 550 and 620); and lineage 6 (L6) lost light between passages 57 and 64 (generations 450 and 550).

8.3.2 *In vitro*-adapted *C. rodentium* has improved growth in DM media

Following 5 months adaptation to modified DM media (supplemented with 1% glucose), I estimated that the *C. rodentium* strains had undergone approximately 1,200 generations. To assess the changes in growth capability of *C. rodentium* strains at generation 1,200 (dubbed L_{1G1200}, L_{2G1200}, L_{3G1200}, L_{4G1200}, L_{5G1200}, and L_{6G1200}), growth curves were performed in the ‘host’ DM media. 50 ml tubes containing 10 ml of either modified DM media or LB were inoculated with 20 µl (~1 x 10⁷ CFU) of *in vitro*-adapted strains and the ancestral ICC180 strain and samples taken at regular intervals to measure bacterial numbers (Fig. 8.2.A) and bioluminescence (Fig. 8.2.B). The *in vitro*-adapted strains all showed improved growth in the media to which they had evolved in (Fig. 8.2). When the bacterial numbers over time are converted to area under the curve (AUC) values, the *in vitro*-adapted strains all have statistically significantly higher AUC than the ancestral strain in this media (Fig. 8.2.C). This improved growth appears most pronounced during exponential phase, and the slopes during exponential phase (between 2 and 8 hours) are significantly steeper for the *in vitro*-adapted strains compared with the ancestor (Fig. 8.2.D).

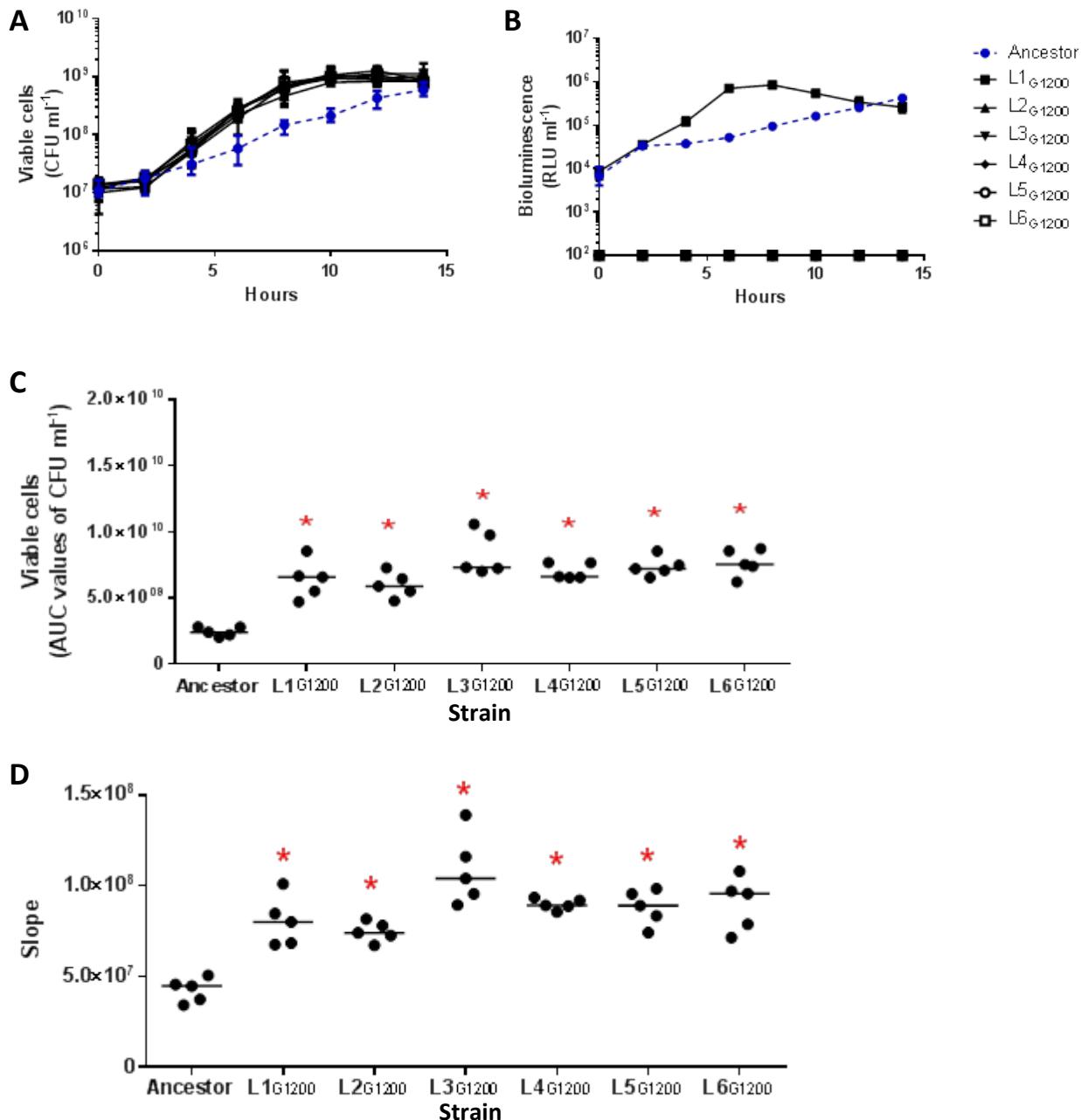


Figure 8.2. *In vitro*-adapted *Citrobacter rodentium* shows improved growth in minimal media. 10 ml of minimal media (modified DM media with 1% glucose) was inoculated with 20 μ l ($\sim 1 \times 10^7$ CFU) of *C. rodentium* and sampled at regular intervals for measuring bacterial numbers (A) and bioluminescence (B). Bacterial numbers were converted into area under the curve (AUC) values (C). The slopes of exponential-phase growth (between 2 and 8 hours) were calculated (D). A non-parametric two-tailed Friedman test was applied to test for variation between all strains for AUC and slope data. A Wilcoxon signed-rank test was used to test for significant differences between the ancestor and each evolved strain; a red star marks statistically significant differences ($p < 0.05$).

8.3.3 *In vitro*-adapted *C. rodentium* has decreased growth in rich media

I also investigated changes in growth capability in rich media (LB). 50 ml tubes containing 10 ml of either modified DM media or LB were inoculated with 20 µl ($\sim 1 \times 10^7$ CFU) of *in vitro*-adapted strains and the ancestral ICC180 strain and samples taken at regular intervals to measure bacterial numbers (Fig. 8.3.A) and bioluminescence (Fig. 8.3.B). The *in vitro*-adapted strains performed slightly worse than the ancestor in rich media (Fig. 8.3). When the bacterial numbers over time are converted to area under the curve values, three out of six lineages have significantly lower AUC values than the ancestor (Fig. 8.3.C), which is reflected in significantly less steep exponential phase slopes (Fig. 8.3.D).

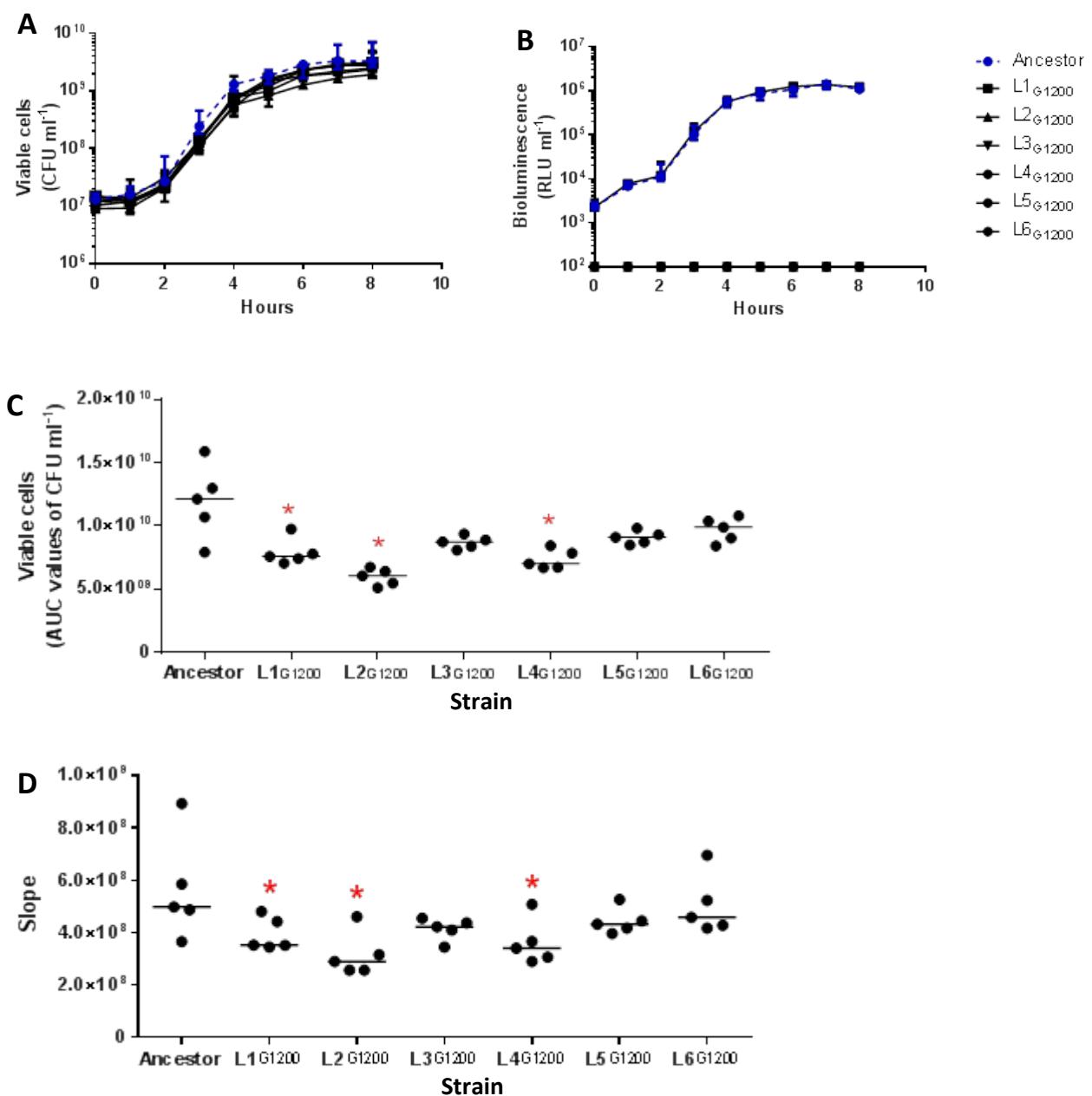


Figure 8.3. *In vitro*-adapted *Citrobacter rodentium* shows reduced growth in rich media. 10 ml of rich media (LB) was inoculated with 20 µl (~1 × 10⁷ CFU) of *C. rodentium* and sampled at regular intervals for measuring bacterial numbers (A) and bioluminescence (B). Bacterial numbers were converted into area under the curve (AUC) values (C). The slopes of exponential-phase growth (between 2 and 4 hours) were calculated (D). A non-parametric two-tailed Friedman test was applied to test for variation between all strains for AUC and slope data. A Wilcoxon signed-rank test was used to test for significant differences between the ancestor and each evolved strain; a red star marks statistically significant differences ($p < 0.05$).

8.4 Discussion

The work that I have presented in this chapter details the complementary *in vitro* evolution experiment which ran alongside the *in vivo* evolution experiment, and which followed traditionally established methods of *in vitro* evolution. I have begun initial investigation into the phenotypes of the *in vitro*-adapted *C. rodentium* by investigating changes in growth of isolated strains. A summary of observed phenotypes is shown in Table 8.1.

8.4.1 Light production appears to confer no benefit *in vitro*

After approximately generation 500, nearly all *in vitro*-adapted *C. rodentium* lineages lost the ability to produce light, along with resistance to kanamycin. Only one strain out of six retained the ability to produce light throughout the five months. This perhaps indicates that light production offers no discernable benefit in an *in vitro* setting compared to the fitness cost required to express the *lux* operon. Previous studies where the *lux* operon has been chromosomally inserted into bacteria have shown promising stability^{263–266}, however not over 5 months of daily passaging.

Table 8.1. Summary of observed changes of *in vitro*-adapted *C. rodentium*.

Strain	Phenotype
L1_{G1200}	Retained light production and kanamycin resistance Improved growth in DM media Reduced growth in LB media
L2_{G1200}	Lost light production and kanamycin resistance Improved growth in DM media Reduced growth in LB media
L3_{G1200}	Lost light production and kanamycin resistance Improved growth in DM media
L4_{G1200}	Lost light production and kanamycin resistance Improved growth in DM media Reduced growth in LB media
L5_{G1200}	Lost light production and kanamycin resistance Improved growth in DM media
L6_{G1200}	Lost light production and kanamycin resistance Improved growth in DM media

8.4.2 Evidence of a growth trade-off for *in vitro*-adapted *C. rodentium*

C. rodentium was adapted to a defined laboratory media (Davis & Mingoli media with 1% glucose supplementation) as a complementary experiment to the *C. rodentium* adapting to the mouse environment. The simplest comparison of their different evolutionary trajectories is assessment of their respective growth in laboratory media, both rich and the same media that the *in vitro*-adapted strains had adapted to. In the DM media, the *in vitro*-adapted strains all consistently grew faster and have steeper exponential phase slopes, have improved overall growth as shown by AUC of CFU ml⁻¹ over time, and reach stationary phase quicker than the ancestral ICC180 strain (Fig. 8.2). However, when grown in rich media (LB), half of the *in vitro*-adapted strains grew slower than the ancestor, with lower AUC of CFU ml⁻¹ over time and less steep exponential phase slopes (Fig. 8.3). The differences are slight, with the slopes varying by less than half (median value of ancestor: 4.99x 10⁸ vs. 3.53 x 10⁸, 2.90 x 10⁸, and 3.41 x 10⁸ for L1_{G1200}, L2_{G1200}, and L4_{G1200}), but are consistent and statistically significant. The growth defect that I observed is perhaps less pronounced due to the nutrient rich LB environment compensating for any reduction in growth due to lower overall concentrations of glucose, which is the carbon source the *in vitro*-adapted strains had adapted to. Trade-offs are a common theme in evolutionary biology²⁶⁷, and the *E. coli* in the LTEE, which evolved to a similar media as the *in vitro*-adapted *C. rodentium*, consistently showed decreased growth on melibiose after 2,000 generations adaptation to the glucose carbon source²¹. Comparing the growth of the *in vitro*-adapted strains in DM media supplemented with an alternative carbon source may highlight its specialisation to growth on glucose and any associated trade-offs.

8.4.3 Conclusions

This work shows that following adaptation to the *in vitro* environment, a moderately predictable and reproducible phenotype emerges: that of a non-bioluminescent *C. rodentium* which has improved growth in the ‘host’ media and slightly reduced growth in a foreign media. Further directions will include further characterisation of the *in vitro*-adapted *C. rodentium*, including further characterisation of any growth trade-offs.

Chapter 9: Discussion and Future Directions

The work detailed in this thesis is the first of its kind: a long-term (5 months) *in vivo* evolution experiment with a complementary *in vitro* evolution experiment, enabling comparison of bacterial adaptation under different conditions, with a study of the phenotypic and genotypic differences present following repeated natural transmissions and adaptation to the mouse environment. While in evolutionary terms, and in comparison to the time which had passed between laboratory outbreaks of *C. rodentium*, 5 months is a relatively short time span, the stored samples allow continuation of each of the evolutionary lineages and further discoveries to be made. This study also shows that measurable and distinguishable phenotypic changes have occurred within the short time frame of this study, indicating that 5 months is sufficient time for *C. rodentium* to adapt to become more infectious.

9.1 The differences between *in vivo* and *in vitro* evolution

One of the important questions that this thesis seeks to answer is what differences can be found between evolution of *C. rodentium* *in vitro* and evolution in a living animal. To address this, I performed a complementary *in vitro* evolution experiment. *C. rodentium* samples were taken at the end of five months of evolution for each experiment; in the case of the *in vitro* experiment this equated to roughly 1,200 generations, while for the *in vivo* experiment calculation of the generation number was deemed too inaccurate to be useful, and so the strains are referred to as their passage number (P_{20}). The number of “passages” that the *in vitro* strains at generation 1,200 corresponded to just over 150, as the samples were “passaged” through media daily. Therefore, the *in vitro* and *in vivo* strains are comparable in terms of absolute time (5 months), but not in passage number or, presumably, generation number. This is an important concept when comparing the strains from each of these two experiments.

9.1.1 Loss of bioluminescence was observed following adaptation to laboratory media, and not following adaptation to a mouse

Remarkably, throughout the five month time period of the *in vivo* experiment, I observed no loss of light production from *C. rodentium* ICC180, with strains detected and isolated from each animal consistently producing light and retaining resistance to kanamycin. This is in contrast to the *C. rodentium* lineages in the *in vitro* evolution experiment, where the majority of lineages (5/6) reproducibly lost both the ability to produce light and resistance to kanamycin after approximately 500 generations adapting to the Davis & Mingoli media. Previous *in vitro* evolution experiments have indicated that the most rapid adaptive changes occur between 500 and 2,000 generations¹⁸. Loss of light production and loss of kanamycin resistance indicates loss or disruption of the entire insertion. The reproducibility and high rate of loss of light production would suggest that light production is not

beneficial in the laboratory environment and so strains which have lost the insertion would gain an advantage and quickly become the predominant population. This selection is not evident in the *in vivo* experiment, perhaps suggesting the converse: that production of light has some beneficial impact on *C. rodentium* in the mouse model, and selection acts to retain the production of bioluminescence at high rates in the population. One potential benefit of light production *in vivo* could be the detoxification of reactive oxygen species (ROS). During *C. rodentium* infection, neutrophils are recruited to the site of infection following recognition of *C. rodentium* by myeloid differentiation primary-response protein 88 (MYD88)-dependent Toll-like receptor (TLR) signalling^{268,269}. When neutrophils encounter pathogens, nicotinamide adenine dinucleotide phosphate oxidase present on the membrane of the neutrophils produce toxic ROS for combating the infection^{270,271}.

However, as is often the case in biology, things are not as simple as they appear, and there are reports indicating that ROS production is beneficial for some pathogenic bacteria. For example, ROS produced during inflammation reacts with thiosulfate present in the intestines forming tetrathionate, a molecule which acts as an electron acceptor and enables *Salmonella typhimurium* to grow faster and outcompete commensal bacteria²⁷². As yet there is no evidence indicating a similar situation for *C. rodentium*, however it is important to be aware that the relationship between ROS and pathogen may not be as clear cut.

Another possible explanation could be that the lineages which have adapted to the *in vivo* environment have not undergone sufficient generations for loss of bioluminescence to emerge, however as the generation number from within a host cannot be quantified with any useful degree of accuracy, there is no way to directly compare the rate of evolution between the *in vivo* and *in vitro* evolution experiments. Interestingly, I have previously found that when *C. rodentium* chromosomally-tagged with the firefly luciferase infects mice, the *C. rodentium* loses the ability to produce light after a single transmission through mice, indicating that the firefly luciferase system is not kept *in vivo* compared to the bacterial luciferase system (data not shown).

9.1.2 Evidence of growth trade-offs in *in vitro*-adapted *C. rodentium* absent in *in vivo*-adapted *C. rodentium*

The *in vitro*-adapted *C. rodentium* strains had a trend of reduced growth in rich media accompanying increased growth in their ‘adaptive’ environment, DM media. This fitness trade-off is a common theme that has come out of many *in vitro* evolution experiments²⁶⁷. However, the *in vivo*-adapted *C. rodentium* all remained at least as fit as their ancestor in the laboratory media tested, with no indication that their adaptation to the mouse environment came at the cost of fitness in the laboratory environments. While it may well be that the *in vivo*-adapted strains have fitness trade-offs present in as

yet untested environments, that they performed as well in both a rich and restricted media does suggest that any fitness costs, if present, will be minor. This finding has interesting implications, as it highlights the ability of a bacterium to adapt to a complex, *in vivo* environment without any impairment in another environment, as would traditionally be expected.

Perhaps due to the complex nature of the *in vivo* environment, I also observed a far greater variety of phenotypes for the *in vivo*-adapted strains compared with the *in vitro*-adapted strains. Further work will involve greater characterisation of the *in vitro*-adapted strains, as well as sequencing to investigate what genetic changes occurred in *C. rodentium* adapted to DM media, and whether there is a stronger selective pressure and more evidence of adaptation to the laboratory environment versus the mouse environment.

9.2 The effect of nalidixic acid treatment on the evolution of *C. rodentium*

The composition of the microbiota present is known to have an impact on infection and disease susceptibility. For example, Firmicutes such as *Lactobacillus* species and *Bacillus subtilis* have been found to be protective against *C. rodentium* infection^{273,274}. Enrichment with *Lactobacillus* species resulted in lessened symptoms of *C. rodentium* disease, and prevented the expansion of *Gammaproteobacteria* and *Actinobacteria* which normally accompanies *C. rodentium* infection²⁷³. Probiotic treatment with the Actinobacterium *Bifidobacterium breve* reduced the number of *C. rodentium* present in infected animals²⁷⁵. In contrast, a member of the family Bacteroidales, *Bacteroides thetaiotaomicron*, has been shown to worsen *C. rodentium* infection in mice with a microflora consisting of *B. thetaiotaomicron* in comparison to mice with no microflora²⁷⁶. While the 16S sequencing method used did not distinguish between different Bacteroidales to the genus level, OTUs 30, 61, 27, which each contained Bacteroidales, were all relatively more abundant following nalidixic acid treatment compared with animals not receiving antibiotic. OTU 50, also containing Bacteroidales, was relatively more abundant in the animals not receiving antibiotic. The OTUs which contained *Lactobacillus* species (2, 33, and 48) were not different between the two treatment groups (Appendix 1). However, a limitation of this method is that the analyses performed provide a measure of relative abundance, and therefore if the antibiotic treatment has altered all of the OTUs present to the same extent, then this would appear as no difference. As only a relatively small number of OTUs were found to be significantly different with antibiotic treatment versus no treatment, I propose that the antibiotic treatment has altered the microflora of the mice.

During the *in vivo* evolution experiment, I observed no differences between the two treatment groups with regards to bacterial shedding, expressed as AUC values of bacteria shed from each infected animal, pooled according to treatment, throughout the five month experimental evolution period.

However, three of the lineages from the antibiotic treated group (N2, N3, and N5) had a decrease in bacterial shedding over the experiment. When I infected animals with the *in vivo*-adapted strains isolated from the end of the *in vivo* evolution experiment (P_{20}), one of the strains which had adapted to an environment with antibiotic present ($N3_{P_{20}}$) was shed from animals at lower numbers compared with the ancestral strain, whereas two of the strains which adapted to mice without antibiotic treatment ($W1_{P_{20}}$ and $W5_{P_{20}}$) were shed from animals at higher numbers compared with the ancestral strain. This suggests a trend of increased shedding of strains adapted to the antibiotic-free environment, and decreased shedding of strains adapted to mice receiving antibiotic. The observation of increased shedding of the “W” strains may be indicative of increased numbers required due to increased competition from commensal bacteria. For the “N” strains, the trend of decreased *C. rodentium* shedding may indicate that, due to the disruption of the microbiota, infection of mice receiving antibiotic was less difficult and that intraspecies competition was a more important driver. I found mutations in the *C. rodentium* type 6 secretion system gene *cts2S* in only one “W” strain ($W1_{P_{20}}$), while these mutations were found in 4/5 of the “N” strains. Type 6 secretion systems have been shown to be important in competition between bacteria²⁴⁵. With respects to other genetic mutations, there were a large number of shared genetic changes between the two treatment groups, with no particular pattern of certain genes or types of genes more likely to be altered depending on antibiotic treatment.

9.3 Phenotypes and genotypes observed for each of the *in vivo*-adapted lineages

9.3.1 $W1_{P_{20}}$ has improved transmissibility and a history of reduced light from the rectum

The $W1$ lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC values of bacteria shed from infected animals throughout the five month experimental evolution period. No transmission failure events were ever observed, with the $W1$ lineage successfully establishing an infection in each successive animal in the infection chain. Over the course of the experimental evolution period, the AUC of light detected from the rectum of each animal significantly decreased, while AUC of light detected from the abdomen of each animal remained stable.

When $W1_{P_{20}}$ was isolated from the end of the evolution experiment and tested for any changes, I found that $W1_{P_{20}}$ was shed at an overall higher level compared with mice infected with the ancestor. When directly competed with $ICC169$ in co-infected animals, $W1_{P_{20}}$ was shed at significantly lower levels ($p<0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, $W1_{P_{20}}$ made up 8% of the total *C. rodentium* recovered from shed stool (1.20×10^7 CFU gram⁻¹ stool), and was able to transmit to 84% of naïve animals. This is a vast improvement in transmissibility over the ancestral $ICC180$ strain, which was shed at 7% and did not transmit. There

were no notable differences between W1_{P20} and the ancestral strain in any of the *in vitro* assays used.

I sequenced the W1_{P20} population and determined that W1_{P20} has 16 unique SNPs and 27 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 16 out of 38 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 6 are unique to W1_{P20}, comprising *ROD_11451*, which encodes for a LysR family transcriptional regulator; *ROD_12451*, which encodes for aconitase hydratase; *ROD_15951*, which encodes for a GntR family transcriptional regulator; *eutK*, which encodes for ethanolamine utilization protein; and *ROD_27781*, which encodes for fimbrial usher protein. LysR is a transcriptional regulator, activation of which leads to transcription of Ler and subsequent transcription of LEE genes important for virulence²⁷⁷. Altering this gene is likely to have major downstream effects. It would be interesting to investigate whether such changes contribute to the improved transmissibility of W1_{P20}.

9.3.2 W2_{P20} has improved transmissibility and a history of reduced light from the rectum

The W2 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC of bacteria shed from infected animals throughout the five month experimental evolution period. A single transmission failure event early in the infection chain occurred, at passage number 6, however this was never detected again and the following population of W2 was able to successfully infect all remaining animals in the infection chain. Over the course of the experimental evolution period, the AUC of light detected from the abdomen of each animal significantly decreased, while AUC of light detected from the rectum of each animal remained stable.

When W2_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that W2_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, W2_{P20} was shed at significantly lower levels ($p < 0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, W2_{P20} comprised 10% of the total *C. rodentium* shed (recovered at 1.95×10^7 CFU gram⁻¹ stool), and was able to transmit to 34% of naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. There were no notable differences between W2_{P20} and the ancestral strain in any of the *in vitro* assays used.

I sequenced the W2_{P20} population and identified 7 unique SNPs and 17 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 12 out of 20 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a

predicted effect on the protein, 3 are unique to W2_{P20}, comprising *ROD_05911*, which encodes for transketolase; and *ROD_34781* and *ROD_41482*, which encode for hypothetical proteins. Transketolase is involved in metabolism, and alterations to it perhaps indicate further adaptation to the mouse gut. Genes involved in metabolism have been shown to be important for virulence in other microorganisms^{278,279}, and one could suggest that improved metabolism in a particular environment would give an advantage to the bacterium, leading towards more efficient colonisation and an observable improvement in transmission between animals.

9.3.3 W3_{P20} performs the same *in vivo* as the ancestor and has the greatest number of genetic changes

The W3 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC of bacteria shed from infected animals throughout the five month experimental evolution period. No transmission failure events were ever observed, with the W3 lineage successfully established an infection in each successive animal in the infection chain. When the AUC values for each animal throughout the experimental evolution period were pooled together and each of the lineages compared, the W3 lineage had the lowest pooled AUC data which was significantly lower than the lineage with the highest pool AUC data, N1. Over the course of the evolution period, the AUC of light detected from the abdomen of each animal significantly decreased, while AUC of light detected from the rectum of each animal remained stable.

When W3_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that W3_{P20} is shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, W3_{P20} was shed at significantly lower levels ($p < 0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, W3_{P20} comprised 19% of the total *C. rodentium* shed (recovered at 1.05×10^7 CFU gram⁻¹ stool), but was unable to transmit any naïve animals. This indicates no improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and also did not transmit. There were no notable differences between W3_{P20} and the ancestral strain in any of the *in vitro* assays used.

I sequenced the population of W3_{P20} and determined that W3_{P20} has the greatest total number of SNPs present, and the greatest number of unique intragenic SNPs. W3_{P20} has 22 unique SNPs and 28 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 29 out of 43 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 15 are unique to W3_{P20}, comprising *cstA*, which encodes carbon starvation protein A; *sucB*, which encodes dihydrolipoamide succinyltransferase component (E2); *astC*, which encodes for succinylornithine transaminase; *ROD_20671*, which encodes

for an integrase; *eutJ*, which encodes for ethanolamine utilization protein; *eutE*, which encodes for ethanolamine utilization aldehyde dehydrogenase; *ROD_35621*, which encodes for radical SMA superfamily protein; *pflC*, which encodes for pyruvate formate-lyase 2 activating enzyme; *ROD_38821*, which encodes for aldose 1-epimerase; *ROD_39511*, which encodes for an amino acid permease; *ggt*, which encodes for gamma-glutamyltranspeptidase; *ROD_48371*, which encodes for a carbon starvation protein, and *ROD_38891*, which encodes for a hypothetical protein. Many of these genes are involved in metabolism, however as no detectable improvement in *in vivo* fitness was observed I suggest that some of these changes may have a neutral or even negative impact on the bacterium.

9.3.4 W4_{P20} has improved transmissibility and no notable unique intragenic SNPs

The W4 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC of bacteria shed from infected animals throughout the five month experimental evolution period. Two transmission failure events occurred in the infection chain, at passage number 7 and 18, however following each event a chain of successful transmission was resumed and so there was little indication that the transmitting populations of W4 were losing their transmissibility. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations. The W4 lineage had the highest AUC of light detected from the abdomen of infected animals over the experimental evolution period, which was statistically significantly higher than all other lineages excluding N1 and N4.

When W4_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that W4_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, W4_{P20} was shed at similar levels to ICC169, which indicates an improvement over the ancestor. At the peak of infection of these co-infected animals, W4_{P20} comprised 23% of the total *C. rodentium* shed (recovered at 4.54×10^7 CFU gram⁻¹ stool), and was able to transmit to all naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. There were no notable differences between W4_{P20} and the ancestral strain in any of the *in vitro* assays used.

I sequenced the population of W4_{P20} at the end of the evolution experiment and determined that W4_{P20} has 10 unique SNPs and 15 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 6 out of 22 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, none are unique to W4_{P20}. This may mean that either the intergenic changes are responsible for the improved transmissibility of W4_{P20}, or that the phenotype is the result of one or more of the changes shared with

another strain. Two SNPs present in the gene *cbiH*, which encodes for cobalt-precorrin-3B C(17)-methyltransferase, are also shared with other strains which exhibit improved transmissibility, namely W5_{P20}, N1_{P20}, N2₂₀, N4_{P20}, and N5_{P20}, suggesting that *cbiH* may be a good starting point for investigating the genetic basis for improved transmissibility. *cbiH* is involved in vitamin B12 biosynthesis²³⁹. It may be that the *in vivo*-adapted bacteria have a reduced need to produce vitamin B12 as they may be able to utilise the vitamin produced by other bacteria. However, as there was no decrease in growth *in vitro* for these strains, they must still be able to synthesise some vitamin.

9.3.5 W5_{P20} is shed at higher levels and has increased growth in DM media

The W5 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC of bacteria shed from infected animals throughout the five month experimental evolution period. No transmission failure events were ever observed, with the W5 lineage successfully establishing an infection in each successive animal in the infection chain. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When W5_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that W5_{P20} was shed from infected mice at an overall higher level compared with the ancestor. When directly competed with ICC169, W5_{P20} was shed at significantly higher levels than ICC169 ($p<0.05$), which indicates major improvement over the ancestor. Significantly higher light was detected from the abdomen and rectum of these same animals, indicating greater numbers of W5_{P20} overall. At the peak of infection of these co-infected animals, W5_{P20} comprised 72% of the total *C. rodentium* shed (recovered at 8.00×10^7 CFU gram⁻¹ stool), and was able to transmit to all naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. W5_{P20} was found to have an increased growth rate and to produce more light when grown in the defined modified DM media.

I sequenced the population of W5_{P20} at the end of the evolution experiment and determined that W5_{P20} has 12 unique SNPs and 32 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 16 out of 39 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 2 are unique to W5_{P20}, comprising *fcl*, which encodes for GDP-L-fucose synthetase, and *ROD_38291*, which encodes for a hypothetical protein. GDP-L-fucose synthetase plays a role in GDP-fucose biosynthesis from GDP-mannose; in *Helicobacter pylori* GDP-L-fucose is suggested to be important for adhesion to gastric epithelium²⁸⁰.

9.3.6 N1_{P20} has improved transmissibility and improved growth in DM media

The N1 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC of bacteria shed from infected animals throughout the five month experimental evolution period. No transmission failure events were ever observed, with the N1 lineage successfully establishing an infection in each successive animal in the infection chain. When the AUC values from each animal throughout the experimental evolution period were pooled together and each of the lineages compared, the N1 lineage had the highest pooled AUC data which was statistically significantly higher than the lineage with the lowest pool AUC data, W3. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When N1_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that N1_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, N1_{P20} was shed at significantly lower levels than ICC169 ($p<0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, N1_{P20} comprised 15% of the total *C. rodentium* shed (recovered at 2.73×10^7 CFU gram⁻¹ stool), and was able to transmit to half of the naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. N1_{P20} was found to have an increased growth rate and to produce more light when grown in the defined modified DM media.

I sequenced the population of N1_{P20} at the end of the evolution experiment and determined that N1_{P20} has 13 unique SNPs and 25 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 22 out of 33 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 10 are unique to N1_{P20}, comprising *ROD_14451*, which encodes for an oxidoreductase; *cbiH*, which encodes for cobalt-precorrin-3B C(17)-methyltransferase; *mgtE*, which encodes for a magnesium transporter; *eutJ*, which encodes for ethanolamine utilization protein; *fimD*, which encodes for fimbrial usher protein FimD; *gcp*, which encodes for O-sialoglycoprotein endopeptidase; *ROD_40471*, which encodes for a sodium solute symporter; and *ROD_16371*, *ROD_20381* and *ROD_34781*, which encode for hypothetical proteins. The combination of a mutated *cbiH*, which was discussed in more detail previously, as well as a number of altered genes involved in metabolism may explain both the improvement in *in vitro* growth as well as improvement in transmission between mice.

9.3.7 N2_{P20} has improved transmissibility while being shed at lower numbers

The N2 lineage demonstrated a significant decrease in the calculated AUC values for bacteria shed from infected animals through the five month experimental evolution period, indicating an overall trend of lower AUC over time. It is important to note that the AUC calculations take into account the peak of infection as well as the duration of high bacterial shedding, and so a reduced AUC value could indicate a lower overall peak or a shortened recorded infection. A single transmission failure event occurred early in the infection chain, at passage number 5, however this was never detected again and the following population of N2 was able to successfully infect all remaining animals in the infection chain. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When N2_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that N2_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, N1_{P20} was shed at significantly lower levels than ICC169 ($p<0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, N1_{P20} comprised 4% of the total *C. rodentium* shed (recovered at 7.42×10^6 CFU gram⁻¹ stool), and was able to transmit to all of the naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. As N2_{P20} was shed at a lower percentage than the ancestor, this improvement in transmissibility cannot be due to the number of N2_{P20} shed but must have another explanation. There were no notable differences between N2_{P20} and the ancestral strain in any of the *in vitro* assays used.

I sequenced the population of N2_{P20} at the end of the evolution experiment and determined that N2_{P20} has 9 unique SNPs and 23 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 16 out of 32 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 6 are unique to N2_{P20}, comprising *cstA*, which encodes for carbon starvation protein A; *fcl*, which encodes for GDP-L-fucose synthetase; *ROD_39711*, which encodes for a permease; and *ROD_48371*, which encodes for a carbon starvation protein. As previously stated, GDP-L-fucose plays a role in *H. pylori* adhesion²⁸⁰, and the *fcl* gene in N2_{P20} has multiple mutations. This suggests that *fcl* would be an interesting candidate gene to investigate in relation to the improved transmissibility of this strain.

9.3.8 N3_{P20} has decreased shedding and an increased mutation rate

Similar to the N2 lineage, the N3 lineage had a statistically significant decrease in the calculated AUC values for bacteria shed from infected animals through the five month experimental evolution period, indicating an overall trend of lower AUC values over time. A single transmission failure event occurred late in the infection chain, at passage number 20. Due to its late appearance, it cannot be said whether this event would be followed by more, indicative of reduced transmissibility of the N3 population. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When N3_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that N3_{P20} was shed at a significantly lower level compared to the ancestor in infected mice. When directly competed with ICC169, N3_{P20} was shed at significantly lower levels than ICC169 ($p<0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, N3_{P20} comprised 3% of the total *C. rodentium* shed (recovered at 8.60×10^6 CFU gram⁻¹ stool), and was unable to transmit to any of the naïve animals. This indicates no improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and also did not transmit. Overall, N3_{P20} appears to have a disadvantage *in vivo*, showing no improvements in the co-infection model or any improvement in ability to transmit, and having lower numbers shed in the single infection model. This is paralleled from the evolution data which showed a downward trend in bacterial shedding from the N3 lineage. When the N3_{P20} was investigated for changes in *in vitro* assays, N3_{P20} was found to have a 10-fold increased mutation rate when exposed to the antibiotics rifampicin and streptomycin.

I sequenced the population of N3_{P20} at the end of the evolution experiment and determined that N3_{P20} has 22 unique SNPs and 7 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 12 out of 21 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 6 are unique to N3_{P20}, comprising: *rluE*, which encodes for ribosomal large subunit pseudouridine synthase E; *pfkB*, which encodes for 6-phosphofructokinase; *rluB*, which encodes for ribosomal large subunit pseudouridine synthase B; *cbiH*, which encodes for cobalt-precorrin-3B C(17)-methyltransferase; *cts2S*, which encodes for the type 6 secretion protein Cts2S; and *radA*; which encodes for a DNA repair protein. The unique mutations in *radA* could potentially be the cause for the increased mutation rate observed *in vitro*. RadA is important for successful recombination, and bacteria with non-functional *radA* genes are recombination-deficient²⁸¹. In *E. coli*, mutations in *radA* also require mutations in *recG* or *ruvABC* to impair recombination due to redundancy in the functions of these proteins²⁴⁸, however in *Pseudomonas aeruginosa* mutations in *radA* have been shown to increase mutation rate by 15-fold²⁸², which is comparable to the 10-fold increase I observed for N3_{P20}.

9.3.9 N4_{P20} transmits preferentially to ICC169 *in vivo* and has increased growth in DM media

The N4 lineage had a fairly stable evolutionary history, with no significant increase or decrease in the AUC values for bacteria shed from infected animals throughout the five month experimental evolution period. No transmission failure events were ever observed, with the N4 lineage successfully established an infection in each successive animal in the infection chain. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When N4_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that N4_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, N4_{P20} was shed at similar levels to ICC169, which indicates an improvement over the ancestor. Significantly higher light was detected from the abdomen, but not the rectum, of these same animals, indicating a greater carriage of N4_{P20} overall. At the peak of infection of these co-infected animals, N4_{P20} comprised 24% of the total *C. rodentium* shed (recovered at 5.35×10^7 CFU gram⁻¹ stool), and was able to transmit to all of the naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. This improvement in transmissibility may be due to the higher percentage of N4_{P20} shed, however notably in these animals ICC169 was unable to transmit to the naïve animals. This suggests that N4_{P20} is in some way preventing ICC169 from transmitting. This does not appear to be due to the number of N4_{P20} present, as ICC169 was present at levels which were shown to be capable to transmitting to all of the naïve animals in another infection group (W5_{P20}).

When investigated for changes in *in vitro* assays, I found that N4_{P20} has an increased growth rate and produces more light when grown in the defined modified DM media. The ability of N4_{P20} to transmit preferentially over ICC169 in the animal studies did not appear to be due to changes in survival in the environment, as shown by the desiccation assay, or overall attachment rate, as assessed using a mouse fibroblast assay. It may be that N4_{P20} preferentially attaches to the fibroblasts and/or displaces the competing ICC169, which is outside the scope of the performed experiment, and which could be investigated by co-infecting the fibroblast cells with N4_{P20} and ICC169. Another possible explanation could be changes in gene expression, an avenue which is currently being investigated.

I sequenced the population of N4_{P20} at the end of the evolution experiment and determined that N4_{P20} has 17 unique SNPs and 28 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 22 out of 39 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 7 are unique to N4_{P20}, comprising: *ROD_05511*, which encodes for a permease; *ROD_15031*, which encodes for an ABC transporter ATP-binding protein; *ROD_20671*, which encodes for an integrase; *cbiH*, which

encodes for cobalt-precorrin-3B C(17)-methyltransferase; *ROD_22911*, which encodes for a nucleoside transporter; *bcsC*, which encodes for cellulose synthase operon protein C; and *pabA*, which encodes for para-aminobenzoate synthase glutamine aminotransferase component I.

The type 6 secretion system (T6SS) which is possessed by a number of Gram-negative bacterial species, including *P. aeruginosa*, *V. cholerae*, *Serratia marcescens* and Enteropathogenic *E. coli*, has recently been shown to play an antibacterial role, and therefore be important in inter- and intra-species competition^{283–286}. In *S. marcescens*, it has been shown that the T6SS is important for intraspecies competition, with two Rhs family antibacterial proteins secreted by the T6SS^{286,287}. Similarly, in *C. rodentium* the *ctsI* T6SS has been shown to be important in interbacterial competition, providing *C. rodentium* a growth advantage when competed with *E. coli*²⁴⁵. While I did not observe any mutations in *ctsI*, N4_{P20} does have two SNPs in *cts2S*, a gene in the second unique T6SS present in *C. rodentium*. However, these mutations alone or in combination were not unique to N4_{P20}, and were also present in other *in vivo*-adapted strains which did not have the same preferential colonisation phenotype. Clearly, there is further work to be done with the N4_{P20} isolate to determine the mechanism of its preferential transmission, and potential intraspecies competition advantage, as well as when its striking phenotype emerged.

9.3.10 N5_{P20} has improved transmissibility and aggregates in LB media

Similar to the N2 and N3 lineages, the N5 lineage had a statistically significant decrease in the calculated AUC values of bacteria shed from infected animals through the five month experimental evolution period, indicating an overall trend of lower AUC over time. Two transmission failure events occurred in the infection chain, at passage number 7 and 15, however following each event a chain of successful transmission was resumed and so there was little indication that the transmitting populations of N5 were losing their ability to transmit. The AUC of light detected from both the abdomen and the rectum of each animal remained stable for the experimental evolution period, with no reduction in light indicating no reduction in bacterial burden for each of these locations.

When N5_{P20} was isolated from the end of the evolution experiment and tested for any changes, I found that N5_{P20} was shed at a similar level to the ancestor in infected mice. When directly competed with ICC169, N5_{P20} was shed at significantly lower levels than ICC169 ($p < 0.05$), which indicates no change from the ancestor. At the peak of infection of these co-infected animals, N5_{P20} comprised 13% of the total *C. rodentium* shed (recovered at 1.66×10^7 CFU gram⁻¹ stool), and was able to transmit to all of the naïve animals. This indicates an improvement in transmissibility over the ancestral ICC180 strain, which was shed at 7%, and did not transmit. This improvement in transmissibility may be due to the higher percentage of N5_{P20} shed.

When investigated for changes in *in vitro* assays, I found that N5_{P20} formed aggregates when grown in LB for 4 hours. The aggregates were incomplete, with approximately 10⁶ CFU ml⁻¹ remaining planktonic in the media. Aggregates were of variable size and had a tendency to float to the liquid/air interface. When stained with calcofluor white, a stain which binds to β-1,3 and β-1,4 polysaccharides, such as cellulose and chitin^{146,147}, the aggregates fluoresced under UV light which is indicative of the presence of cellulose. However, the aggregates were not dispersed or prevented by the addition of cellulase, and their composition remains unclear. Under light microscopy, the aggregates appear to be made of condensed packs of N5_{P20} cells, and gram stained slides of N5_{P20} show diffuse lightly stained areas around the cells, perhaps indicating the presence of an excreted substance.

I sequenced the population of N5_{P20} at the end of the evolution experiment and determined that N5_{P20} has 8 unique SNPs and 18 SNPs shared with at least one of the other lineages. Of the intragenic SNPs, 14 out of 23 had a PROVEAN score of -2.5 or below, and therefore have a predicted deleterious or non-neutral effect on the protein. Of these SNPs with a predicted effect on the protein, 2 are unique to N5_{P20}, comprising: *narG*, which encodes for respiratory nitrate reductase 1 subunit alpha, and *ROD_47661*, which encodes for a hypothetical protein. Neither of these genes are among those known to have a role in bacterial aggregation^{288–293}.

While further investigation is needed to determine the genetic basis of clumping in *C. rodentium*, I have not yet determined whether the clumping phenotype displayed by N5_{P20} in LB media also occurs *in vivo*, and if so, whether it provides any advantage during infection or transmission. One could imagine the formation of aggregates inside the colon, with the outer cells intimately attached to intestinal cells, and the aggregate forming a thick biofilm layer within the colon. Such a biofilm could lead to a larger and more stable *C. rodentium* population *in vivo* with more sustained bacterial shedding. However, my experiments do not currently support this scenario and only an improved transmission efficiency was observed for N5_{P20}.

9.4 Future directions

As with any body of research, the number of questions answered is less than the number of questions raised. Matching genotype to phenotype, which is the holy grail of experimental evolution, is a massive undertaking with respects to this study given the number and variety of mutations present for each strain. An epistasis effect is likely to be occurring, with many of the combinations of mutations working together to produce a particular phenotype. In order to tease apart which combination of mutations result in what phenotype, a wide range of deletion mutants would need to be constructed, with each mutation made in the ancestral strain to confirm its effect. In addition, the presence of subpopulations which may be working together, adds greater complexity to untangling the phenotypes observed and the genetic changes that may be responsible for them. I would propose to begin initial experiments focussing on SNPs unique to a given strain and common to a given phenotype.

The length of the evolution experiments, while long for a given experiment, is only a blink of an eye in evolutionary terms, and so the potential for further adaptation would be worth exploring. Several questions remain: have the isolates reached their ‘peak’ fitness? Will their fitness plateau, and if so, when? What further phenotypic changes and branching of lineages may occur? As well as looking to the future, looking back into the past may also hold important lessons. I anticipate that investigating the historical isolates stored in the freezer would clarify when the observed changes occurred and if any other phenotypes were present before the observed phenotypes from the isolates at the five month time point. Further investigation around the transmission failure events in the *in vivo* evolution experiment would also allow us to determine if these events really were sporadic and due to host effects, or whether there was any particular genetic background which disposed those populations to failure.

Further investigation and sequencing of the *in vitro*-adapted *C. rodentium* strains would allow comparison of the genetic changes present and selected for in the *in vivo* and *in vitro* environments, and would perhaps give insight into what changes are unnecessary in the evolution to the *in vivo* environment. Preliminary experiments on two of the *in vitro*-adapted *C. rodentium* isolates indicates that following adaptation to laboratory media, the *C. rodentium* isolates have lost the ability to infect C57BL/6 mice (data not shown). Given the two very different evolutionary trajectories, a further comparison between the *in vitro*-adapted and *in vivo*-adapted isolates would be worth investigating.

It is important to note that I did not observe any sign of increased severity of disease in the mice infected with any *in vivo*-adapted strains. Adaptation and efficient reproduction, survival, and spread to new hosts did not require increased disease severity, symptom severity, or a more virulent disease. This may not be the case for other hosts, and as *C. rodentium* infection has a different prognosis depending on the genetic background of the host animal, testing the *in vivo*-adapted *C. rodentium* isolates in different strains of mice may offer further insights. I would begin by testing the *in vivo*-adapted strains

in BALB/C mice, which are relatively more resistant to *C. rodentium* infection than C57BL/6 mice, as well as testing in mice which lack the p50 subunit of the transcription factor nuclear factor kappa B, which are unable to clear *C. rodentium* infection²⁹⁴. Testing the *in vivo*-adapted strains in a breed of mice which suffer severe infections and mortality, such as C3H/HeJ mice²⁹⁵, may provide clues as to whether *in vivo* adaptation led to greater virulence or perhaps reduced virulence.

9.5 Conclusions

I have shown that phenotypic differences can be observed following adaptation of *C. rodentium* in mice over a five month experimental evolution period, including improvements in transmissibility. I have also created a wealth of resources for future analysis, with a complete fossil record for ten independent lineages evolving to laboratory mice, and six independent lineages evolving to laboratory media. This work clearly demonstrated that evolution in a tube versus evolution in a real-world environment differs, and assumptions made from one condition cannot necessarily apply to adaptation to another. I have shown in this thesis how *C. rodentium* gains different phenotypic traits when adapted to *in vitro* and *in vivo* environments, the most noticeable being the loss of bioluminescence following *in vitro* adaptation.

Within the *in vivo* evolution experiment, I initially had three hypotheses, which I again list here:

- 1) Low-dose nalidixic acid treatment will alter normal microflora, reducing the amount of genetic information available for recombination purposes and thus limiting adaptation,
- 2) Low-dose nalidixic acid treatment will alter normal microflora, removing the barrier for colonisation of the mice and enabling efficient infection and faster adaptation
- 3) Low-dose nalidixic acid treatment will have no impact on the *in vivo* adaptation of *C. rodentium*

I have evidence that the antibiotic treatment did indeed alter the normal microflora, although what impact this has had on the adaptation of *C. rodentium* to the mouse host remains unclear. I did observe a greater variety of phenotypes present in the strains which had adapted to the presence of nalidixic acid compared with those that had adapted without antibiotic, although whether this is a true reflection of faster adaptation and employment of a greater variety of adaptive processes, or if this is only an artefact of the types of assays which were used to assess phenotypic changes, requires further experimental work. The majority of strains from each treatment group showed evidence of improved transmissibility to new hosts, highlighting *C. rodentium*'s already efficient ability to infect and adapt, and this trait did not require the assistance of antibiotic treatment to emerge. Investigation of earlier time points in the *in vivo* evolution experiment may offer insight to the remaining hypotheses.

The work detailed in this thesis contains the longest and largest *in vivo* evolution experiment that follows the natural infection, adaptation, and transmission of a bacterial pathogen through its natural mammalian host. Samples throughout the experiment were frozen and stored, providing a complete “fossil record” for future analysis. The experiment is also able to be restarted and continued in the future, providing a wealth of information for evolutionary biologists. Much like evolution itself, unravelling the secrets of bacterial adaptation is an ongoing process.

Appendices

Appendix 1. Complete filtered list of OTUs present in C57BL/6 mice.

OTU	Phylum; Class	Order; Family	Genus species
1	Verrucomicrobia; Verrucomicrobiae	Verrucomicrobiales; Verrucomicrobiaceae	<i>Akkermansia muciniphila</i>
2	Firmicutes; Bacilli	Lactobacillales; Lactobacillaceae	<i>Lactobacillus</i> spp.
3	Bacteroidetes; Bacteroidia	Bacteroidales	
4	Bacteroidetes; Bacteroidia	Bacteroidales	
5	Proteobacteria; Betaproteobacteria	Burkholderiales; Alcaligenaceae	<i>Sutterella</i> spp.
7	Bacteroidetes; Bacteroidia	Bacteroidales	
8	Bacteroidetes; Bacteroidia	Bacteroidales	
10	Bacteroidetes; Bacteroidia	Bacteroidales	
11	Bacteroidetes; Bacteroidia	Bacteroidales; Bacteroidaceae	<i>Bacteroides acidifaciens</i>
12	Bacteroidetes; Bacteroidia	Bacteroidales	
13	Firmicutes; Clostridia	Clostridiales	
14	Bacteroidetes; Bacteroidia	Bacteroidales	
15	Bacteroidetes; Bacteroidia	Bacteroidales	
16	Bacteroidetes; Bacteroidia	Bacteroidales; Bacteroidaceae	<i>Bacteroides</i> spp.
18	Bacteroidetes; Bacteroidia	Bacteroidales	
19	Bacteroidetes; Bacteroidia	Bacteroidales; Prevotellaceae	<i>Prevotella</i> spp.
20	Bacteroidetes; Bacteroidia	Bacteroidales	
21	Bacteroidetes; Bacteroidia	Bacteroidales	
22	Bacteroidetes; Bacteroidia	Bacteroidales; Odoribacteraceae	<i>Odoribacter</i> spp.
23	Bacteroidetes; Bacteroidia	Bacteroidales; Prevotellaceae	<i>Prevotella</i> spp.
24	Bacteroidetes; Bacteroidia	Bacteroidales	
25	Bacteroidetes; Bacteroidia	Bacteroidales	
26	Bacteroidetes; Bacteroidia	Bacteroidales	
27	Bacteroidetes; Bacteroidia	Bacteroidales	
28	Bacteroidetes; Bacteroidia	Bacteroidales	
29	Bacteroidetes; Bacteroidia	Bacteroidales	
30	Bacteroidetes; Bacteroidia	Bacteroidales	
31	Firmicutes; Clostridia	Clostridiales	
32	Actinobacteria; Actinobacteria	Bifidobacteriales; Bifidobacteriaceae	<i>Bifidobacterium pseudolongum</i>

33	Firmicutes; Bacilli	Lactobacillales; Lactobacillaceae	<i>Lactobacillus reuteri</i>
34	Bacteroidetes; Bacteroidia	Bacteroidales	
35	Bacteroidetes; Bacteroidia	Bacteroidales	
36	Firmicutes; Clostridia	Clostridiales	
37	Bacteroidetes; Bacteroidia	Bacteroidales	
38	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
39	Bacteroidetes; Bacteroidia	Bacteroidales	
40	Bacteroidetes; Bacteroidia	Bacteroidales	
42	Bacteroidetes; Bacteroidia	Bacteroidales	
43	Bacteroidetes; Bacteroidia	Bacteroidales; Prevotellaceae	
44	Firmicutes; Clostridia	Clostridiales	
45	Bacteroidetes; Bacteroidia	Bacteroidales; Rikenellaceae	
46	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
47	Bacteroidetes; Bacteroidia	Bacteroidales; Rikenellaceae	
48	Firmicutes; Bacilli	Lactobacillales; Lactobacillaceae	<i>Lactobacillus</i> spp.
49	Bacteroidetes; Bacteroidia	Bacteroidales	
50	Bacteroidetes; Bacteroidia	Bacteroidales	
51	Firmicutes; Clostridia	Clostridiales	
52	Bacteroidetes; Bacteroidia	Bacteroidales; Rikenellaceae	
53	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
54	Firmicutes; Bacilli	Turicibacterales; Turicibacteraceae	<i>Turicibacter</i> spp.
55	Bacteroidetes; Bacteroidia	Bacteroidales; Rikenellaceae	
56	Firmicutes; Clostridia	Clostridiales	
57	Firmicutes; Clostridia	Clostridiales	
58	Firmicutes; Clostridia	Clostridiales; Clostridiaceae	<i>Candidatus</i> <i>Arthromitus</i> spp.
59	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
60	Firmicutes; Clostridia	Clostridiales	
61	Bacteroidetes; Bacteroidia	Bacteroidales	
62	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
63	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
64	Bacteroidetes; Bacteroidia	Bacteroidales	
65	Firmicutes;	Clostridiales;	<i>Oscillospira</i> spp.

	Clostridia	Ruminococcaceae	
66	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
67	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
68	Firmicutes; Clostridia	Clostridiales; Clostridiaceae	
72	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
74	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
75	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Ruminococcus</i> spp.
76	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Anaerotruncus</i> spp.
77	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	
78	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Ruminococcus</i> spp.
79	Proteobacteria; Deltaproteobacteria	Desulfovibrionales; Desulfovibrionaceae	<i>Bilophila</i> spp.
80	Bacteroidetes; Bacteroidia	Bacteroidales	
85	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
87	Firmicutes; Clostridia	Clostridiales	
90	Firmicutes; Clostridia	Clostridiales	
92	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
94	Firmicutes; Clostridia	Clostridiales	
100	Firmicutes; Clostridia	Clostridiales	
101	Firmicutes; Clostridia	Clostridiales	
105	Actinobacteria; Coriobacteriia	Coriobacterales; Coriobacteriaceae	<i>Adlercreutzia</i> spp.
106	Firmicutes; Clostridia	Clostridiales	
107	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Ruminococcus</i> spp.
109	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
112	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	<i>Ruminococcus</i> <i>gnavus</i>
113	Firmicutes; Clostridia	Clostridiales	
117	Firmicutes; Clostridia	Clostridiales	
118	Bacteroidetes; Bacteroidia	Bacteroidales	
119	Firmicutes; Clostridia	Clostridiales	
121	Firmicutes; Clostridia	Clostridiales	
124	Firmicutes; Clostridia	Clostridiales	
126	Firmicutes; Clostridia	Clostridiales	
129	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.

133	Firmicutes; Clostridia	Clostridiales; Clostridiaceae	
134	Firmicutes; Clostridia	Clostridiales	
137	Firmicutes; Clostridia	Clostridiales	
138	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
141	Firmicutes; Clostridia	Clostridiales	
153	Firmicutes; Clostridia	Clostridiales	
155	Firmicutes; Clostridia	Clostridiales	
163	Firmicutes; Clostridia	Clostridiales	
186	Actinobacteria; Coriobacteriia	Coriobacteriales; Coriobacteriaceae	
187	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	<i>Ruminococcus gnavus</i>
199	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
242	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	<i>Dorea</i> spp.
251	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
292	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	
332	Firmicutes; Clostridia	Clostridiales	
353	Firmicutes; Clostridia	Clostridiales; Lachnospiraceae	<i>Ruminococcus gnavus</i>
388	Firmicutes; Clostridia	Clostridiales	
563	Bacteroidetes; Bacteroidia	Bacteroidales	
564	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.
828	Firmicutes; Clostridia	Clostridiales; Ruminococcaceae	<i>Oscillospira</i> spp.

Appendix 2. Unmapped contigs present in *in vivo*-adapted *C. rodentium* strains.

References

1. Bardell, D. The roles of the sense of taste and clean teeth in the discovery of bacteria by Antoni van Leeuwenhoek. *Microbiological reviews* 47, 121–6 (1983).
2. Gustafsson, B.E. The physiological importance of the colonic microflora. *Scandinavian journal of gastroenterology. Supplement* 77, 117–131 (1982).
3. Zhang, C., Li, S., Yang, L., Huang, P., Li, W., Wang, S., Zhao, G., Zhang, M., Pang, X., Yan, Z., Liu, Y., Zhao, L. Structural modulation of gut microbiota in life-long calorie-restricted mice. *Nature communications* 4, 2163 (2013).
4. Renz-Polster, H., David, M.R., Buist, A.S., Vollmer, W.M., O'Connor, E.A., Frazier, E.A., Wall, M.A. Caesarean section delivery and the risk of allergic disorders in childhood. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 35, 1466–72 (2005).
5. Alkanani, A., Hara, N., Gottlieb, P.A., Ir, D., Robertson, C.E., Wagner, B.D., Frank, D.N., Zipris, D. Alterations in Intestinal Microbiota Correlate with Susceptibility to Type 1 Diabetes. *Diabetes* (2015). doi:10.2337/db14-1847
6. Ussar, S., Griffin, N.W., Bezy, O., Fujisaka, S., Vienberg, S., Softic, S., Deng, L., Bry, L., Gordon, J.I., Kahn, C.R. Interactions between Gut Microbiota, Host Genetics and Diet Modulate the Predisposition to Obesity and Metabolic Syndrome. *Cell metabolism* (2015).
7. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–14 (2012).
8. Mackenzie, B.W., Waite, D.W., Taylor, M.W. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Frontiers in Microbiology* 6, 130 (2015).
9. Buffie, C., Bucci, V., Stein, R.R., McKenney, P.T., Ling, L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A., Littmann, E., van den Brink, M.R., Jenq, R.R., Taur, Y., Sander, C., Cross, J.R., Toussaint, N.C., Xavier, J.B., Pamer, E.G. Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* 517, 205–8 (2015).
10. Hershkovitz, I., Donoghue, H.D., Minnikin, D.E., Besra, G.S., Lee, O.Y., Gernaey, A.M., Galili, E., Eshed, V., Greenblatt, C.L., Lemma, E., Bar-Gal, G.K., Spigelman, M. Detection and molecular characterization of 9000-year-old *Mycobacterium tuberculosis* from a Neolithic settlement in the Eastern Mediterranean. *PloS one* 3, e3426 (2008).
11. Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., Carniel, E. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* 96, 14043–8 (1999).
12. Kupferschmidt, K. Germany. Scientists rush to study genome of lethal *E. coli*. *Science* (New York, N.Y.) 332, 1249–50 (2011).
13. Scheutz, F., Nielsen, E.M., Frimodt-Møller, J., Boisen, N., Morabito, S., Tozzoli, R., Nataro, J.P., Caprioli, A. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro surveillance : bulletin Européen sur les maladies transmissibles = European communicable disease bulletin* 16, (2011).
14. Brzuszkiewicz, E., Thürmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F.D., Boelter, J., Petersen, H., Gottschalk, G., Daniel, R. Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Archives of microbiology* 193, 883–91 (2011).

15. Bennett, A., Hughes, B. Microbial experimental evolution. *American journal of physiology. Regulatory, integrative and comparative physiology* 297, R17–25 (2009).
16. Wu, T.T. Experimental evolution in bacteria. *CRC critical reviews in microbiology* 6, 33–51 (1978).
17. Dallinger, W.H. On the Life-History of a Minute Septic Organism: With an Account of Experiments Made to Determine Its Thermal Death Point. *Proceedings of the Royal Society of London* 27, 332–350 (1878).
18. Lenski, R.E., Rose, M.R., Simpson, S.C., Tadler, S.C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. (1991).
19. Vasi, F., Travisano, M., Lenski, R.E. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. (1994).
20. Travisano, M., Vasi, F., Lenski, R.E. Long-term experimental evolution in *Escherichia coli*. III. Variation among replicate populations in correlated responses to novel environments. (1995).
21. Travisano, M., Lenski, R.E. Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. *Genetics* 143, 15–26 (1996).
22. Cooper, V.S., Schneider, D., Blot, M., Lenski, R.E. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *Journal of Bacteriology* (2001).
23. Cooper, V.S., Lenski, R.E. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407, 736–9 (2000).
24. Cooper, V.S., Bennett, A., Lenski, R.E. Evolution of thermal dependence of growth rate of *Escherichia coli* populations during 20,000 generations in a constant environment. *Evolution* 55, 889–896 (2001).
25. Sleight, S., Wigginton, N., Lenski, R.E. Increased susceptibility to repeated freeze-thaw cycles in *Escherichia coli* following long-term evolution in a benign environment. *BMC evolutionary biology* 6, 104 (2006).
26. Rainey, P.B., Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* 394, 69-72. (1998).
27. MacLean, R.C., Bell, G., Rainey, P.B. The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. *Proceedings of the National Academy of Sciences in the United States of America* 25 (2004).
28. Spiers, A. Wrinkly-Spreader fitness in the two-dimensional agar plate microcosm: maladaptation, compensation and ecological success. *PloS one* 2, e740 (2007).
29. Ellis, C.N., Cooper, V.S. Experimental adaptation of *Burkholderia cenocepacia* to onion medium reduces host range. *Applied and Environmental Microbiology* 76 (2010).
30. Bennett, A.F., Lenski, R.E., Mittler, J.E. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution* 46, 16-30 (1992).
31. Mongold, Bennett & Lenski. Evolutionary adaptation to temperature. IV. Adaptation of *Escherichia coli* at a niche boundary. *Evolution* 50, 35-43 (1996).
32. Mongold, J.A., Bennett, A.F., Lenski, R.E. Evolutionary adaptation to temperature. VII. Extension of the upper thermal limit of *Escherichia coli*. *Evolution* 53, 386-394 (1999).
33. Bennett, A.F., Lenski, R.E. Evolutionary adaptation to temperature II. Thermal niches of experimental lines of *Escherichia coli*. *Evolution* 47, 1-12 (1993).

34. Zhou, A., Hillesland, K.L., He, Z., Schackwitz, W., Tu, Q., Zane, G.M., Ma, Q., Qu, Y., Stahl, D.A., Wall, J.D., Hazen, T.C., Fields, M.W., Arkin, A.P., Zhou, J. Rapid selective sweep of pre-existing polymorphisms and slow fixation of new mutations in experimental evolution of *Desulfovibrio vulgaris*. *The ISME journal* (2015).
35. Weller, C., Wu, M. A Generation Time Effect on the Rate of Molecular Evolution in Bacteria. *Evolution. International journal of organic evolution* (2015).
36. Sniegowski, P.D., Gerrish, P.J., Lenski, R.E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387, 703–5 (1997).
37. Souza, V., Turner, P.E., Lenski, R.E. Long-term experimental evolution in *Escherichia coli*. V. Effects of recombination with immigrant genotypes on the rate of bacterial evolution. *Journal of Evolutionary Biology* 10, 743–769 (1997).
38. Blount, Z.D., Borland, C.Z., Lenski, R.E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 105, 7899–906 (2008).
39. Quandt, E., Deatherage, D., Ellington, A., Georgiou, G., Barrick, J. Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* (2013).
40. Wiser, M.J. The man who bottled evolution. *Science* 342, 790-793 (2013).
41. Lenski, R.E. Evolution in action: a 50,000-generation salute to Charles Darwin. *Microbe Magazine* 6 (2011).
42. Van Hofwegen, D.J., Hovde, C.J., Minnich, S.A. Rapid evolution of citrate utilization by *Escherichia coli* by direct selection requires *citT* and *dctA*. *Journal of Bacteriology* 198 (2016).
43. Roth, J., Maisnier-Patin, S. Reinterpreting Long-Term Evolution Experiments: Is Delayed Adaptation an Example of Historical Contingency or a Consequence of Intermittent Selection? *Journal of bacteriology* 198, 1009–12 (2016).
44. Rainey, P.B., Rainey, K. Evolution of cooperation and conflict in experimental bacterial populations. *Nature* 425, 72-74 (2003).
45. Brockhurst, M.A., Morgan, A.D., Rainey, P.B., Buckling, A. Population mixing accelerates coevolution. *Ecology Letters* 6 (2003).
46. Callahan, B.J., Fukami, T. Fisher, D.S. Rapid evolution of adaptive niche construction in experimental microbial populations. *Evolution* 68, 3307-3316 (2014).
47. Cutter, A., Jovelin, R., Dey, A. Molecular hyperdiversity and evolution in very large populations. *Molecular ecology* 22, 2074–95 (2013).
48. Poltak, S.R., Cooper, V.S. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *The ISME Journal* 5, 369-378 (2011).
49. Korona, R., Nakatsu, C.H., Forney, L.J., Lenski, R.E. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proceedings of the National Academy of Sciences of the United States of America* 13 (1994).
50. Korona, R. Adaptation to structurally different environments. *Proceedings of the Royal Society of London B: Biological Sciences* 263, 1665–1669 (1996).

51. Riley, M.S., Cooper, V.S., Lenski, R.E., Forney, L.J., Marsh, T.L. Rapid phenotypic change and diversification of a soil bacterium during 1000 generations of experimental evolution. *Microbiology* 147, 995–1006 (2001).
52. MacLean, R.C., Vogwill, T. Limits to compensatory adaptation and the persistence of antibiotic resistance in pathogenic bacteria. *Evolution, medicine, and public health* (2014).
53. Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M., Fons, M., Taddei, F. Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut. *Science* 291, 2606–2608 (2001).
54. Nilsson, A.I., Kugelberg, E., Berg, O.G., Andersson, D.I. Experimental adaptation of *Salmonella typhimurium* to mice. *Genetics* 168, 1119–30 (2004).
55. Lee, S., Wyse, A., Lesher, A., Everett, M.L., Lou, L., Holzknecht, Z.E., Whitesides, J.F., Spears, P.A., Bowles, D.E., Lin, S.S., Tonkonogy, S.L., Orndorff, P.E., Bollinger, R.R., Parker, W. Adaptation in a mouse colony monoassociated with *Escherichia coli* K-12 for more than 1,000 days. *Applied and environmental microbiology* 76, 4655–63 (2010).
56. Lüttich, A., Brunke, S., Hube, B., Jacobsen, I. Serial passaging of *Candida albicans* in systemic murine infection suggests that the wild type strain SC5314 is well adapted to the murine kidney. *PloS one* 8, e64482 (2013).
57. McCarthy, A., Loeffler, A., Witney, A.A., Gould, K.A., Lloyd, D.H., Lindsay, J.A. Extensive horizontal gene transfer during *Staphylococcus aureus* co-colonization in vivo. *Genome biology and evolution* 6, 2697–708 (2014).
58. Mikonranta, L., Mappes, J., Laakso, J., Ketola, T. Within-host evolution decreases virulence in an opportunistic bacterial pathogen. *BMC evolutionary biology* 15, 165 (2015).
59. Martins, N., Faria, V., Teixeira, L., Magalhães, S., Sucena, É. Host adaptation is contingent upon the infection route taken by pathogens. *PLoS pathogens* 9, e1003601 (2013).
60. Merrell, D.S., Butler, S.M., Qadri, F., Dolganov, N.A., Alam, A., Cohen, M.B., Calderwood, S.B., Schoolnik, G.K., Camilli, A. Host-induced epidemic spread of the cholera bacterium. *Nature* 417, 642–5 (2002).
61. Price, E.P., Sarovich, D.S., Mayo, M., Tuanyok, A., Drees, K.P., Kaestli, M., Beckstrom-Sternberg, S.M., Babic-Sternberg, J.S., Kidd, T.J., Bell, S.C., Keim, P., Pearson, T., Currie, B.J. Within-Host Evolution of *Burkholderia pseudomallei* over a Twelve-Year Chronic Carriage Infection. *mBio* 4, (2013).
62. Levin, B.R., Bull, J.J. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends in Microbiology* 2, 76–81 (1994).
63. Schoen, C., Kischkies, L., Elias, J., Ampattu, B. Metabolism and virulence in *Neisseria meningitidis*. *Frontiers in cellular and infection microbiology* 4, 114 (2014).
64. Moxon, E.R., Jansen, V.A. Phage variation: understanding the behaviour of an accidental pathogen. *Trends in microbiology* 13, 563–5 (2005).
65. Gagneux, S. Host-pathogen coevolution in human tuberculosis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 367, 850–9 (2012).

66. Casali, N., Nikolayevskyy, V., Balabanova, Y., Ignatyeva, O., Kontsevaya, I., Harris, S.R., Bentley, S.D., Parkhill, J., Nejentsev, S., Hoffner, S.E., Horstmann, R.D., Brown, T., Drobniowski, F. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome research* 22, 735–45 (2012).
67. Moorman, D.R., Mandell, G.L. Characteristics of rifampin-resistant variants obtained from clinical isolates of *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy* 20, 709–13 (1981).
68. Björkman, J., Hughes, D., Andersson, D.I. Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America* 95, 3949–53 (1998).
69. Bina, X., Provenzano, D., Nguyen, N., Bina, J. Vibrio cholerae RND family efflux systems are required for antimicrobial resistance, optimal virulence factor production, and colonization of the infant mouse small intestine. *Infection and immunity* 76, 3595–605 (2008).
70. Peleg, A.Y., Monga, D., Pillai, S., Mylonakis, E., Moellering, R.C. Jr, Eliopoulos, G.M. Reduced susceptibility to vancomycin influences pathogenicity in *Staphylococcus aureus* infection. *The Journal of infectious diseases* 199, 532–6 (2009).
71. Webber, M.A., Bailey, A.M., Blair, J.M., Morgan, E., Stevens, M.P., Hinton, J.C., Ivens, A., Wain, J., Piddock, L.J. The global consequence of disruption of the AcrAB-TolC efflux pump in *Salmonella enterica* includes reduced expression of SPI-1 and other attributes required to infect the host. *Journal of bacteriology* 191, 4276–85 (2009).
72. Padilla, E., Llobet, E., Doménech-Sánchez, A., Martínez-Martínez, L., Bengoechea, J.A., Albertí, S. *Klebsiella pneumoniae* AcrAB efflux pump contributes to antimicrobial resistance and virulence. *Antimicrobial agents and chemotherapy* 54, 177–83 (2010).
73. Havelaar, A.H., Kirk, M.D., Torgerson, P.R., Gibb, H.J., Hald, T., Lake, R.J., Praet, N., Bellinger, D.C., de Silva, N.R., Gargouri, N., Speybroeck, N., Cawthorne, A., Mathers, C., Stein, C., Angulo, F.J., Devleesschauwer, B., World Health Organization Foodborne Disease Burden Epidemiology Reference Group. World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010. *PLoS Med* 12, e1001923 (2015).
74. Kosek, M., Bern, C., Guerrant, R.L. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bulletin of the World Health Organization* 81, 197–204 (2003).
75. Nataro, J.P., Kaper, J.B. Diarrheagenic *Escherichia coli*. *Clinical microbiology reviews* 11, 142–201 (1998).
76. Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., Tauxe, R.V. Food-related illness and death in the United States. *Emerging infectious diseases* 5, 607–25 (1999).
77. Mundy, R., Girard, F., FitzGerald, A., Frankel, G. Comparison of colonization dynamics and pathology of mice infected with enteropathogenic *Escherichia coli*, enterohaemorrhagic *E. coli* and *Citrobacter rodentium*. *FEMS microbiology letters* 265, 126–32 (2006).
78. Brenna, P.C., Fritz, T.E., Flynn, R.J., Poole, C.M. *Citrobacter freundii* associated with diarrhea in a laboratory mice. *Laboratory animal care* 15, 266–75 (1965).
79. Muto, T., Nakagawa, M., Isobe, Y., Saito, M., Nakano, T. Infectious megaenteron of mice. I. Manifestation and pathological observation. *Jpn. J. Med. Sci. Biol.* 22, 363–74 (1969).
80. Ediger, R. D., Kovatch, R. M., Rabstein, M. M. Colitis in mice with a high incidence of rectal prolapse. *Lab. Anim. Sci.* 24, 488–94 (1974).
81. Itoh, K., Maejima, K., Ueda, K., Fujiwara, K. Effect of intestinal flora on megaenteron of mice. *Microbiol. Immunol.* 22, 661–72 (1978).

82. Barthold, S.W., Coleman, G.L., Bhatt, P.N., Osbaldiston, G.W., Jonas, A.M. The etiology of transmissible murine colonic hyperplasia. *Lab. Anim. Sci.* 26, 889–94 (1976).
83. Schauer, D.B., Zabel, B.A., Pedraza, I.F., O'Hara, C.M., Steigerwalt, A.G., Brenner, D.J. Genetic and biochemical characterization of *Citrobacter rodentium* sp. nov. *J. Clin. Microbiol.* 33, 2064–8 (1995).
84. Barthold, S.W., Osbaldiston, G.W., Jonas, A.M. Dietary, bacterial, and host genetic interactions in the pathogenesis of transmissible murine colonic hyperplasia. *Laboratory animal science* 27, 938–45 (1977).
85. de la Puente-Redondo, V.A., Gutiérrez-Martín, C.B., Pérez-Martínez, C., del Blanco, N.G., García-Iglesias, M.J., Pérez-García, C.C., Rodríguez-Ferri, E.F. Epidemic infection caused by *Citrobacter rodentium* in a gerbil colony. *The Veterinary record* 145, 400–3 (1999).
86. Brenner, D.J., Grimont, P.A., Steigerwalt, A.G., Fanning, G.R., Ageron, E., Riddle, C.F. Classification of citrobacteria by DNA hybridization: designation of *Citrobacter farmeri* sp. nov., *Citrobacter youngae* sp. nov., *Citrobacter braakii* sp. nov., *Citrobacter werkmanii* sp. nov., *Citrobacter sedlakii* sp. nov., and three unnamed *Citrobacter* genomospecies. *International journal of systematic bacteriology* 43, 645–58 (1993).
87. Luperchio, S.A., Newman, J.V., Dangler, C.A., Schrenzel, M.D., Brenner, D.J., Steigerwalt, A.G., Schauer, D.B. *Citrobacter rodentium*, the causative agent of transmissible murine colonic hyperplasia, exhibits clonality: synonymy of *C. rodentium* and mouse-pathogenic *Escherichia coli*. *J. Clin. Microbiol.* 38, 4343–50 (2000).
88. Mundy, R., MacDonald, T., Dougan, G., Frankel, G., Wiles, S. *Citrobacter rodentium* of mice and man. *Cellular microbiology* 7, 1697–706 (2005).
89. Wales, A.D., Woodward, M.J., Pearson, G.R. Attaching-effacing bacteria in animals. *Journal of comparative pathology* 132, 1–26 (2005).
90. Barthold, S.W., Coleman, G.L., Jacoby, R.O., Livestone, E.M., Jonas, A.M. Transmissible murine colonic hyperplasia. *Vet. Pathol.* 15, 223–36 (1978).
91. Luperchio, S.A., Schauer, D.B. Molecular pathogenesis of *Citrobacter rodentium* and transmissible murine colonic hyperplasia. *Microbes and infection / Institut Pasteur* 3, 333–40 (2001).
92. Borenshtein, D., Fry, R.C., Groff, E.B., Nambiar, P.R., Carey, V.J., Fox, J.G., Schauer, D.B. Diarrhea as a cause of mortality in a mouse model of infectious colitis. *Genome biology* 9, R122 (2008).
93. Wiles, S., Clare, S., Harker, J., Huett, A., Young, D., Dougan, G., Frankel, G. Organ specificity, colonization and clearance dynamics in vivo following oral challenges with the murine pathogen *Citrobacter rodentium*. *Cellular microbiology* 6, 963–72 (2004).
94. Wiles, S., Pickard, K., Peng, K., MacDonald, T., Frankel, G. In vivo bioluminescence imaging of the murine pathogen *Citrobacter rodentium*. *Infection and immunity* 74, 5391–6 (2006).
95. Yang, J., Tauschek, M., Hart, E., Hartland, E., Robins-Browne, R. Virulence regulation in *Citrobacter rodentium*: the art of timing. *Microbial Biotechnology* (2010).
96. Barthold, S.W. The microbiology of transmissible murine colonic hyperplasia. *Laboratory animal science* 30, 167–73 (1980).
97. Ghaem-Maghami, M., Simmons, C.P., Daniell, S., Pizza, M., Lewis, D., Frankel, G., Dougan, G. Intimin-specific immune responses prevent bacterial colonization by the attaching-effacing pathogen *Citrobacter rodentium*. *Infection and immunity* 69, 5597–605 (2001).

98. Schauer, D.B., Falkow, S. The eae gene of *Citrobacter freundii* biotype 4280 is necessary for colonization in transmissible murine colonic hyperplasia. *Infection and immunity* 61, 4654–61 (1993).
99. Schauer, D.B., Falkow, S. Attaching and effacing locus of a *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infection and immunity* 61, 2486–92 (1993).
100. Deng, W., Vallance, B.A., Li, Y., Puente, J.L., Finlay, B.B. *Citrobacter rodentium* translocated intimin receptor (Tir) is an essential virulence factor needed for actin condensation, intestinal colonization and colonic hyperplasia in mice. *Molecular microbiology* 48, 95–115 (2003).
101. Vallance, B.A., Finlay, B.B. Exploitation of host cells by enteropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 97, 8799–806 (2000).
102. Petty, N.K., Bulgin, R., Crepin, V.F., Cerdeño-Tárraga, A.M., Schroeder, G.N., Quail, M.A., Lennard, N., Corton, C., Barron, A., Clark, L., Toribio, A.L., Parkhill, J., Dougan, G., Frankel, G., Thomson, N.R. The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. *Journal of bacteriology* 192, 525–38 (2010).
103. McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., Kaper, J.B. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proceedings of the National Academy of Sciences of the United States of America* 92, 1664–8 (1995).
104. Hueck, C.J. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and molecular biology reviews : MMBR* 62, 379–433 (1998).
105. Jerse, A.E., Yu, J., Tall, B.D., Kaper, J.B. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proceedings of the National Academy of Sciences of the United States of America* 87, 7839–43 (1990).
106. Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., Finlay, B.B. Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 91, 511–20 (1997).
107. Abe, A., de Grado, M., Pfuetzner, R.A., Sánchez-Sanmartín, C., Devinney, R., Puente, J.L., Strynadka, N.C., Finlay, B.B. Enteropathogenic *Escherichia coli* translocated intimin receptor, Tir, requires a specific chaperone for stable secretion. *Molecular microbiology* 33, 1162–75 (1999).
108. Elliott, S.J., Hutcheson, S.W., Dubois, M.S., Mellies, J.L., Wainwright, L.A., Batchelor, M., Frankel, G., Knutton, S., Kaper, J.B. Identification of CesT, a chaperone for the type III secretion of Tir in enteropathogenic *Escherichia coli*. *Molecular microbiology* 33, 1176–89 (1999).
109. Wilson, R.K., Shaw, R.K., Daniell, S., Knutton, S., Frankel, G. Role of EscF, a putative needle complex protein, in the type III protein translocation system of enteropathogenic *Escherichia coli*. *Cellular microbiology* 3, 753–62 (2001).
110. Donnenberg, M.S., Yu, J., Kaper, J.B. A second chromosomal gene necessary for intimate attachment of enteropathogenic *Escherichia coli* to epithelial cells. *Journal of bacteriology* 175, 4670–80 (1993).
111. Kenny, B., Lai, L.C., Finlay, B.B., Donnenberg, M.S. EspA, a protein secreted by enteropathogenic *Escherichia coli*, is required to induce signals in epithelial cells. *Molecular microbiology* 20, 313–23 (1996).
112. Lai, L.C., Wainwright, L.A., Stone, K.D., Donnenberg, M.S. A third secreted protein that is encoded by the enteropathogenic *Escherichia coli* pathogenicity island is required for transduction of signals and for attaching and effacing activities in host cells. *Infection and immunity* 65, 2211–7 (1997).
113. McNamara, B.P., Donnenberg, M.S. A novel proline-rich protein, EspF, is secreted from enteropathogenic *Escherichia coli* via the type III export pathway. *FEMS microbiology letters* 166, 71–8 (1998).

114. Elliott, S.J., Sperandio, V., Girón, J.A., Shin, S., Mellies, J.L., Wainwright, L., Hutcheson, S.W., McDaniel, T.K., Kaper, J.B. The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infection and immunity* 68, 6115–26 (2000).
115. Mellies, J.L., Elliott, S.J., Sperandio, V., Donnenberg, M.S., Kaper, J.B. The Per regulon of enteropathogenic *Escherichia coli*: identification of a regulatory cascade and a novel transcriptional activator, the locus of enterocyte effacement (LEE)-encoded regulator (Ler). *Molecular microbiology* 33, 296–306 (1999).
116. Sperandio, V., Mellies, J.L., Delahay, R.M., Frankel, G., Crawford, J.A., Nguyen, W., Kaper, J.B. Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler. *Molecular microbiology* 38, 781–93 (2000).
117. Haack, K., Robinson, C., Miller, K., Fowlkes, J., Mellies, J. Interaction of Ler at the LEE5 (tir) operon of enteropathogenic *Escherichia coli*. *Infection and immunity* 71, 384–92 (2003).
118. Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vázquez, A., Barba, J., Ibarra, J.A., O'Donnell, P., Metalnikov, P., Ashman, K., Lee, S., Goode, D., Pawson, T., Finlay, B.B. Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proceedings of the National Academy of Sciences of the United States of America* 101, 3597–602 (2004).
119. Lio, J., Syu, W. Identification of a negative regulator for the pathogenicity island of enterohemorrhagic *Escherichia coli* O157:H7. *Journal of biomedical science* 11, 855–63 (2004).
120. Tsai, N.P., Wu, Y.C., Chen, J.W., Wu, C.F., Tzeng, C.M., Syu, W.J. Multiple functions of l0036 in the regulation of the pathogenicity island of enterohaemorrhagic *Escherichia coli* O157:H7. *The Biochemical journal* 393, 591–9 (2006).
121. Mundy, R., Pickard, D., Wilson, R.K., Simmons, C.P., Dougan, G., Frankel, G. Identification of a novel type IV pilus gene cluster required for gastrointestinal colonization of *Citrobacter rodentium*. *Molecular microbiology* 48, 795–809 (2003).
122. Mundy, R., Petrovska, L., Smollett, K., Simpson, N., Wilson, R.K., Yu, J., Tu, X., Rosenshine, I., Clare, S., Dougan, G., Frankel, G. Identification of a novel *Citrobacter rodentium* type III secreted protein, EspI, and roles of this and other secreted proteins in infection. (2004).
123. Donnenberg, M.S., Tacket, C.O., James, S.P., Lososky, G., Nataro, J.P., Wasserman, S.S., Kaper, J.B., Levine, M.M. Role of the eaeA gene in experimental enteropathogenic *Escherichia coli* infection. *The Journal of clinical investigation* 92, 1412–7 (1993).
124. Tacket, C.O., Sztein, M.B., Lososky, G., Abe, A., Finlay, B.B., McNamara, B.P., Fantry, G.T., James, S.P., Nataro, J.P., Levine, M.M., Donnenberg, M.S. Role of EspB in experimental human enteropathogenic *Escherichia coli* infection. *Infection and immunity* 68, 3689–95 (2000).
125. Newman, J.V., Zabel, B.A., Jha, S.S., Schauer, D.B. *Citrobacter rodentium* espB is necessary for signal transduction and for infection of laboratory mice. *Infection and immunity* 67, 6019–25 (1999).
126. Elliott, S.J., Krejany, E.O., Mellies, J.L., Robins-Browne, R.M., Sasakwa, C., Kaper, J.B. EspG, a novel type III system-secreted protein from enteropathogenic *Escherichia coli* with similarities to VirA of *Shigella flexneri*. *Infection and immunity* 69, 4027–33 (2001).
127. Tu, X., Nisan, I., Yona, C., Hanski, E., Rosenshine, I. EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic *Escherichia coli*. *Molecular microbiology* 47, 595–606 (2003).

128. Crane, J.K., McNamara, B.P., Donnenberg, M.S. Role of EspF in host cell death induced by enteropathogenic *Escherichia coli*. *Cellular microbiology* 3, 197–211 (2001).
129. McNamara, B.P., Koutsouris, A., O'Connell, C.B., Nougayréde, J.P., Donnenberg, M.S., Hecht, G. Translocated EspF protein from enteropathogenic *Escherichia coli* disrupts host intestinal barrier function. *The Journal of clinical investigation* 107, 621–9 (2001).
130. Kenny, B., Jepson, M. Targeting of an enteropathogenic *Escherichia coli* (EPEC) effector protein to host mitochondria. *Cellular microbiology* 2, 579–90 (2000).
131. Kenny, B., Ellis, S., Leard, A.D., Warawa, J., Mellor, H., Jepson, M.A. Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic *Escherichia coli* effector molecules. *Molecular microbiology* 44, 1095–1107 (2002).
132. Tan, A., Petty, N.K., Hocking, D., Bennett-Wood, V., Wakefield, M., Praszkier, J., Tauschek, M., Yang, J., Robins-Browne, R. Evolutionary adaptation of an AraC-like regulatory protein in *Citrobacter rodentium* and *Escherichia* species. *Infection and immunity* (2015). doi:10.1128/IAI.02697-14
133. Tauschek, M., Yang, J., Hocking, D., Azzopardi, K., Tan, A., Hart, E., Praszkier, J., Robins-Browne, R.M. Transcriptional analysis of the *grlRA* virulence operon from *Citrobacter rodentium*. *Journal of bacteriology* 192, 3722–34 (2010).
134. Petty, N.K., Feltwell, T., Pickard, D., Clare, S., Toribio, A.L., Fookes, M., Roberts, K., Monson, R., Nair, S., Kingsley, R.A., Bulgin, R., Wiles, S., Goulding, D., Keane, T., Corton, C., Lennard, N., Harris, D., Willey, D., Rance, R., Yu, L., Choudhary, J.S., Churcher, C., Quail, M.A., Parkhill, J., Frankel, G., Dougan, G., Salmond, G.P., Thomson, N.R. *Citrobacter rodentium* is an unstable pathogen showing evidence of significant genomic flux. *PLoS pathogens* 7, e1002018 (2011).
135. Petty, N.K., Bulgin, R., Crepin, V.F., Cerdeño-Tárraga, A.M., Schroeder, G.N., Quail, M.A., Lennard, N., Corton, C., Barron, A., Clark, L., Toribio, A.L., Parkhill, J., Dougan, G., Frankel, G., Thomson, N.R. The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. *Journal of bacteriology* 192, 525–38 (2010).
136. Gueguen, E., Wills, N., Atkins, J., Cascales, E. Transcriptional frameshifting rescues *Citrobacter rodentium* type VI secretion by the production of two length variants from the prematurely interrupted *tssM* gene. *PLoS genetics* 10, e1004869 (2014).
137. Petty, N.K., Toribio, A.L., Goulding, D., Foulds, I., Thomson, N., Dougan, G., Salmond, G.P. A generalized transducing phage for the murine pathogen *Citrobacter rodentium*. *Microbiology (Reading, England)* 153, 2984–8 (2007).
138. Wiles, S., Dougan, G., Frankel, G. Emergence of a ‘hyperinfectious’ bacterial state after passage of *Citrobacter rodentium* through the host gastrointestinal tract. *Cellular microbiology* 7, 1163–72 (2005).
139. Smith, A., Bhagwat, A. Hypervirulent-host-associated *Citrobacter rodentium* cells have poor acid tolerance. *Current microbiology* 66, 522–6 (2013).
140. Smith, A., Yan, X., Chen, C., Dawson, H., Bhagwat, A. Understanding the host-adapted state of *Citrobacter rodentium* by transcriptomic analysis. *Archives of microbiology* 198, 353–62 (2016).
141. Caballero-Flores, G., Croxen, M., Martínez-Santos, V., Finlay, B.B., Puente, J. Identification and regulation of a novel *Citrobacter rodentium* gut colonization fimbria (Gcf). *Journal of bacteriology* 197, 1478–91 (2015).
142. Wiles, S., Clare, S., Harker, J., Huett, A., Young, D., Dougan, G., Frankel, G. Organ-specificity, colonization and clearance dynamics in vivo following oral challenges with the murine pathogen *Citrobacter rodentium*. *Cellular Microbiology* 7, 459–459 (2005).

143. Knapp, S., Hacker, J., Jarchau, T., Goebel, W. Large, unstable inserts in the chromosome affect virulence properties of uropathogenic *Escherichia coli* O6 strain 536. *Journal of bacteriology* 168, 22–30 (1986).
144. Rodríguez-Beltrán, J., Tourret, J., Tenaillon, O., López, E., Bourdelier, E., Costas, C., Matic, I., Denamur, E., Blázquez, J. High recombinant frequency in extra-intestinal pathogenic *Escherichia coli* strains. *Molecular biology and evolution* (2015). doi:10.1093/molbev/msv072
145. Coico, R. Gram staining. *Current protocols in microbiology Appendix 3, Appendix 3C* (2005).
146. Rasconi, S., Jobard, M., Jouve, L., Sime-Ngando, T. Use of calcofluor white for detection, identification, and quantification of phytoplanktonic fungal parasites. *Applied and environmental microbiology* 75, 2545–53 (2009).
147. Rowe, M., Withers, H., Swift, S. Uropathogenic *Escherichia coli* forms biofilm aggregates under iron restriction that disperse upon the supply of iron. *FEMS microbiology letters* 307, 102–9 (2010).
148. Foster, P. Methods for determining spontaneous mutation rates. *Methods in enzymology* 409, 195–213 (2006).
149. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* 166, 557–580 (1983).
150. Hanahan, D., Jesse, J., Bloom, F.R. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods in enzymology* 204, 63–113 (1991).
151. Field, D., Tiwari, B., Booth, T., Houten, S., Swan, D., Bertrand, N., Thurston, M. Open software for biologists: from famine to feast. *Nature Biotechnology* 24 (2006).
152. Chevreux, B., Wetter, T., Suhai, S. Genome sequence assembly using trace signals and additional sequence information. *German conference on bioinformatics* (1999).
153. Read, H.M., Mills, G., Johnson, S., Tsai, P., Dalton, J., Barquist, L., Print, C.G. Patrick, W.M., Wiles, S.. The *in vitro* and *in vivo* effects of constitutive light expression on a bioluminescent strain of the mouse enteropathogen *Citrobacter rodentium*. (2016).
154. Vencl, F.V. Allometry and proximate mechanisms of sexual selection in photinus fireflies, and some other beetles. *Integrative and comparative biology* 44, 242–9 (2004).
155. Meyer-Rochow, V. B. Glowworms: a review of Arachnocampa spp. and kin. *Luminescence* 22, 251–65 (2007).
156. Jones, B. W. & Nishiguchi, M. K. Counterillumination in the Hawaiian bobtail squid, *Euprymna scolopes* Berry (Mollusca: Cephalopoda). *Mar. Biol.* 144, 1151–1155 (2004).
157. De Wet, J. R., Wood, K. V., Helinski, D. R. & DeLuca, M. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82, 7870–3 (1985).
158. Engebrecht, J., Nealon, K. & Silverman, M. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. *Cell* 32, 773–81 (1983).
159. Szittner, R. & Meighen, E. Nucleotide sequence, expression, and properties of luciferase coded by *lux* genes from a terrestrial bacterium. *J. Biol. Chem.* 265, 16581–7 (1990).
160. Andreu, N., Zelmer, A. & Wiles, S. Noninvasive biophotonic imaging for studies of infectious disease. *FEMS Microbiol. Rev.* 35, 360–94 (2011).

161. Andreu, N., Zelmer, A., Sampson, S. L., Ikeh, M., Bancroft, G. J., Schaible, U. E., Wiles, S. & Robertson B. D. Rapid in vivo assessment of drug efficacy against *Mycobacterium tuberculosis* using an improved firefly luciferase. *J. Antimicrob. Chemother.* 68, 2118–27 (2013).
162. Andreu, N., Fletcher, T., Krishnan, N., Wiles, S. & Robertson, B. D. Rapid measurement of antituberculosis drug activity in vitro and in macrophages using bioluminescence. *J. Antimicrob. Chemother.* 67, 404–14 (2012).
163. Hastings & Presswood. Bacterial luciferase: FMNH₂-aldehyde oxidase. *Meth. Enzymol.* 53, 558–70 (1978).
164. Collins, J. W., Keeney, K. M., Crepin, V. F., Rathinam, V. A., Fitzgerald, K. A., Finlay, B. B. & Frankel, G. *Citrobacter rodentium*: infection, inflammation and the microbiota. *Nat. Rev. Microbiol.* 12, 612–23 (2014).
165. Winson, M. K., Swift, S., Hill, P. J., Sims, C. M., Griesmayr, G., Bycroft, B. W., Williams, P. & Stewart, G.S. Engineering the *luxCDABE* genes from *Photobacterium luminescens* to provide a bioluminescent reporter for constitutive and promoter probe plasmids and mini-Tn5 constructs. *FEMS Microbiol. Lett.* 163, 193–202 (1998).
166. Bochner, B. R., Gadzinski, P. & Panomitros, E. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Res.* 11, 1246–55 (2001).
167. Davis, B. D. The isolation of biochemically deficient mutants of bacteria by means of penicillin. *Proc. Natl. Acad. Sci. U.S.A.* 35, 1–10 (1949).
168. Cox, M. P., Peterson, D. A. & Biggs, P. J. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11, 485 (2010).
169. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–95 (2010).
170. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–9 (2009).
171. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* 1151, 165–88 (2014).
172. Hernandez, D., François, P., Farinelli, L., Osterås, M. & Schrenzel, J. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res.* 18, 802–9 (2008).
173. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–9 (2012).
174. Reuter, S., Connor, T. R., Barquist, L., Walker, D., Feltwell, T., Harris, S. R., Fookes, M., Hall, M. E., Petty, N. K., Fuchs, T. M., Corander, J., Dufour, M., Ringwood, T., Savin, C., Bouchier, C., Martin, L., Miettinen, M., Shubin, M., Riehm, J. M., Laukkonen-Ninios, R., Sihvonen, L. M., Siitonens, A., Skurnik, M., Falcão, J. P., Fukushima, H., Scholz, H. C., Prentice, M. B., Wren, B. W., Parkhill, J., Carniel, E., Achtman, M., McNally, A. & Thomson, N.R. Parallel independent evolution of pathogenicity within the genus *Yersinia*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6768–73 (2014).
175. Homann, O. R., Cai, H., Becker, J. M. & Lindquist, S. L. Harnessing natural diversity to probe metabolic pathways. *PLOS Genet.* 1, e80 (2005).

176. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, Article3 (2004).
177. Galardini, M., Mengoni, A., Biondi, E. G., Semeraro, R., Florio, A., Bazzicalupo, M., Benedetti, A. & Mocali, S. DuctApe: A suite for the analysis and correlation of genomic and OmniLogTM Phenotype Microarray data. *Genomics* 103, 1-10 (2014).
178. Freter, R., O'Brien, P. C. & Macsai, M. S. Role of chemotaxis in the association of motile bacteria with intestinal mucosa: *in vivo* studies. *Infect. Immun.* 34, 234–40 (1981).
179. Taylor, R. K., Miller, V. L., Furlong, D. B. & Mekalanos, J. J. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2833–7 (1987).
180. Sun, Y., Connor, M. G., Pennington, J. M. & Lawrenz, M. B. Development of bioluminescent bioreporters for *in vitro* and *in vivo* tracking of *Yersinia pestis*. *PLOS One* 7, e47123 (2012).
182. Steinhuber, A., Landmann, R., Goerke, C., Wolz, C. & Flückiger, U. Bioluminescence imaging to study the promoter activity of *hla* of *Staphylococcus aureus* *in vitro* and *in vivo*. *Int. J. Med. Microbiol.* 298, 599–605 (2008).
183. Kassem, I. I., Splitter, G. A., Miller, S. & Rajashekara, G. Let there be light! Bioluminescent imaging to study bacterial pathogenesis in live animals and plants. *Adv. Biochem. Eng. Biotechnol.* 154, 119–45 (2016).
184. Sanz, P., Teel, L. D., Alem, F., Carvalho, H. M., Darnell, S. C. & O'Brien, A. D. Detection of *Bacillus anthracis* spore germination *in vivo* by bioluminescence imaging. *Infect. Immun.* 76, 1036–47 (2008).
185. Alam, F. M., Bateman, C., Turner, C. E., Wiles, S. & Sriskandan, S. Non-invasive monitoring of *Streptococcus pyogenes* vaccine efficacy using biophotonic imaging. *PLOS One* 8, e82123 (2013).
186. Bergmann, S., Rohde, M., Schughart, K. & Lengeling, A. The bioluminescent *Listeria monocytogenes* strain Xen32 is defective in flagella expression and highly attenuated in orally infected BALB/cJ mice. *Gut Pathog.* 5, 19 (2013).
187. Falls, K., Williams, A., Bryksin, A. & Matsumura, I. *Escherichia coli* deletion mutants illuminate trade-offs between growth rate and flux through a foreign anabolic pathway. *PLOS One* 9, e88159 (2014).
188. Soo, V. W., Hanson-Manful, P. & Patrick, W. M. Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1484–9 (2011).
189. Vogel, H., Altincicek, B., Glöckner, G. & Vilcinskas, A. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. *BMC Genomics* 12, 308 (2011).
190. Park, S., Kim, C. H., Jeong, W. H., Lee, J. H., Seo, S. J., Han, Y. S. & Lee, I. H. Effects of two hemolymph proteins on humoral defense reactions in the wax moth, *Galleria mellonella*. *Dev. Comp. Immunol.* 29, 43–51 (2005).
191. Bergin, D., Reeves, E., Renwick, J., Wientjes, F. & Kavanagh, K. Superoxide production in *Galleria mellonella* hemocytes: identification of proteins homologous to the NADPH oxidase complex of human neutrophils. *Infect. Immun.* 73, 4161–4170 (2005).
192. Wiles, S., Hanage, W., Frankel, G., Robertson, B. Modelling infectious disease - time to think outside the box? *Nature reviews. Microbiology* 4, 307–12 (2006).

193. Grave, K., Hansen, M.K., Kruse, H., Bangen, M., Kristoffersen, A.B. Prescription of antimicrobial drugs in Norwegian aquaculture with an emphasis on ‘new’ fish species. *Preventive Veterinary Medicine* 83 (2008).
194. Wise, R. Antimicrobial resistance: priorities for action. *The Journal of Antimicrobial Chemotherapy* 49, 585-586 (2002).
195. Kümmerer, K., Henninger, A. Promoting resistance by the emission of antibiotics from hospitals and households into effluent. *Clin Microbiol Infect* 9, 1203-1214 (2003).
196. Gaskins, H.R., Collier, C.T., Anderson, D.B. Antibiotics as growth promotants: mode of action. *Animal Biotechnology* 13, 29-42 (2002).
197. Cromwell, G.L. Why and how antibiotics are used in swine production. *Animal Biotechnology* 13, 7-27 (2002).
198. Chang, H., Hu, J., Asami, M., Kunikane, S. Simultaneous analysis of 16 sulfonamide and trimethoprim antibiotics in environmental waters by liquid chromatography–electrospray tandem mass spectrometry. *Journal of Chromatography A* 1190, 390-393 (2008).
199. Duong, H.A., Pham, N.H., Nguyen, H.T., Hoang, T.T., Pham, H.V., Pham, V.C., Berg, M., Giger, W., Alder, A.C. Occurrence, fate and antibiotic resistance of fluoroquinolone antibacterials in hospital wastewaters in Hanoi, Vietnam. *Chemosphere* 72, 968-973 (2008).
200. Peng, X., Tan, J., Tang, C., Yu, Y., Wang, Z. Multiresidue determination of fluoroquinolone, sulfonamide, trimethoprim, and chloramphenicol antibiotics in urban waters in China. *Environmental Toxicology and Chemistry* 27, 73-79 (2008).
201. Hernández, F., Sancho, J.V., Ibáñez, M., Guerrero, C. Antibiotic residue determination in environmental waters by LC-MS. *TrAC Trends in Analytical Chemistry* 26, 466-485 (2007).
202. Botitsi, E., Frosyni, C., Tsipi, D. Determination of pharmaceuticals from different therapeutic classes in wastewaters by liquid chromatography–electrospray ionization–tandem mass spectrometry. *Analytical and Bioanalytical Chemistry* 387, 1317-1327 (2007).
203. Batt, A.L., Aga, D.S. Simultaneous analysis of multiple classes of antibiotics by ion trap LC/MS/MS for assessing surface water and groundwater contamination. *Analytical Chemistry* 77, 2940-2947 (2005).
204. Goneau, L.W., Hannan, T.J., MacPhee, R.A., Schwartz, D.J., Macklaim, J.M., Gloor, G.B., Razvi, H., Reid, G., Hultgren, S.J., Burton, J.P. Subinhibitory antibiotic therapy alters recurrent urinary tract infection pathogenesis through modulation of bacterial virulence and host immunity. *mBio* 6, (2015).
205. Modi, S., Lee, H., Spina, C., Collins, J. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499, 219–222 (2013).
206. Hoggard, M., Biswas, K., Zoing, M., Mackenzie, B.W., Taylor, M.W., Douglas, R.G. Evidence of microbiota dysbiosis in chronic rhinosinusitis. *International Forum of Allergy & Rhinology* (2016).
207. Owens, R., Donskey, C., Gaynes, R., Loo, V., Muto, C. Antimicrobial-associated risk factors for *Clostridium difficile* infection. *Clinical Infectious Diseases* 46, S19–S31 (2008).
208. Luzzatto, L., Apirion, D., Schlessinger, D. Mechanism of action of streptomycin in *E. coli*: interruption of the ribosome cycle at the initiation of protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 60, 873–80 (1968).
209. Wehrli, W. Rifampin: mechanisms of action and resistance. *Reviews of infectious diseases* 5 Suppl 3, S407–11 (1983).

210. Qi, W., Joshi, S., Weber, C.R., Wali, R.K., Roy, H.K., Savkovic, S.D. Polyethylene glycol diminishes pathological effects of *Citrobacter rodentium* infection by blocking bacterial attachment to the colonic epithelia. *Gut microbes* 2, 267–73 (2011).
211. Himathongkham, S., Riemann, H. Destruction of *Salmonella typhimurium*, *Escherichia coli* O157:H7 and *Listeria monocytogenes* in chicken manure by drying and/or gassing with ammonia. *FEMS microbiology letters* 171, 179–82 (1999).
212. Sinton, L., Braithwaite, R., Hall, C., Mackenzie, M. Survival of indicator and pathogenic bacteria in bovine feces on pasture. *Applied and environmental microbiology* 73, 7917–25 (2007).
213. Méhi, O., Bogos, B., Csörgo, B., Pál, C. Genomewide Screen for Modulators of Evolvability under Toxic Antibiotic Exposure. *Antimicrobial agents and chemotherapy* 57, 3453–6 (2013).
214. Criss, A.K., Bonney, K.M., Chang, R.A., Duffin, P.M., LeCuyer, B.E., Seifert, H.S. Mismatch correction modulates mutation frequency and pilus phase and antigenic variation in *Neisseria gonorrhoeae*. *Journal of bacteriology* 192, 316–25 (2010).
215. Giraud, A., Matic, I., Radman, M., Fons, M., Taddei, F. Mutator bacteria as a risk factor in treatment of infectious diseases. *Antimicrobial agents and chemotherapy* 46, 863–5 (2002).
216. Ozaki, M., Mizushima, S., Nomura, M. Identification and functional characterization of the protein controlled by the streptomycin-resistant locus in *E. coli*. *Nature* 222, 333–9 (1969).
217. Springer, B., Kidan, Y.G., Prammananan, T., Ellrott, K., Böttger, E.C., Sander, P. Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. *Antimicrobial agents and chemotherapy* 45, 2877–84 (2001).
218. Timms, A.R., Steingrimsdottir, H., Lehmann A.R., Bridges, B.A. Mutant sequences in the rpsL gene of *Escherichia coli* B/r: mechanistic implications for spontaneous and ultraviolet light mutagenesis. *Molecular & general genetics : MGG* 232, 89–96 (1992).
219. Mindlin, S.Z., Ilyina, T.S., Voeykova, T.A., Velkov, V.V. Genetical analysis of rifampicin resistant mutants of *E. coli* K 12. *Molecular & general genetics : MGG* 115, 115–21 (1972).
220. Rodríguez-Verdugo, A., Gaut, B., Tenaillon, O. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC evolutionary biology* 13, 50 (2013).
221. Xu, M., Zhou, Y., Goldstein, B., Jin, D. Cross-resistance of *Escherichia coli* RNA polymerases conferring rifampin resistance to different antibiotics. *Journal of bacteriology* 187, 2783–2792 (2005).
222. Moxon, E.R., Rainey, P.B., Nowak, M.A., Lenski, R.E. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current biology : CB* 4, 24–33 (1994).
223. Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M.E., Frapy, E., Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguénec, C., Lescat, M., Mangenot, S., Martinez-Jéhanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M.A., Pichon, C., Rouy, Z., Ruf, C.S., Schneider, D., Tourret, J., Vacherie, B., Vallenet, D., Médigue, C., Rocha, E.P., Denamur, E. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS genetics* 5, e1000344 (2009).
224. Tenaillon, O., Le Nagard, H., Godelle, B., Taddei, F. Mutators and sex in bacteria: conflict between adaptive strategies. *Proceedings of the National Academy of Sciences of the United States of America* 97, 10465–70 (2000).
225. Deatherage, D.E., Traverse, C.C., Wolf, L.N., Barrick, J.E. Detecting rare structural variation in evolving microbial populations from new sequence junctions using breseq. *Front Genet* 5, 468 (2014).

226. Kimura, M. The neutral theory of molecular evolution. *Scientific American* 241, 98–100, 102, 108 passim (1979).
227. Miyata, T., Yasunaga, T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *Journal of Molecular Evolution* 16, 23–36 (1980).
228. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. BLAST+: architecture and applications. *BMC Bioinformatics* (2009).
229. Madden, T.L., Tatusov, R.L., Zhang, J. Applications of network BLAST server. *Methods in Enzymology* (1996).
230. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410 (1990).
231. The UniProt Consortium. UniProt: a hub for protein information. *Nucleic acids research* gku989 (2014). doi:10.1093/nar/gku989
232. Pujic, P., Dervyn, R., Sorokin, A., Ehrlich, S.D. The kdgRKAT operon of *Bacillus subtilis*: detection of the transcript and regulation by the kdgR and ccpA genes. *Microbiology (Reading, England)* 144 (Pt 11), 3111–8 (1998).
233. Allen, C., Reverchon, S., Robert-Baudouy, J. Nucleotide sequence of the *Erwinia chrysanthemi* gene encoding 2-keto-3-deoxygluconate permease. *Gene* 83, 233–41 (1989).
234. Gavín, R., Rabaan, A.A., Merino, S., Tomás, J.M., Gryllos, I., Shaw, J.G. Lateral flagella of *Aeromonas* species are essential for epithelial cell adherence and biofilm formation. *Molecular microbiology* 43, 383–97 (2002).
235. Gumaa, K.A., McLean, P. The pentose phosphate pathway of glucose metabolism. Enzyme profiles and transient and steady-state content of intermediates of alternative pathways of glucose metabolism in Krebs ascites cells. *The Biochemical Journal* 155, 1009–1029 (1969).
236. Schultz, J.E., Matin, A. Molecular and functional characterization of a carbon starvation gene of *Escherichia coli*. *Journal of molecular biology* 218, 129–40 (1991).
237. Spencer, M.E., Guest, J.R. Transcription analysis of the sucAB, aceEF and lpd genes of *Escherichia coli*. *Molecular & general genetics : MGG* 200, 145–54 (1985).
238. Reeves, R.E., Sols, A. Regulation of *Escherichia coli* phosphofructokinase in situ. *Biochemical and biophysical research communications* 50, 459–66 (1973).
239. Santander, P.J., Roessner, C.A., Stolowich, N.J., Holderman, M.T., Scott, A.I. How corrinoids are synthesized without oxygen: nature's first pathway to vitamin B12. *Chemistry & biology* 4, 659–66 (1997).
240. Nakayama, K., Maeda, Y., Jigami, Y. Interaction of GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase with GDP-mannose-4,6-dehydratase stabilizes the enzyme activity for formation of GDP-fucose from GDP-mannose. *Glycobiology* 13, 673–80 (2003).
241. Stojiljkovic, I., Bäumler, A.J., Heffron, F. Ethanolamine utilization in *Salmonella typhimurium*: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster. *Journal of bacteriology* 177, 1357–66 (1995).
242. Klemm, P., Christiansen, G. The fimD gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. *Molecular & general genetics : MGG* 220, 334–8 (1990).

243. Orndorff, P.E., Falkow, S. Organization and expression of genes responsible for type 1 piliation in *Escherichia coli*. *Journal of bacteriology* 159, 736–44 (1984).
244. Mundy, R., Petrovska, L., Smollett, K., Simpson, N., Wilson, R.K., Yu, J., Tu, X., Rosenshine, I., Clare, S., Dougan, G., Frankel, G. Identification of a novel *Citrobacter rodentium* type III secreted protein, EspI, and roles of this and other secreted proteins in infection. *Infection and immunity* 72, 2288–302 (2004).
245. Gueguen, E., Cascales, E. Promoter swapping unveils the role of the *Citrobacter rodentium* CTS1 type VI secretion system in interbacterial competition. *Applied and environmental microbiology* 79, 32–8 (2013).
246. Grass, G., Wong, M., Rosen, B., Smith, R., Rensing, C. ZupT is a Zn(II) uptake system in *Escherichia coli*. *Journal of bacteriology* 184, 864–6 (2002).
247. Borukhov, S., Sagitov, V., Goldfarb, A. Transcript cleavage factors from *E. coli*. *Cell* 72, 459–66 (1993).
248. Lovett, S. Replication arrest-stimulated recombination: Dependence on the RecA paralog, RadA/Sms and translesion polymerase, DinB. *DNA repair* 5, 1421–7 (2006).
249. Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., Chan, A.P. Predicting the functional effect of amino acid substitutions and indels. *PLOS one* (2012). doi:10.1371/journal.pone.0046688
250. Choi, Y. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine* (2012).
251. Choi, Y., Chan, A.P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 15, 2745–2747 (2015).
252. Gish, W., States, D.J. Identification of protein coding regions by database similarity search. *Nature Genetics* 3, 266–272 (1993).
253. Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., Stockinger, H. ExPASy: SIB bioinformatics resource portal. (2012).
254. Herron, M., Doebeli, M. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. *PLoS biology* 11, e1001490 (2013).
255. Lang, G.I., Rice, D.P., Hickman, M.J., Sodergren, E., Weinstock, G.M., Botstein, D., Desai, M.M. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500, 571–4 (2013).
256. Barrick, J.E., Lenski, R.E. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harbor symposia on quantitative biology* 74, 119–29 (2009).
257. Kvitek, D., Sherlock, G. Whole genome, whole population sequencing reveals that loss of signaling networks is the major adaptive strategy in a constant environment. *PLoS genetics* 9, e1003972 (2013).
258. Wren, B. The Yersiniae — a model genus to study the rapid evolution of bacterial pathogens. *Nature Reviews Microbiology* 1, 55–64 (2003).

259. Kendall, M., Gruber, C., Parker, C., Sperandio, V. Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7. *mBio* 3, (2012).
260. Hershberg, R., Petrov, D. Evidence that mutation is universally biased towards AT in bacteria. *PLoS genetics* 6, e1001115 (2010).
261. Jackson, R., Vinatzer, B., Arnold, D., Dorus, S., Murillo, J. The influence of the accessory genome on bacterial pathogen evolution. *Mobile genetic elements* 1, 55–65 (2011).
262. Phillips, P. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. *Nature reviews. Genetics* 9, 855–67 (2008).
263. Howe, K., Karsi, A., Germon, P., Wills, R.W., Lawrence, M.L., Bailey, R.H. Development of stable reporter system cloning luxCDABE genes into chromosome of *Salmonella enterica* serotypes using Tn7 transposon. *BMC microbiology* 10, 197 (2010).
264. Maoz, A., Mayr, R., Bresolin, G., Neuhaus, K., Francis, K.P., Scherer, S. Sensitive *in situ* monitoring of a recombinant bioluminescent *Yersinia enterocolitica* reporter mutant in real time on Camembert cheese. *Applied and environmental microbiology* 68, 5737–40 (2002).
265. Riedel, C.U., Casey, P.G., Mulcahy, H., O'Gara, F., Gahan, C.G., Hill, C. Construction of p16Slux, a novel vector for improved bioluminescent labeling of gram-negative bacteria. *Applied and environmental microbiology* 73, 7092–5 (2007).
266. Flaherty, C., Begley, M., Hill, C. Investigation of the use of a cocktail of lux-tagged *Cronobacter* strains for monitoring growth in infant milk formulae. *Journal of food protection* 76, 1359–65 (2013).
267. Ferenci, T. Trade-off Mechanisms Shaping the Diversity of Bacteria. *Trends in microbiology* (2015). doi:10.1016/j.tim.2015.11.009
268. Lebeis, S.L., Bommarius, B., Parkos, C.A., Sherman, M.A., Kalman, D. TLR signaling mediated by MyD88 is required for a protective innate immune response by neutrophils to *Citrobacter rodentium*. *Journal of Immunology* 179, 566-577 (2007).
269. Gibson, D.L., Ma, C., Bergstrom, K.S., Huang, J.T., Man, C., Vallance, B.A. MyD88 signalling plays a critical role in host defence by controlling pathogen burden and promoting epithelial cell homeostasis during *Citrobacter rodentium*-induced colitis. *Cell Microbiol* 10, 618-631 (2008).
270. Cross, A.R., Segal, A.W. The NADPH oxidase of professional phagocytes—prototype of the NOX electron transport chain systems. *Biochimica et biophysica acta* (2004).
271. Fournier, B.M., Parkos, C.A. The role of neutrophils during intestinal inflammation. *Mucosal Immunology* 5, 354-366 (2012).
272. Winter, S.E., Thiennimitr, P., Winter, M.G., Butler, B.P., Huseby, D.L., Crawford, R.W., Russell, J.M., Bevins, C.L., Adams, L.G., Tsolis, R.M., Roth, J.R., Bäumler, A.J. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature* 467, 426–9 (2010).
273. Vong, L., Pinnell, L.J., Määttänen, P., Yeung, C.W., Lurz, E., Sherman, P.M. Selective enrichment of commensal gut bacteria protects against *Citrobacter rodentium*-induced colitis. *American journal of physiology. Gastrointestinal and liver physiology* 309, G181–92 (2015).
274. Jones, S., Knight, K. *Bacillus subtilis*-mediated protection from *Citrobacter rodentium*-associated enteric disease requires espH and functional flagella. *Infection and immunity* 80, 710–9 (2012).
275. Fanning, S., Hall, L.J., Cronin, M., Zomer, A., MacSharry, J., Goulding, D., Motherway, M.O., Shanahan, F., Nally, K., Dougan, G., van Sinderen, D. Bifidobacterial surface-exopolysaccharide

- facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences of the United States of America* 109, 2108–13 (2012).
276. Curtis, M.M., Hu, Z., Klimko, C., Narayanan, S., Deberardinis, R., Sperandio, V. The gut commensal *Bacteroides thetaiotomicron* exacerbates enteric infection through modification of the metabolic landscape. *Cell host & microbe* 16, 759–69 (2014).
277. Sperandio, V., Li, C., Kaper, J. Quorum-sensing *Escherichia coli* regulator A: a regulator of the LysR family involved in the regulation of the locus of enterocyte effacement pathogenicity island in enterohemorrhagic *E. coli*. *Infection and immunity* 70, 3085–93 (2002).
278. Schouler, C., Koffmann, F., Amory, C., Leroy-Sétrin, S. & Moulin-Schouleur, M. Genomic subtraction for the identification of putative new virulence factors of an avian pathogenic *Escherichia coli* strain of O2 serogroup. *Microbiology (Reading, England)* 150, 2973–84 (2004).
279. Gaurivaud, P., Danet, J.L., Laigret, F., Garnier, M., Bové, J.M. Fructose utilization and phytopathogenicity of *Spiroplasma citri*. *Molecular plant-microbe interactions : MPMI* 13, 1145–55 (2000).
280. Järvinen, N., Mäki, M., Räbinä, J., Roos, C., Mattila, P., Renkonen, R. Cloning and expression of *Helicobacter pylori* GDP-l-fucose synthesizing enzymes (GMD and GMER) in *Saccharomyces cerevisiae*. *European journal of biochemistry / FEBS* 268, 6458–64 (2001).
281. Woods, W.G., Dyall-Smith, M.L. Construction and analysis of a recombination-deficient (*radA*) mutant of *Haloferax volcanii*. *Molecular microbiology* 23, 791–7 (1997).
282. Wiegand, I., Marr, A.K., Breidenstein, E.B., Schurek, K.N., Taylor, P., Hancock, R.E. Mutator genes giving rise to decreased antibiotic susceptibility in *Pseudomonas aeruginosa*. *Antimicrobial agents and chemotherapy* 52, 3810–3 (2008).
283. Brunet, Y., Espinosa, L., Harchouni, S., Mignot, T., Cascales, E. Imaging type VI secretion-mediated bacterial killing. *Cell reports* 3, 36–41 (2013).
284. Hood, R.D., Singh, P., Hsu, F., Güvener, T., Carl, M.A., Trinidad, R.R., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L., Li, M., Schwarz, S., Wang, W.Y., Merz, A.J., Goodlett, D.R., Mougous, J.D. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell host & microbe* 7, 25–37 (2010).
285. MacIntyre, D., Miyata, S., Kitaoka, M., Pukatzki, S. The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proceedings of the National Academy of Sciences of the United States of America* 107, 19520–4 (2010).
286. Murdoch, S.L., Trunk, K., English, G., Fritsch, M.J., Pourkarimi, E., Coulthurst, S.J. The opportunistic pathogen *Serratia marcescens* utilizes type VI secretion to target bacterial competitors. *Journal of bacteriology* 193, 6057–69 (2011).
287. Alcoforado Diniz, J., Coulthurst, S.J. Intraspecies competition in *Serratia marcescens* is mediated by type VI-secreted Rhs effectors and a conserved effector-associated accessory protein. *Journal of Bacteriology* 197, 2350–2360 (2015).
288. Godefroid, M. Characterization of the clumping phenotype displayed by a Quorum sensing deficient strain in *Brucella melitensis* 16M and research of genes involved in its exopolysaccharide biosynthesis. FUNDP (2010).
289. Uzureau, S., Godefroid, M., Deschamps, C., Lemaire, J., De Bolle, X., Letesson, J-J. Mutations of the quorum sensing-dependent regulator VjbR lead to drastic surface modifications in *Brucella melitensis*. *Journal of bacteriology* 189, 6035–47 (2007).

290. Heighway, J., Santibanez-Koref, M.F. Unexpected effects on bacterial phenotype induced by expression of a tumour-amplified human sequence. *Nucleic acids research* 17, 6893–901 (1989).
291. Kachlany, S.C., Planet, P.J., Bhattacharjee, M.K., Kollia, E., DeSalle, R., Fine, D.H., Figurski, D.H. Nonspecific adherence by *Actinobacillus actinomycetemcomitans* requires genes widespread in bacteria and archaea. *Journal of bacteriology* 182, 6169–76 (2000).
292. Bernard, C., Bordi, C., Termine, E., Filloux, A., de Bentzmann, S. Organization and PprB-dependent control of the *Pseudomonas aeruginosa* tad Locus, involved in Flp pilus biology. *Journal of bacteriology* 191, 1961–73 (2009).
293. Sherlock, O., Schembri, M., Reisner, A., Klemm, P. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *Journal of bacteriology* 186, 8058–65 (2004).
294. Dennis, A., Kudo, T., Kruidenier, L., Girard, F., Crepin, V.F., MacDonald, T.T., Frankel, G., Wiles, S. The p50 subunit of NF- κ B is critical for *in vivo* clearance of the noninvasive enteric pathogen *Citrobacter rodentium*. *Infect. Immun.* 76, 4978–88 (2008).
295. Vallance, B., Deng, W., Jacobson, K., Finlay B.B. Host susceptibility to the attaching and effacing bacterial pathogen *Citrobacter rodentium*. *Infection and immunity* 71, 3443–53 (2003).