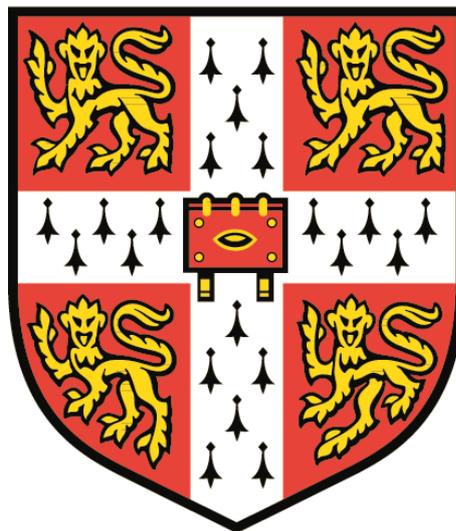


**A functional genomic-based study of the
streptomycin mouse model of human
Salmonella Typhimurium gastroenteritis**

Jennifer Louise Hill



University of Cambridge

Darwin College

August 2015

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and is unlike any work I have previously submitted for any other qualification. Work performed in collaboration is declared below and/or specified in the materials and methods section. This thesis does not exceed the word limit of 60,000 words (excluding bibliography, figures and appendices) required by the University of Cambridge School of Biological Sciences.

Cordelia Brandt, Katherine Harcourt and Leanne Kane performed cervical dislocation and cardiac puncture, and assisted with the collection of tissue from mice. George Notley weighed mice and monitored their wellbeing. David Goulding performed three-dimensional confocal imaging and took conventional confocal microscope images with my assistance. RNAseq library preparation, RNAseq and DNA sequencing were performed by the Wellcome Trust Sanger Institute (WTSI) sequencing core facility. Alignment of sequence reads with the mouse reference genome and generation of read counts was performed by the WTSI pathogen informatics facility as part of the RNAseq transcript mapping pipeline. Dr Lu Yu assisted in protocol development for extraction of protein from mouse caecum for mass spectrometry (MS), and performed MS and database searching. Dr James Wright performed analysis on MS data to produce fold changes in protein abundance and T-test p-values. Prof. Mark Arends (University of Edinburgh Division of Pathology) performed pathological scoring of mouse tissue. Dr Maria Duque performed qPCR analysis of tissues from naïve *IL22ra1*^{tm1a/tm1a} mice presented in Figure 6.4B.

Jennifer Hill

August 2015

Acknowledgements

I would like to thank my supervisor Prof. Gordon Dougan for giving me the opportunity to carry out this work, and for his guidance and support throughout. For their invaluable suggestions and direction I wish to thank the members of my thesis committee; Prof. Paul Kellam and Prof. Arthur Kaser; also Prof. Cal MacLennan for many very helpful conversations about complement. I am highly grateful to Dr Simon Clare for sharing his expertise in the area of mouse experiments and Dr Jyoti Choudhary for her thoughts and advice in the area of proteomics.

I would like thank the many people who have given their valuable time to teach me techniques both in the lab and at the computer, to offer advice and suggestions, and to assist me with experiments. There is of course a very large number and I hope to have remembered most here - thank you to Dr Anneliese Speak, Emma Cambridge, Yvette Hooks, Mark Stares, David Harris, Dr Nitin Kumar, Kaur Alasoo, Dr Christine Hale, Dr Elizabeth Klemm, Jessica Forbester, Dr Tu Anh Pham, Dr Lynda Mottram, Dr Mercedes Pardo Calvo, Dr Robert Kingsley and Dr Maria Duque. Also many thanks to those mentioned in the declaration section, in particular the MGP team for their help with mouse work on countless occasions; your skill and dexterity I aspire to achieve myself one day. Also a special thanks to Dr Lu Yu for advice and support with all things concerning MS combined with an excellent sense of humour.

I wish to extend my heartfelt thanks to all those who have made my time at the WTSI an enjoyable one, I will miss you. Del, your everlasting supply of chocolate in the office was on occasion vital to my survival. Thanks to those of you who helped me to plate out organs into the evening, and all who joined me for a laugh, a gossip and a cup of tea.

I thank my mum for her dedicated proofreading. I apologise sincerely for the fact you now know far more about *Salmonella* infection than you ever hoped to! To Tom and my family, I thank you for all your encouragement and reassurance throughout my PhD, without this I would not have enjoyed the past four years as I have; perhaps I'd never have finished this at all!

Finally I acknowledge the Wellcome Trust whose funding made this work possible.

Abstract

Antibiotic treatment abolishes resistance to invading microbes conferred by the natural murine microflora, creating an opportunity for *Salmonella* Typhimurium to colonise the gut. Pathological changes occurring in intestinal tissue during infection in mice mirror aspects of the inflammatory effects of *S. Typhimurium* upon the human intestinal mucosa. The streptomycin mouse model has emerged as a valuable tool to investigate both the host response to *Salmonella* in an intestinal setting, and bacterial virulence factors important for intestinal colonisation.

The Wellcome Trust Sanger Institute has established a phenotypic screening platform using novel mutant mice that incorporates a pathogen challenge component. This screen includes a systemic but not an oral *Salmonella* challenge. In this thesis I explore the potential of the murine *Salmonella* oral streptomycin treatment model as a secondary phenotyping component of such a screen. Using a combination of functional genomic approaches including RNAseq and proteomics I catalogue molecular changes which occur in caecal tissue during *S. Typhimurium* infection. Pathway analysis was used to aid interpretation of these large datasets and gain mechanistic insight into aspects of the host response. I found upregulated genes overrepresented in numerous immune-related pathways whereas downregulated genes were overrepresented in metabolic pathways; indicating infection leads to extensive disruption of local host metabolism.

Significantly overrepresented during infection at both the level of RNA and protein, the complement pathway was selected for further investigation in light of limited understanding of its role in mucosal infection. By Western blotting I demonstrated proteolytic activation of the complement protein C3 in intestinal tissue and using immunofluorescence staining showed patterns of C3 localisation in the mucosa. Using mutant mice, I identified genes with potential involvement in susceptibility to oral infection with *S. Typhimurium* and applied functional genomic approaches to describe the roles of these genes. In summary, this work explores the combination of high throughput approaches for identification of key signatures of infection with hypothesis-driven experiments in a model of *Salmonella* gastroenteritis, aiming to advance our understanding of host factors involved in the immune response to gastrointestinal infection.

Abbreviations

A/E	Attaching & effacing
APR	Acute phase response
BCR	B cell receptor
BMDM	Bone marrow-derived macrophage
cDNA	Complementary DNA
CFU	Colony forming units
CSA	Common structural antigens
DDA	Data-dependent acquisition
DE	Differentially expressed
DIA	Data-independent acquisition
DSS	Dextran sodium sulphate
EPEC	Enteropathogenic <i>Escherichia coli</i>
ES cell	Embryonic stem cell
EHEC	Enterohaemorrhagic <i>Escherichia coli</i>
FAE	Follicle-associated epithelium
FCS	Foetal calf serum
FDR	False discovery rate
GALT	Gut-associated lymphoid tissue
GEMS	Global Enteric Multicentre Study
GO	Gene ontology
GPCR	G-protein coupled receptor
GWAS	Genome-wide association study
HPLC	High pressure liquid chromatography
IBD	Inflammatory bowel disease
IEL	Intraepithelial lymphocyte
IKMC	International Knockout Mouse Consortium
ILC	Innate lymphoid cell
ILF	Innate lymphoid follicle

iNOS	Inducible nitric oxide synthase
iNTS	Invasive non-typhoidal <i>Salmonella</i>
IP	Intraperitoneal
LEE	Locus of enterocyte effacement
LPS	Lipopolysaccharide
LB	Luria Bertani
M cell	Microfold cell
MAC	Membrane attack complex
MBL	Mannose-binding lectin
MGP	Mouse genetics project
miRNA	microRNA
MLEE	Multi-locus enzyme electrophoresis
mLN	Mesenteric lymph node
MLST	Multi-locus sequence typing
mRNA	messengerRNA
MS	Mass spectrometry
NLR	Nod-like receptor
NTS	Non-typhoidal <i>Salmonella</i>
ORA	Overrepresentation analysis
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PI	Post-infection
PMN	Polymorphonuclear leukocyte
PRR	Pattern recognition receptor
qPCR	Quantitative polymerase chain reaction
RNP	Ribonucleoprotein

ROS	Reactive oxygen species
rRNA	Ribosomal RNA
SCV	<i>Salmonella</i> -containing vacuole
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sIgA	Secretory immunoglobulin A
SILT	Solitary isolated lymphoid tissue
SNP	Single nucleotide polymorphism
SPF	Specific pathogen free
SPI	Salmonella pathogenicity island
ST	Sequence type
TCR	T cell receptor
TLR	Toll-like receptor
T3SS	Type 3 secretion system
WGS	Whole genome sequencing
WHO	World Health Organisation
WTSI	Wellcome Trust Sanger Institute

Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
Abbreviations	iv
Contents	vii
List of figures	xiv
List of tables	xvii

1 Introduction.....	1
1.1 The global burden of infectious disease	1
1.1.1 Incidence, morbidity and mortality of diarrhoeal disease	2
1.2 Biology of the gastrointestinal tract	4
1.2.1 The gastrointestinal mucosa	4
1.2.1.1 The epithelium.....	4
1.2.1.2 The lamina propria and muscularis mucosae	6
1.2.2 The gut-associated immune system	7
1.2.2.1 mLN and secondary lymphoid tissue	7
1.2.2.2 Tertiary lymphoid tissue.....	8
1.2.2.3 Individual immune cells	8
1.2.3 The intestinal microbiota	8
1.2.3.1 Colonisation resistance.....	9
1.2.3.2 The microbiota in intestinal development and maintenance	10
1.2.3.3 Wider impact of the microbiota	11
1.2.3.4 The microbiota in diarrhoeal disease	12

1.3	<i>Salmonella</i>	13
1.3.1	Classification and phylogeny.....	13
1.3.2	Global distribution of <i>Salmonella</i> serotypes	15
1.3.3	Typhoidal and non-typhoidal <i>Salmonella</i> (NTS)	15
1.3.3.1	Clinical features and burden of typhoid fever	15
1.3.3.2	Clinical features and burden of NTS disease	16
1.3.3.3	Invasive non-typhoidal <i>Salmonella</i> (iNTS).....	16
1.3.4	<i>Salmonella</i> transmission	17
1.3.5	Molecular phylogeny of <i>S. enterica</i>	18
1.3.5.1	Serological and DNA-based classification.....	18
1.3.5.2	The genome of <i>S. enterica</i>	21
1.3.5.2.1	The genome of <i>S. Typhimurium</i>	21
1.3.5.2.2	The genome of <i>S. Typhi</i>	21
1.3.6	Virulence mechanisms of <i>S. Typhimurium</i>	23
1.3.6.1	Type 3 secretion systems (T3SS).....	24
1.3.6.2	Nutrition and energy sources.....	26
1.3.6.3	Resistance to host antimicrobial peptides	27
1.3.7	Animal models of human <i>S. Typhimurium</i> infection	28
1.3.7.1	Bovine models.....	30
1.3.7.2	The streptomycin mouse model	30
1.4	Murine models of intestinal inflammation.....	33
1.4.1	<i>Citrobacter rodentium</i>	34
1.4.2	Dextran sodium sulphate (DSS)-induced colitis	36
1.5	Immune response to <i>S. Typhimurium</i> infection in the intestinal mucosa.....	37
1.5.1	Innate immune surveillance and signalling	37
1.5.2	Three major innate effector responses to <i>S. Typhimurium</i>	39
1.5.2.1	Macrophage activation	39
1.5.2.2	Neutrophil recruitment	39
1.5.2.3	Antimicrobial peptides	40
1.5.3	The adaptive immune response	41
1.5.4	Genetic susceptibility to intestinal inflammation in humans.....	43
1.5.4.1	Infection	43
1.5.4.2	Inflammatory bowel disease.....	44

1.6	Aims of the thesis	44
2	Materials and methods	45
2.1	Materials	45
2.1.1	Bacterial strains	45
2.1.2	Oligonucleotides.....	45
2.1.2.1	Sequencing of 16S rRNA genes.....	45
2.1.2.2	Primers used in quantitative polymerase chain reaction (qPCR).....	45
2.1.3	Antibodies.....	46
2.1.4	Mice.....	46
2.2	Methods	47
2.2.1	<i>S. Typhimurium</i> culture and preparation of inoculums.....	47
2.2.2	<i>In vivo</i> experiments and tissue harvesting.....	47
2.2.2.1	Streptomycin pre-treatment and infection with <i>S. Typhimurium</i>	47
2.2.2.2	Enumeration of <i>S. Typhimurium</i> counts in tissue.....	47
2.2.2.3	Harvesting of tissue for RNAseq	48
2.2.2.4	Harvesting of intestinal content and tissue for microbiota analysis.....	48
2.2.2.5	Harvesting and culture of peritoneal macrophages	48
2.2.2.6	Harvesting and culture of bone marrow-derived macrophages (BMDM) ..	49
2.2.3	Peripheral blood leukocyte analysis	49
2.2.4	Microscopic analysis	50
2.2.4.1	Histopathological analysis.....	50
2.2.4.2	Immunofluorescence staining	51
2.2.4.3	Three-dimensional confocal imaging.....	51
2.2.5	RNA methods	52
2.2.5.1	RNA extraction	52
2.2.5.2	Reverse transcription and qPCR	52
2.2.5.3	RNAseq.....	52
2.2.5.3.1	Library preparation and sequencing.....	52
2.2.5.3.2	RNAseq data analysis.....	53
2.2.6	16S rRNA gene sequencing for microbiota analysis.....	53
2.2.6.1	DNA extraction, library preparation and sequencing.....	53
2.2.6.2	16S rRNA sequencing data analysis	54

2.2.7	Proteins	54
2.2.7.1	Analysis of tissue extracts by MS	54
2.2.7.1.1	Extraction of protein from caecal tissue.....	54
2.2.7.1.2	Protein concentration determination	54
2.2.7.1.3	Preparation for MS	54
2.2.7.1.4	MS analysis	55
2.2.7.1.5	MS data analysis.....	56
2.2.7.2	Western blotting	57
2.2.7.2.1	Preparation of intestinal content and faecal samples	57
2.2.7.2.2	Preparation of tissue samples	57
2.2.7.2.3	Preparation of plasma.....	57
2.2.7.2.4	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) 57	
2.2.7.2.5	Blotting and protein visualisation	57
2.2.8	Statistical tests	58

3 Pathological and transcriptional changes in the streptomycin mouse model of human

S. Typhimurium gastroenteritis	59	
3.1	Introduction.....	59
3.1.1	Characterisation of the streptomycin mouse model	59
3.1.2	Transcriptional profiling.....	60
3.1.2.1	Pathway analysis	61
3.2	Aims of the work described in this chapter	62
3.3	Results.....	62
3.3.1	Infection with <i>S. Typhimurium</i> results in weight loss and colonisation of gastrointestinal and systemic organs	62
3.3.2	<i>S. Typhimurium</i> infection results in major changes in caecal morphology	65
3.3.3	Severe inflammation and leukocyte infiltration in infected caecal tissue	65
3.3.4	Changes in the indigenous microflora in response to streptomycin and <i>S. Typhimurium</i> infection	68
3.3.4.1	Microbiota samples cluster according to treatment group in microbial composition analysis	68
3.3.4.2	Species of the indigenous mouse microflora are displaced by <i>Salmonella</i> .69	
3.3.4.3	Response to streptomycin and <i>S. Typhimurium</i> infection at the level of phyla	69

3.3.4.4	Changes in proportional abundance of specific genera in response to infection	70
3.3.4.5	Effects of streptomycin treatment and <i>S. Typhimurium</i> infection upon microbial diversity	71
3.3.5	Blood leukocyte populations respond to infection in the streptomycin mouse model	75
3.3.6	Transcriptional changes in caecal tissue during <i>S. Typhimurium</i> -induced inflammation	78
3.3.6.1	Naïve and infected caecal tissue produce distinct transcriptional profiles	78
3.3.6.2	Biological functions of genes most highly regulated in <i>S. Typhimurium</i> infection	81
3.3.6.3	Pathway analysis of transcripts regulated in infection with <i>S. Typhimurium</i>	83
3.4	Discussion	87

4 Proteomic analysis of tissues from the streptomycin mouse model and integration with transcriptomic data..... 94

4.1	Introduction	94
4.1.1	Label-free mass spectrometry for large scale tissue proteomics	94
4.1.2	Post-transcriptional control of gene activity	96
4.1.3	Concordance between RNA transcript and protein abundances	98
4.2	Aims of the work described in this chapter	99
4.3	Results	100
4.3.1	Proteomic analysis of <i>S. Typhimurium</i> -infected caecal tissue	100
4.3.2	Comparative analysis of transcriptomic and proteomic datasets	104
4.3.2.1	The caecal transcriptome is more dramatically changed in <i>S. Typhimurium</i> infection than the proteome	104
4.3.2.2	Changes in abundance of proteins encoded by transcripts most highly regulated in <i>S. Typhimurium</i> infection	107
4.3.2.3	Correlation between changes in transcript and protein abundance during <i>S. Typhimurium</i> infection	108
4.3.3	Pathway analysis of consensus regulated genes in transcriptome and proteome datasets	110
4.4	Discussion	115

5 Signatures of the streptomycin mouse model: the complement system in mucosal <i>S. Typhimurium</i> infection	121
5.1 Introduction.....	121
5.1.1 The complement system in the immune response to infection.....	121
5.1.1.1 Three complement activation pathways converge upon a C3 convertase.	121
5.1.1.2 The three major effector functions of the complement system	122
5.1.1.3 Regulation of the complement cascade	124
5.1.1.4 Proteolytic cleavage of the major complement protein C3	125
5.1.1.5 Complement control of <i>Salmonella</i>	127
5.1.1.6 Local complement protein synthesis	128
5.1.1.7 The role of complement in the gastrointestinal mucosa.....	128
5.2 Aims of the work described in this chapter	131
5.3 Results.....	131
5.3.1 Complement upregulation during <i>S. Typhimurium</i> infection in the gastrointestinal mucosa	131
5.3.2 Detection of C3 and C3d by Western blotting	132
5.3.2.1 Failure to detect C3 in luminal content or faeces by Western blot	133
5.3.2.2 C3 and C3 activation fragments in caecal tissue and plasma.....	134
5.3.3 Localisation of C3 in caecal tissue by immunofluorescence staining	138
5.4 Discussion	142
6 Signatures of the streptomycin mouse model: investigating host susceptibility using mutant mice.....	147
6.1 Introduction.....	147
6.1.1 Systematic screening of knockout mice for the detection of novel phenotypes in infection susceptibility	147
6.1.2 The role of <i>IL10rb</i> in infection and immunity.....	151
6.1.3 The role of <i>IL22ra1</i> in the intestinal mucosa.....	153
6.2 Aims of the work described in this chapter	154
6.3 Results.....	154
6.3.1 Expression of <i>IL22ra1</i> and <i>IL10rb</i> in mice homozygous for the <i>tm1a</i> mutant allele	154
6.3.2 <i>S. Typhimurium</i> colonisation and dissemination in <i>IL10rb</i> ^{tm1a/tm1a} mice following streptomycin pre-treatment.....	159

6.3.3	<i>S. Typhimurium</i> infection of <i>IL22ra1</i> ^{tm1a/tm1a} mice in the streptomycin model	160
6.3.4	The transcriptome of <i>IL22ra1</i> ^{tm1a/tm1a} and <i>IL10rb</i> ^{tm1a/tm1a} mice in <i>S. Typhimurium</i> infection	163
6.3.4.1	Signatures of T cell activation in the caecum of <i>IL10rb</i> mutant mice during <i>S. Typhimurium</i> infection.....	165
6.3.4.2	Common genes are differentially expressed in <i>S. Typhimurium</i> -infected caecum from <i>IL10rb</i> and <i>IL22ra1</i> mutant mice	166
6.3.5	<i>S. Typhimurium</i> infection of <i>BC017643</i> ^{tm1a/tm1a} mice in the streptomycin model	167
6.4	Discussion.....	169
7	Final discussion	174
7.1	New opportunities presented by high throughput technologies.....	174
7.2	A role for complement in gastrointestinal <i>S. Typhimurium</i> infection.....	176
7.3	Insight into genes involved in the host defence to infection from the streptomycin mouse model.....	177
8	References.....	181
	Appendices	205
	Appendix 1. Reactome pathways significantly associated with transcripts upregulated in wild type <i>S. Typhimurium</i> infection.	205
	Appendix 2. Reactome pathways significantly associated with transcripts downregulated in wild type <i>S. Typhimurium</i> infection.	207
	Appendix 3. GO terms significantly associated with transcripts regulated in wild type <i>S. Typhimurium</i> infection	210
	Appendix 4. Reactome pathways significantly associated with genes upregulated at the RNA and protein level in wild type <i>S. Typhimurium</i> infection.....	211
	Appendix 5. Reactome pathways significantly associated with genes downregulated at the RNA and protein level in wild type <i>S. Typhimurium</i> infection.....	213
	Appendix 6. Differentially expressed genes in caecal tissue from mutant versus wild type mice, as determined by RNAseq.....	215

List of figures

Figure 1.1. Structures and cell types of the gastrointestinal mucosa.....	5
Figure 1.2. Altered microbiota composition in intestinal inflammation.	13
Figure 1.3. Classification of the genus <i>Salmonella</i>	14
Figure 1.4. Phylogenetic relationships within serovars <i>S. Typhimurium</i> and <i>S. Typhi</i>	21
Figure 1.5. Interactions between <i>S. Typhimurium</i> , the microbiota and the intestinal mucosa.	24
Figure 1.6. Epithelial cell invasion by <i>S. Typhimurium</i>	26
Figure 1.7. Adaptations of <i>S. Typhimurium</i> to the inflamed intestinal mucosa.....	28
Figure 1.8. Antibiotic treatment in the streptomycin mouse model overcomes colonisation resistance.....	31
Figure 1.9. Microscopic changes in the mouse colon during infection with <i>C. rodentium</i>	35
Figure 1.10. The major pathways in the mucosal response to <i>S. Typhimurium</i>	41
Figure 3.1 Weight loss and tissue colonisation in the streptomycin mouse model	64
Figure 3.2. Striking changes in caecum morphology in <i>S. Typhimurium</i> infection.....	65
Figure 3.3. Inflammatory changes in <i>S. Typhimurium</i> -infected caecal tissue	67
Figure 3.4. Effects of streptomycin treatment and infection with <i>S. Typhimurium</i> SL1344 on the community structure of the microbiota.....	75
Figure 3.5. Flow cytometry analysis of leukocyte populations in peripheral blood during <i>S. Typhimurium</i> infection.....	78
Figure 3.6. Principal component analysis of caecal transcriptional profiles	79
Figure 3.7. Infection with <i>S. Typhimurium</i> remodels the transcriptome in the mouse caecum.	81
Figure 3.8. Expression profiles of the 30 most differentially expressed genes during <i>S. Typhimurium</i> infection of caecum	83
Figure 3.9. Pathway overrepresentation analysis of genes regulated in <i>S. Typhimurium</i> infection	86
Figure 4.1. Quantitative shotgun proteomics.....	96

Figure 4.2. MS analysis of the murine caecal proteome at day 4 PI with <i>S. Typhimurium</i> SL1344.....	102
Figure 4.3. Comparison of transcripts and proteins differentially regulated during infection with <i>S. Typhimurium</i> in murine caecum	105
Figure 4.4. Quantile plot of log2 fold change in protein abundance during <i>S. Typhimurium</i> infection	106
Figure 4.5. Correlation between fold changes in transcript and protein abundance at day 4 PI with <i>S. Typhimurium</i> in mouse caecum	109
Figure 4.6. Venn diagram to show overlap between significantly regulated transcripts and proteins	110
Figure 4.7. Regulation of the KEGG pathway ‘Complement and coagulation cascades’ in <i>S. Typhimurium</i> infection.....	115
Figure 5.1. Complement activation pathways and major effector functions of activated complement.....	124
Figure 5.2. Complement activity is controlled by multiple regulatory mechanisms.....	125
Figure 5.3. Proteolytic processing of C3 is important for both activation and termination of effector functions	126
Figure 5.4. Western blotting for C3 in intestinal content and faeces	134
Figure 5.5. Western blotting for C3 and the activation fragment C3d	137
Figure 5.6. Immunofluorescence staining of C3 and CD34 in caecal tissue.....	139
Figure 5.7. Immunofluorescence staining of C3d and CSA in caecal tissue.....	140
Figure 5.8. Three-dimensional confocal imaging of C3 and CSA in caecal tissue.....	141
Figure 6.1. Mutant allele design used in the IKMC programme to mutate all protein coding genes in the mouse strain C57Bl/6N	148
Figure 6.2. Summary of current screening for phenotypes relating to infection and immunity directed by the 3i consortium	149
Figure 6.3. IL10 inhibits the major steps in activation of an inflammatory immune response	152
Figure 6.4. Analysis of <i>IL10rb</i> and <i>IL22ra1</i> expression levels in tm1a mutant mouse lines.	159

Figure 6.5. <i>Salmonella</i> burden in <i>IL10rb</i> ^{tm1a/tm1a} mice at day 4 PI with <i>S. Typhimurium</i> following streptomycin treatment.....	160
Figure 6.6. Burden of <i>Salmonella</i> in organs, weight loss, and intestinal inflammation in <i>IL22ra1</i> ^{tm1a/tm1a} mice at day 4 PI with <i>S. Typhimurium</i> following streptomycin treatment	163
Figure 6.7. Overlap between genes differentially expressed in caecal tissue from <i>S. Typhimurium</i> -infected <i>IL22ra1</i> ^{tm1a/tm1a} and <i>IL10rb</i> ^{tm1a/tm1a} mice compared with wild type controls	167
Figure 6.8. Burden of <i>Salmonella</i> in organs of <i>BC017643</i> ^{tm1a/tm1a} and wild type mice at day 3 PI.....	168

List of tables

Table 2.1. Antibodies used in this study.....	45
Table 4.1. The 50 most highly up- or down-regulated genes in the transcriptome for which protein abundance was also significantly regulated.	101
Table 4.2. The 10 Reactome pathways most significantly associated with genes upregulated in the transcriptome and proteome during <i>S. Typhimurium</i> infection.....	104
Table 5.1. Complement pathway genes upregulated both in RNA and protein in <i>S. Typhimurium</i> -infected caecal tissue.....	122

1 Introduction

1.1 The global burden of infectious disease

Despite increasing availability of therapeutics and direct targeting by healthcare systems, infectious disease remains a major global challenge [1, 2]. The spectrum of infections differs geographically according to climate and environmental factors, and is closely related to economic development [2, 3]. Common infections in economically developed countries are often associated with healthcare settings, where there are higher concentrations of compromised individuals [4]. In contrast, many classical infectious diseases such as typhoid and cholera remain common in less economically developed regions [5, 6]. Infectious diseases can transmit directly between humans, from the environment, or from zoonotic sources, and the contribution of different transmission routes varying according to region.

Infection-related illness represents a dynamic problem. In an age of globalisation, with local and international travel greater than ever before, there are a growing number of infectious agents moving readily between healthcare institutions, countries and even continents, presenting an integrated global threat [7, 8]. Microbes can evolve rapidly, in part due to their short generation time, with selective pressures such as human therapeutic interventions (including antibiotics) driving change. The emergence of antibiotic resistant microbes and pathogen evasion of current vaccines further compound the challenge of preventing and treating infectious disease, and there is increasing recognition that new approaches are required to combat such threats [9].

Fortunately we live in an era in which new technologies and research themes are generating deeper insight into the molecular basis of infection and the epidemiology of disease, providing understanding, which we hope will lead to future effective disease interventions. Epidemiological work has helped to understand the makeup of the human (and zoonotic) population exposed to the infection threat, and guides policy decisions, such as prioritising between therapies and infrastructure. We are beginning to discover how colonisation by microbes (the microbiota) can influence health in a multitude of ways, from protection against infection, to involvement in diseases previously classified as non-communicable. For example, conditions such as inflammatory bowel disease and certain

allergies may be associated with dysbiosis, involving host colonisation by ‘unhealthy’ microbial communities which drive pathogenic immune responses [10]. Indeed, pathogenic microbes may be evolving from such communities [11].

1.1.1 Incidence, morbidity and mortality of diarrhoeal disease

Diarrhoeal diseases remain a major cause of illness and mortality. Diarrhoea can be caused by many classes of pathogen, including viruses (e.g. *Rotavirus*), bacteria (e.g. certain strains of *Escherichia coli* and *Salmonella*) and parasites (e.g. *Giardia* spp. and *Cryptosporidium* spp.); by the microbes themselves, and through intoxication by the toxins they produce. Thus the epidemiology of diarrhoea is complex, although there may be common underlying mechanisms associated with this syndromic event. Despite the large number of causes of diarrhoea, paediatric cases at healthcare centres in the developing world can be described by a just a few clinical syndromes. Simple gastroenteritis, accounting for around 80% of episodes, is characterised by loose watery stool, low grade fever, occasional vomiting, anorexia, weakness and discomfort. A further 5 - 15% of cases are described as overt dysentery, with obvious bloody diarrhoea indicating extensive damage to the intestinal mucosa, accompanied by fever, sometimes high. More rare are cases of profuse watery diarrhoea leading rapidly to severe dehydration, persistent diarrhoea, and diarrhoea accompanying extensive vomiting [12].

Young children are most severely affected; in 2011 diarrhoeal disease was estimated to account for 10% of the 6.9 million deaths in children under five, and of these 72% of deaths occurred in infants under two years of age [13, 14]. In developing countries classical childhood diarrhoea is extremely common; in 2010 there were an estimated 1.7 billion episodes of diarrhoea globally, with the highest incidence of disease in Africa and Southeast Asia [14]. Just 15 countries, including some of the poorest nations of the world, account for over half of the known global cases of diarrhoea. Progression to severe episodes of disease and case-fatality rates are much higher in low-income regions; Southeast Asia and Africa each contributed 26% to the total number of severe cases of diarrhoea in 2011 [14]. The intimate relationship between under-nutrition and diarrhoeal disease is a major factor in determining the distribution of severe disease. Chronic malnutrition predisposes children to severe diarrhoeal disease, which in effect often causes injury to the gut, leading to impaired nutrient absorption and slower growth. Consequently diarrhoea presents a major obstacle to child development in poorer nations [15].

Though a continuing obstacle to human health, reductions in diarrhoeal disease seen in recent years give cause for optimism. In the first 10 years of the millennium deaths in children under five decreased by 2 million globally, with reductions in diarrhoea, measles and pneumonia contributing most significantly to this decrease, diarrhoea individually contributing 18% [16]. The advent of the new millennium brought several initiatives with commitments to reduce global diarrhoeal disease including the UN Millennium Declaration and Development Goal 4 to reduce the under-five mortality rate, and the launch of the Bill and Melinda Gates Foundation, and Global Alliance for Vaccines and Immunization (GAVI). Improvements in hygiene and sanitation infrastructure are helping to cut transmission of enteric pathogens, and the improved availability of vaccines is contributing also [17]. Many significant challenges remain. The need to increase coverage of positive interventions is holding up progress. Vaccines against enteric pathogens with high efficacy in developed countries have in several cases failed to recreate the same levels of protection in developing nations [18].

A lack of detailed investigations into the burden and aetiology of diarrhoeal disease has until recently presented a barrier to implementation of effective interventions to combat it. The Global Enteric Multicentre Study (GEMS) was funded in 2006 to address this need, enrolling over 9,000 children with moderate diarrhoea and over 13,000 control children without diarrhoea across four sites in Africa and three in Asia [12, 19]. The study presented many interesting findings regarding the distribution and prevalence of enteric pathogens. Most cases of moderate to severe diarrhoea in children aged zero to five attributable to a specific pathogen were caused by *Rotavirus*, *Cryptosporidium* spp., Enterotoxigenic *Escherichia coli* (ETEC) producing heat-stable toxin, and *Shigella* spp. Some pathogens were strongly associated with the presence of diarrhoea but many were weakly associated, resulting in difficulties in determining the aetiological agent behind many diarrhoeal episodes. Clear patterns emerged in pathogen distribution; certain pathogens such as *Rotavirus* and *Cryptosporidium* spp. were found to be a common cause of diarrhoea at all test sites, whilst others, including the genus *Salmonella*, showed a more restricted distribution.

In developed countries, infectious diarrhoea is often zoonotic in origin and transmitted via food contaminated with animal waste, or improperly cooked meat products [20, 21]. Diarrhoea is also associated with antibiotic treatment and hospitalisation, with *Norovirus* and *Clostridium difficile* major causes of hospital diarrhoea. The economic impact of these is sizeable; in 2002 it was estimated that the cost of *C. difficile* disease in the United States

exceeds \$1.1 billion per year [22]. Mortality from these infections is largely restricted to the elderly and patients with comorbidities [23, 24].

1.2 Biology of the gastrointestinal tract

1.2.1 The gastrointestinal mucosa

The human gastrointestinal tract runs from the mouth to the anus with an approximate length of nine metres, including multiple organs and regions with specialised structure and function in the absorption of nutrients from the diet. The innermost region surrounding the lumen, the mucosa, comprises three layers; at the centre the epithelium, surrounded by lamina propria followed by the muscularis mucosae. The following sections describe the mucosa of the small and large intestine.

1.2.1.1 The epithelium

The epithelium comprises a single layer of polarized cells arranged in crypt and villus structures that forms a physical barrier separating host tissues from the external environment. With a surface area of $\sim 400 \text{ m}^2$ the intestinal mucosa is the largest mucosal surface of the body. The majority of intestinal epithelial cells are absorptive enterocytes, the major function of which is the transport of dietary nutrients such as glucose and amino acids across the epithelial barrier. In addition, enterocytes perform important functions in sensing and responding to microbial stimuli to trigger appropriate immune responses; both tolerance to dietary antigens and immune activation for pathogen responses [25]. Positioned between the absorptive enterocytes are epithelial cell types of specialised function, including intestinal epithelial stem cells and secretory cells including paneth, goblet, and enteroendocrine cells.

Intestinal epithelial stem cells reside at the base of the mucosal crypts and divide indefinitely to replace the huge number of epithelial cells which die and are sloughed off the surface of the epithelium. The environment of the intestinal lumen is harsh. Toxic molecules originating from both the diet and the microbial community are in continual contact with cells, and consequently the turnover time of the gastrointestinal epithelium is as little as ≤ 3 days [26]. Division of intestinal adult stem cells produces highly proliferative progenitor cells. As these migrate upwards from the crypt base they differentiate to form all the different epithelial cell lineages, with the exception of paneth cells which develop through a separate dedicated pathway [27].

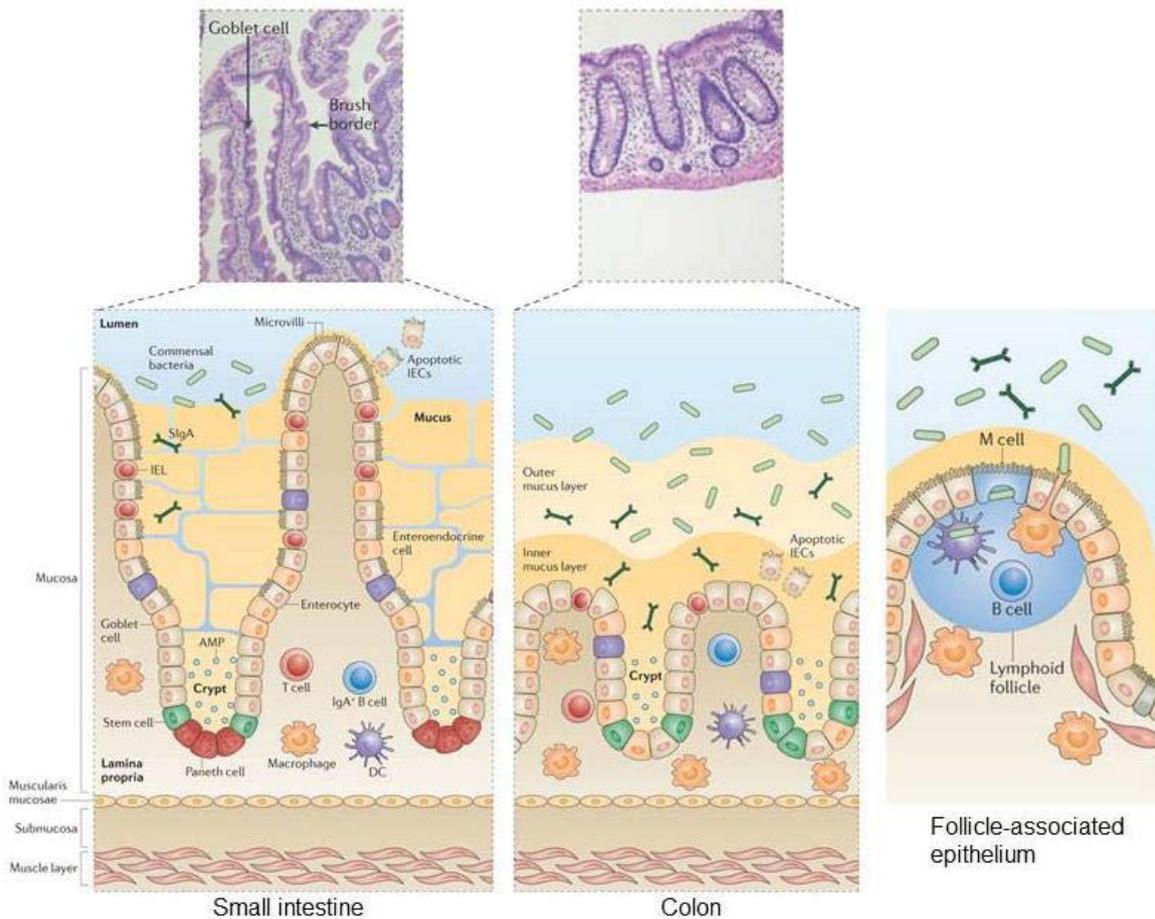


Figure 1.1. Structures and cell types of the gastrointestinal mucosa. Upper - Hematoxylin and eosin stained histological sections of human jejunum and colon. Lower - Schematic representations of the mucosa in the small intestine and colon, and the follicle-associated epithelium. Small intestine is characterised by deep crypts and villi while colonic crypts are much shorter. The intestinal epithelium in both organs consists of absorptive enterocytes and specialised cell types including adult stem cells, enteroendocrine cells and intestinal epithelial lymphocytes. Beneath the epithelial barrier the lamina propria contains an assortment of immune cells positioned to respond rapidly to microbes which breach the barrier and enter this region. A thick coating of mucus on the epithelial surface limits contact between intestinal content and the epithelial cells and contains secretory IgA antibodies which can bind and aggregate bacteria to prevent epithelial translocation. Follicle-associated epithelium are highly specialised regions containing microfold cells for sampling luminal antigens and closely associated phagocytes for uptake of these antigens and their presentation to adaptive immune cells. Adapted from [28] and [25].

Paneth cells remain localised to the crypts where they reside in small intestine. They play an important role in restriction of bacteria to the lumen through intensive secretion of antimicrobial peptides; both constitutively to prevent bacterial invasion of the crypt and epithelial mucus layer, and in response to specific microbial triggers. The presence of antimicrobial peptides influences the composition of the

microbiota at the mucosal surface, an important aspect of homeostasis to prevent excessive pro-inflammatory responses [29].

Goblet cells are located throughout the small and large intestine, increasing in density from proximal to distal colon. Goblet cells synthesise and secrete high molecular weight glycoproteins called mucins which form a thick protective glycocalyx over the surface of the epithelium. Mucins possess a net negative charge as a result of their acidic carbohydrate content and upon secretion form a network of ionic interactions. They can directly bind and trap certain bacteria through the carbohydrate moieties, some of which mimic epithelial cell surface molecules targeted by bacterial lectins. Pathogens may also become trapped in the mucus layer through binding by secretory immunoglobulin A (sIgA) within the mucus [30].

A third secretory cell type, enteroendocrine cells, comprise < 1% of the epithelial cell layer and these can be divided into multiple subtypes classified by the hormone and peptide content of their secretory vesicles. Molecules secreted by the enteroendocrine cells are important for regulation of a wide range of activities from gut motility to lipid absorption [31, 32].

In addition to the secretions from the cells within, the structure of the epithelium itself is adapted to prevent bacteria from accessing the tissue beneath. Sealing structures called tight junctions formed from integral membrane proteins directly below the apical surface encircle polarized epithelial cells. These connect adjacent cells to form a barrier relatively impermeable to small molecules and larger objects such as bacteria.

1.2.1.2 The lamina propria and muscularis mucosae

Surrounding the epithelial cell layer are multiple layers of connective tissue and muscle; first the lamina propria and the muscularis mucosae of the mucosa, followed by the submucosa, muscularis externa and the outermost layer of adventitia or serosa. The connective tissue of the lamina propria provides a structural base for the crypt and villus structures above, and contains the blood supply, local nervous system, and lymph drainage for the mucosa and immune cells. The muscle layers of the muscularis mucosae and muscularis externa are important to aid agitation and passage of lumen contents through peristalsis.

1.2.2 The gut-associated immune system

The intestinal immune system is challenged with the complex task of permitting commensal microbiota to safely inhabit the gut, avoiding the triggering of detrimental inflammation, yet providing appropriate protective immune responses against pathogens. A system of multiple specialised immune compartments with different immune cell compositions and functions has evolved to meet this demand [33]. The intestinal immune system comprises the mesenteric lymph nodes (mLN), and organised secondary and tertiary lymphoid tissues within the mucosa and submucosa. Together, these are referred to as gut-associated lymphoid tissue (GALT), with immune effector cells diffusely localised throughout the lamina propria and epithelium.

1.2.2.1 mLN and secondary lymphoid tissue

The mLN and organised structures of the GALT are the main sites of immune induction and regulation in the intestine. The formation of mLN and the patches of the small intestine (Peyer's patches), caecum and colon, begins early in embryonic development [34]. Like lymph nodes elsewhere in the body, the mLN sample antigens delivered via an afferent lymph vessel. The intestinal patches are located within the mucosa and submucosa, and in contrast to mLN, antigens are sampled from the intestinal lumen through a region of specialised columnar epithelium, the follicle-associated epithelium (FAE). FAE contains large numbers of highly specialised epithelial cells called microfold (M) cells. An opening in the glycocalyx assists M cells in endocytic uptake of luminal antigen at the apical surface, followed by transfer to a pocket at the basal side of the cell contacting dendritic cells, macrophages and lymphocytes in the subepithelial dome region beneath [35, 36]. Here dendritic cells process antigen and present it to lymphocytes located within the patch [37].

Intestinal patches consist of multiple B cell follicles in germinal centres surrounded by a smaller T cell zone. Germinal centres facilitate the replication, differentiation and mutation of antibody encoding genes in B cells in response to antigen stimulation within the patch. IgA-producing antigen-specific plasma cells differentiate following stimulation and are the main immune effector cell type. The total number of Peyer's patches and colonic and caecal patches varies between species but the pattern of fewer colonic and caecal patches compared with the small intestine is conserved. In mice there are typically < 10 colonic and caecal patches while small intestinal patches in humans number around 240 [38, 39].

1.2.2.2 Tertiary lymphoid tissue

Smaller tertiary intestinal lymphoid tissues, referred to collectively as solitary isolated lymphoid tissues (SILTs), perform similar functions to the secondary lymphoid structures; however these develop only after birth, and intestinal colonisation by the microbiota is required for normal development. Several types of follicle of varying size and cell composition have been described. Recent work has questioned the validity of strict definitions, suggesting that a changeable continuum modified in response to environmental signals better describes these structures compared with classical categories of cryptopatch and innate lymphoid follicle (ILF) [40]. SILTs are more numerous than the larger secondary lymphoid structures; the mouse small intestine is estimated to contain ~ 1,000 evenly distributed follicles [40]. Colonic SILTs are concentrated in the distal colon where the bacterial load is greater [38]. Tertiary intestinal lymphoid tissue growth and expansion is induced in response to microbes, resulting in increased production of antigen-specific IgA, potentially keeping microbial growth in check; this relationship acts as a feedback loop maintaining homeostasis [41]. Imbalance results in drastic effects, as seen in mice lacking these structures which show a 10-fold expansion in the microflora [42].

1.2.2.3 Individual immune cells

The epithelium and the lamina propria contain large numbers of individual immune effector cells, collectively the largest population of T cells, plasma cells and macrophages in the body [28]. Intraepithelial lymphocytes (IELs) sit within the epithelium with a frequency of around 10 - 15 per 100 epithelial cells, and these are primarily T cells [43]. Unlike most healthy tissues the lamina propria contains large numbers of plasma cells, of which 75% in the duodenum, and 90% in the colon, are IgA-producing. In addition the lamina propria contains T cells and plentiful innate cells including dendritic cells, macrophages, eosinophils, mast cells, and the recently discovered innate lymphoid cells (ILCs).

1.2.3 The intestinal microbiota

The human intestinal microbiota is a largely stable collection of at least 500 - 1,000 microbial species including bacteria, viruses, and eukaryotes [44, 45]. The relationship between the host and the microbial community is normally mutually beneficial. The host provides an environment favourable for bacterial survival; nutrient rich and at a constant

temperature. In return the microbiota performs a wide range of valuable functions. Commensal microbes metabolise dietary nutrients the host is otherwise unable to digest, and produce essential vitamins. Collectively their presence can resist the overgrowth of pathogenic bacteria; a phenomenon known as colonisation resistance. Bacterial population density and diversity in the GI tract increases from the stomach to the large intestine, with $\sim 10^3$ organisms per gram of content in the duodenum increasing to $\sim 10^{12}$ per gram in the distal colon. Total bacterial cells outnumber those in the human host 10:1 [46]. The vast majority of bacterial species are strictly anaerobic; there are 100 - 1,000-fold fewer facultative anaerobes; and belong to one of a relatively small number of bacterial phyla. Most abundant by far are species of the phyla Bacteroidetes and Firmicutes, accompanied by smaller numbers belonging to the Proteobacteria, Verrumicrobia, Actinobacteria, Fusobacteria and Cyanobacteria [45, 47].

Until recently the significance of the microbiota in development and homeostasis in the healthy gut was poorly understood, however the past decade has uncovered many processes where the microbiota are crucial, and a much wider role for the microbiota than previously imagined [48].

1.2.3.1 Colonisation resistance

Colonisation resistance is an essential protective mechanism provided by the microbiota. Studies of *Salmonella* Typhimurium infection in mice have determined that the microbiota composition strongly dictates the degree of colonisation resistance conferred; in germ-free mice and mice with a defined microbiota of 4 - 20 species, colonisation resistance is absent. Co-housing of these mice with conventional mice, allowing the acquisition of a normal microbiota, restores resistance. Further study of mice with varying degrees of colonisation resistance determined that the presence of closely related species increases the chance of colonisation by an incoming species [49]. Clearance of *S. Typhimurium* from the gut after infection relies on the re-establishment of a complex microbiota, and similar to the loss of colonisation resistance, is deficient in mice with a defined low complexity microbiota [50]. Whilst these examples show that certain important aspects of colonisation resistance are being exposed, more work is needed to uncover exactly how resistance to colonisation is achieved and to determine the relative importance of individual species, phyla, and overall complexity in protection.

Antibiotics are essential for treatment of infections the body fails to resolve independently, however antibiotic therapy frequently impacts the species which make up the microbiota in addition to the pathogen target, reducing microbial diversity and shifting the composition at the levels of species and phyla [51, 52]. These changes collapse colonisation resistance and can render patients more susceptible to gastrointestinal infection with opportunistic pathogens. *C. difficile* is carried asymptotically in the microbiota of some healthy individuals yet it is the most important nosocomial cause of diarrhoea in adults [53]. Disruption of the microbiota can allow rapid growth of the bacterium resulting in disease ranging from diarrhoea to pseudomembranous colitis, and in a small proportion of patients more serious complications including death [54]. Treatment of *C. difficile* infection involves cessation of treatment with the inciting antibiotic and employs antimicrobials targeting *C. difficile*, but whilst these can kill the *C. difficile* bacteria, indigenous microbial species are afforded little chance of recovery and relapsing infection is common. Utilising the natural protection in colonisation resistance, faecal transplantation is a promising therapeutic intervention for the treatment of *C. difficile* infection. This approach reintroduces species which comprise a healthy microbiota to the gastrointestinal tract, leading to suppression of pathogen growth [55].

1.2.3.2 The microbiota in intestinal development and maintenance

Studies in germ-free mice have vastly increased our understanding of the important functions of the microbiota in development of the gastrointestinal mucosa, in particular the full development of lymphoid tissues and production of mucus. Germ-free mice develop an enlarged caecum and defective small intestinal brush border, with a reduced total intestinal surface area relative to conventionally housed mice [56, 57]. As mentioned earlier, differentiation of secondary lymphoid tissues of the intestinal mucosa begins in the embryo, however microbial colonisation is required after birth for the increase in size and development of germinal centres during maturation [58]. The microbiota is critical for the generation and development of ILFs; the peptidoglycan of Gram-negative bacteria is needed to induce ILF genesis, and further maturation requires the detection of bacteria by toll-like receptors (TLRs) [42]. Specifically the differentiation of B cells to sIgA-producing plasma cells requires detection of Pathogen-Associated Molecular Patterns (PAMPs) such as bacterial flagellin by TLR5 on lamina propria dendritic cells [59]. sIgA targeting commensal antigens is essential

to maintain the balance in the microbiota and physical separation between the microbiota and host [60].

The microbiota is necessary to induce and maintain a strong epithelial barrier, affecting multiple components involved in integrity. Compared with conventionally housed mice, germ-free mice possess fewer ROR γ t⁺ NKp46⁺ ILCs, a key cell type for production of IL22 which signals to induce production of antimicrobial peptides, epithelial repair and barrier function [61]. In conventionally-housed mice the mucus which lines the intestinal mucosa is several hundred μ m thick in the colon. This generous blanket physically limits bacterial contact with host cells at the epithelial surface. Germ-free mice possess fewer goblet cells resulting in a thinner mucus layer, and the mucus composition is skewed towards a higher proportion of neutral mucin molecules [62]. Finally many bacterial species of the microbiota have been found to impact barrier integrity through effects on tight junctions, for example several *Lactobacillus* rigidify tight junctions via signalling through pattern recognition receptors (PRR) [63].

Certain groups and even individual species of bacteria have been linked with specific aspects of intestinal immune development and function. For example tryptophan metabolism by *Lactobacillus* spp. triggers pathways leading to Th17 cell production of IL22, and *Clostridia* spp. promote a TGF β -rich environment which leads to the preferential generation of regulatory T cells (Tregs) in the colon [64, 65].

1.2.3.3 Wider impact of the microbiota

Recent studies have begun to uncover important roles for the microbiota beyond development and maintenance of the gastrointestinal tract and immune compartment. Changes in the microbiota have been associated with tissue homeostasis at remote sites, including changes in bone density [66]. The metabolic potential of the microbiota is far greater than that of the host and it is therefore unsurprising that the microbiota is associated with metabolic disease and adiposity. The microbiota of obese mice was found to possess greater potential for metabolism of polysaccharides and synthesis of short chain fatty acids, and hence a greater ability to unlock the energy potential within food compared with mice of normal weight [67]. Future work in this area is likely to further broaden the range of host biological processes we know to be influenced by the intestinal microbial community.

1.2.3.4 The microbiota in diarrhoeal disease

The relationship of the microbiota with diarrhoeal disease is complex. Perturbation of the microbiota can lead to diarrhoeal disease, and in turn diarrhoea has large effects on microbiota composition, hence in many cases disambiguating cause and effect is a difficult task [68]. Pathogenic species significantly impact the population structure of the resident microbiota, creating a state of 'dysbiosis'. Similar changes are observed in inflammatory bowel disease (IBD) and other autoimmune diseases, indicating these changes are a characteristic of inflammation occurring in the gut rather than the result of pathogen-specific mechanisms. Studies of gastrointestinal infection with diverse pathogens and in IBD have shown outgrowth of the class Gammaproteobacteria, and in particular the *Enterobacteriaceae* family is a common feature [69-71]. Conversely fermentative bacteria of the phylum Firmicutes and in particular the class Clostridia are lost during inflammation [71]. The shift in composition reflects differences in adaptation of commensal species to the hostile environment of the inflamed gut. Species which thrive are often closely related to obligate pathogens and are able to express genes which help them to take advantage of the environment, for example to allow them to use metabolites such as nitrates [72]. Inflammation leads to dysbiosis, and in turn this shapes the immune system, creating positive feedback. Many of the species seen to overgrow or 'bloom' in the inflammatory conditions are inflammatory themselves, particularly in the environment of the damaged mucosa, and contribute to the immune activation taking place. Compounding problems further is the loss of normal inflammation dampening mechanisms. Fermenting bacteria reduced in inflammation are often important immunoregulators, for example Treg-inducing *Clostridia* spp. are amongst those whose numbers are diminished [65].

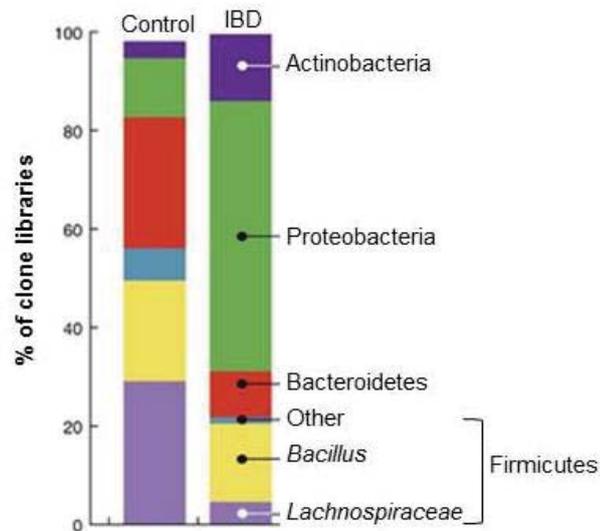


Figure 1.2. Altered microbiota composition in intestinal inflammation. Representative microbiota composition data for IBD patients and healthy controls. Bar shading represents bacterial groups and bar height the proportion of cloned sequences in the sample belonging to the group. Expansion of the phylum Proteobacteria and a decline in the phylum Firmicutes is typical in inflamed intestine. Adapted from [73].

1.3 *Salmonella*

1.3.1 Classification and phylogeny

The salmonellae are Gram-negative bacteria belonging to the class Gammaproteobacteria and the family *Enterobacteriaceae*. They are facultative intracellular bacteria and a model organism for the study of bacterial genetics and pathogenesis. The *Salmonella* genus diverged from closely related *E. coli* in the region of 100 - 140 million years ago, and diverged again between 71 and 100 million years ago to form two lineages which define the *Salmonella* species *S. bongori* and *S. enterica*. Two key genetic elements associated with virulence, *Salmonella* pathogenicity islands 1 & 2 (SPI-1 and SPI-2), were acquired by horizontal gene transfer after divergence from *E. coli*, in the case of SPI-1 prior to the split between *S. bongori* and *S. enterica*, whilst SPI-2 was acquired by the *S. enterica* lineage exclusively [74].

S. enterica is divided into seven subspecies, the first six originally defined by biochemical traits and the seventh distinguished by multi-locus enzyme electrophoresis (MLEE), though DNA-sequence based analyses confirms these subspecies are indeed evolving independently [75-77]. The salmonellae are further divided into serotypes based upon the O-antigen of LPS and flagellar or H- antigen. Combining the seven subspecies with

the 46 O-antigen groups and 114 H-antigen groups defines all recognised serotypes of *Salmonella*. *S. bongori* and *S. enterica* subspecies II, IIIa, IIIb, IV, and VI are mainly associated with cold-blooded vertebrates whilst members of *S. enterica* subspecies I often infect mammals and birds. Greater than 99% of *Salmonella* isolates from human and mammalian hosts belong to subspecies I [78, 79].

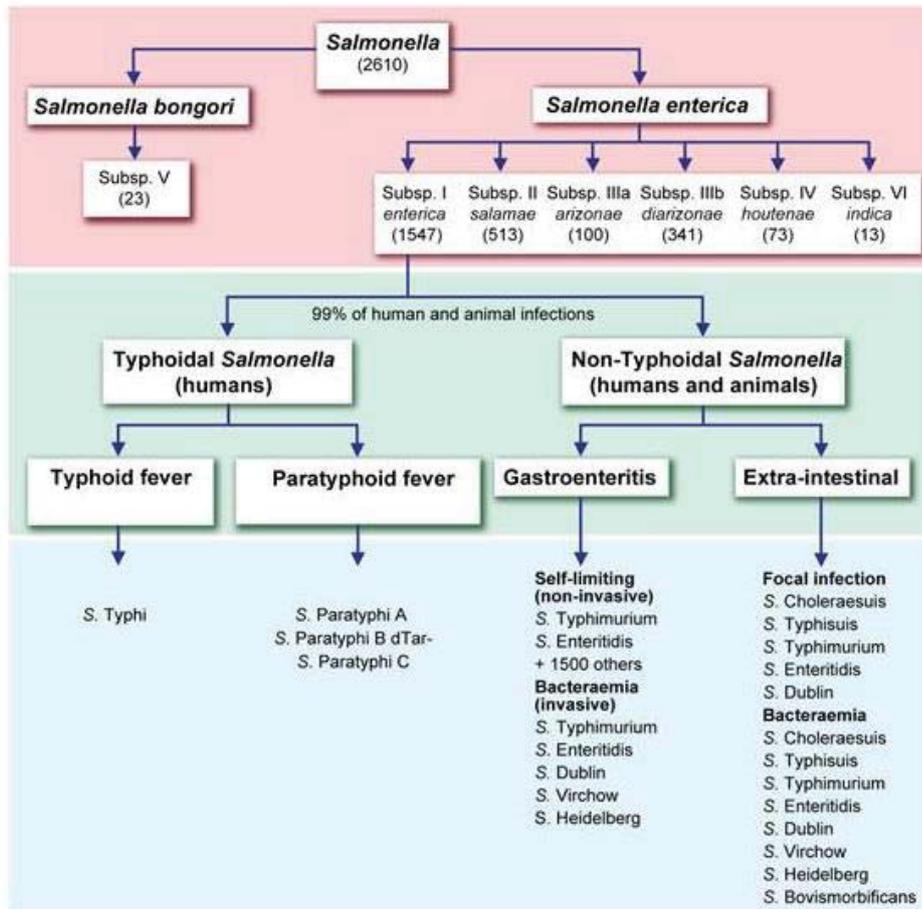


Figure 1.3. Classification of the genus *Salmonella*. Two species form the *Salmonella* genus, of these the majority of isolates belong to the species *S. enterica*. *S. enterica* is divided into subspecies according to biochemical and serological characteristics. *S. enterica* subspecies I accounts for the vast majority of human infections and is further divided into typhoidal and non-typhoidal salmonellae by the disease symptoms these bacteria generate. Within these two groups are around 1,500 serovars, characterised by the O- and H- antigens of the bacterial surface. Taken from [80].

1.3.2 Global distribution of *Salmonella* serotypes

Data collected through the World Health Organisation (WHO) Global Salm-Surv describe the distribution of *Salmonella* serotypes isolated by its members. 41 countries entered results for human isolates contributing to the report for 2000 - 2002. Of the > 350,000 isolates, *S. Enteritidis* was most common, accounting for 65% of all isolates, followed by *S. Typhimurium* at 12% and *S. Newport* at 4%. The vast majority of contributing isolates were North American and European and unfortunately the results are highly biased toward the industrialised world. Regional statistics show that serotypes vary greatly in proportion but more data is needed for developing nations. Serotypes *S. Enteritidis*, *S. Typhimurium*, *S. Infantis*, *S. Montevideo*, and *S. Typhi* displayed a broad geographical distribution whilst serotypes *S. Rissen* and *S. Weltevreden* were region specific. Overall the majority of human cases are caused by a small number of serovars [81, 82].

1.3.3 Typhoidal and non-typhoidal *Salmonella* (NTS)

Salmonella serotypes responsible for human infection are categorized by the clinical manifestations of disease and host range, into invasive (typhoidal) and non-invasive (non-typhoidal) strains.

1.3.3.1 Clinical features and burden of typhoid fever

Typhoidal salmonellae cause life-threatening systemic infection and are adapted to a human host. Infection is characterised by high fever and potentially complicated by sepsis and shock, gastrointestinal bleeding or perforation, encephalopathy and enlargement of the mLN, liver and spleen accompanied by granulomatous lesions [83]. *Salmonella enterica* serovar Typhi causes the largest proportion of typhoidal disease, with serotypes Paratyphi A, B and C also responsible for significant numbers of infections. Historically typhoidal *Salmonella* infections were a worldwide problem but now have been largely eradicated in developed countries as a result of improved sanitation. In the developing world infections remain a significant cause of morbidity and mortality. In 2000 there were an estimated 21 million cases of typhoid fever resulting in around 200,000 deaths, and over 5 million cases of paratyphoid fever [6]. Multiple typhoid vaccines are available but the most effective licensed vaccines achieve incomplete protection at around 70%, and immunization coverage in developing nations is poor [84].

1.3.3.2 Clinical features and burden of NTS disease

NTS serovars are more common than *S. Typhi* and are globally distributed; *Salmonella* gastroenteritis is a significant burden in both developed and developing countries. Globally there are an estimated 93.8 million cases of *Salmonella* gastroenteritis resulting in 155,000 deaths annually [85]. 1.4 million cases occur in the United States [86]. Non-typhoidal salmonellae infect mammals and birds in addition to humans, and the majority of human infections are thought to be the result of zoonotic transmission. Typical NTS infections are cases of self-limiting gastrointestinal disease, typified by diarrhoea and intestinal inflammation. Symptoms also include fever, and abdominal cramps. Following ingestion bacteria pass through the stomach to colonise the terminal ileum and colon; typically generating symptoms of gastroenteritis within 24 h. Early inflammatory signals are amplified to promote continued exudative inflammation in the mucosa. Vascular and epithelial permeability increases to allow neutrophils to enter the tissue and attack *Salmonella* that has invaded beyond the epithelial layer, and further leave the tissue to act on luminal *Salmonella*. Fluid leakage results from the increased permeability and contributes to diarrhoea, clearing out the contents of the gut and leading to dehydration.

1.3.3.3 Invasive non-typhoidal *Salmonella* (iNTS)

In developed nations bacteremia in NTS infection is a relatively rare occurrence, largely limited to individuals with inherited immunodeficiencies, or immunosuppression as a result of other factors such as steroid use or diabetes. In contrast NTS are a common cause of bloodstream infections in adults and children presenting with fever in many parts of sub-Saharan Africa, with a case fatality rate of 20 - 25%. iNTS is closely associated with specific risk factors. 95% of adult cases are in people with HIV, whilst in children HIV, malaria, and malnutrition are all closely associated with this type of infection [87]. The major feature of typical NTS infection, enterocolitis, is often absent and replaced by diverse clinical presentations including respiratory symptoms, fever and hepatosplenomegaly. Most cases of iNTS are caused by the serotypes *S. Enteritidis* and *S. Typhimurium* but there is a general lack of data describing whether the same strains responsible for NTS are also causing iNTS disease. Recent evidence points to the emergence of new strains in sub-Saharan Africa, which may be adapted to the niche provided by immunocompromised individuals [88]. The dominant regional genotype of iNTS in Malawi and Kenya is *S. Typhimurium* ST313 and

recent work has identified signatures of genomic degradation and convergence with phylogenetically distinct *S. Typhi* (see section 1.3.5.2.2) [89].

The association between HIV and iNTS may be explained by the viral destruction of CD4⁺ T cells. In particular the loss of Th17 cells results in decreased cytokines critical for epithelial barrier function, such as IL22, and those for the recruitment, activation and survival of neutrophils, for example TNF α and GM-CSF. These weaknesses in host defence combine to enable *Salmonella* to disseminate more effectively, as has been demonstrated in an SIV rhesus macaque model. The diminished enterocolitis in iNTS infection is thought in part due to the failure to recruit neutrophils to the mucosa [90].

1.3.4 *Salmonella* transmission

Salmonella is predominantly transmitted by the faecal-oral route in contaminated food and water, and through consumption of animal products from species that harbour the bacteria. The relative importance of food- and water-borne transmission is thought to vary with regional economic development. In developing nations a lack of basic sanitation significantly aids transmission, providing increased opportunity for bacteria from the faeces of infected animals or individuals to contaminate water used for drinking and food preparation. In contrast the most common sources of *Salmonella* in developed nations are meat and egg products, and processed food contaminated with faeces from animal sources.

In the majority of cases of NTS *Salmonella* is completely eliminated from the gastrointestinal tract in recovery from infection, though antibiotic treatment may increase the chance of persistence [91]. On the other hand *S. Typhi* has a human reservoir in the form of asymptomatic carriers. Approximately 2 - 5% of typhoid patients fail to fully clear the bacterium within one year, asymptotically carrying *Salmonella* in the gall bladder and biliary tract [92, 93]. *S. Typhi* is thought to enter the gall bladder directly from the liver during the systemic phase of infection. Persistence is closely associated with gall stones; ~90% of chronic carriers of *S. Typhi* can have them; and it is thought that formation of biofilms on gall stones is a key survival strategy employed by the bacteria [94, 95].

1.3.5 Molecular phylogeny of *S. enterica*

1.3.5.1 Serological and DNA-based classification

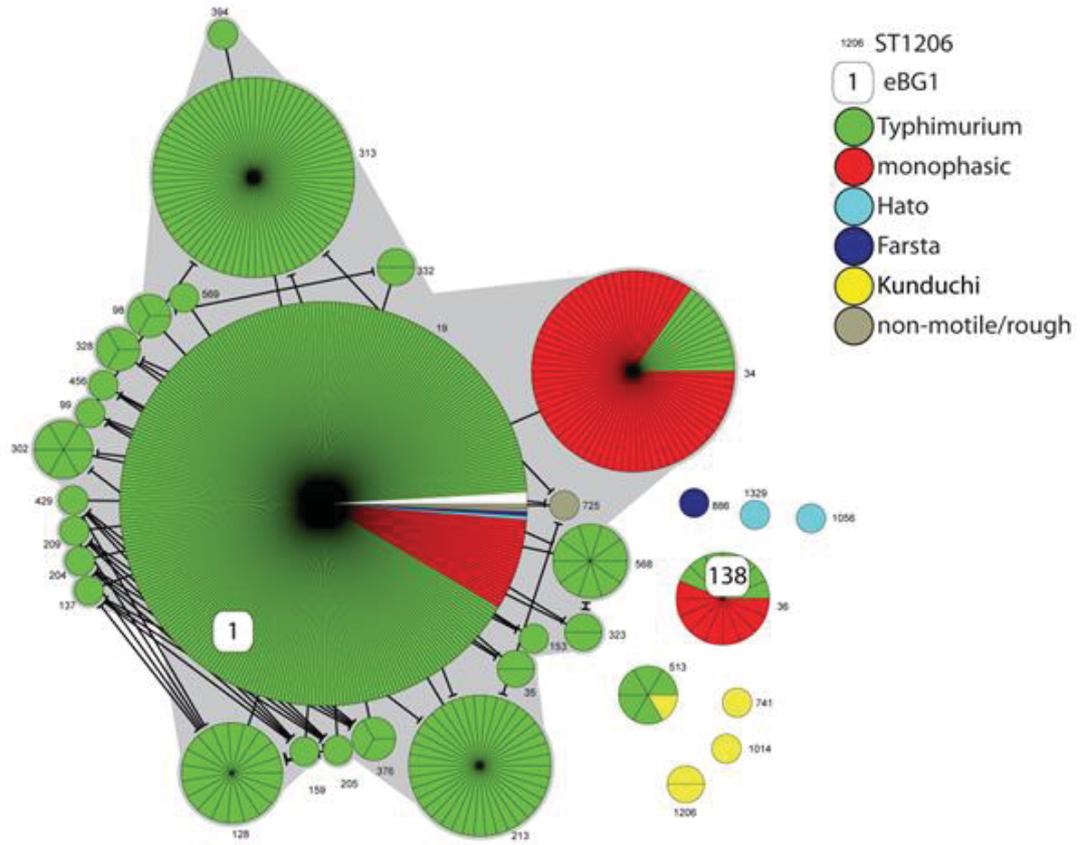
For almost a century, isolates of *Salmonella* have been classified according to serological and nutritional characteristics into serovars (see section 1.3.1). This globally accepted system, known as the Kauffmann-White scheme, has served to facilitate the epidemiological analysis of *Salmonella* infections. The scheme exploits diversity in the O (lipopolysaccharide) and H (flagella) antigens of *Salmonella*. However serological analysis lacks the ability to discriminate between closely related isolates and provides limited phylogenetic information compared to new approaches such as whole genome DNA-based typing systems. Further, although they have been tremendously useful, serotyping methods are comparatively slow, costly and low throughput.

Multi-locus Sequence Typing (MLST) was proposed to replace traditional serological classification of *Salmonella* in 2012 [80]. Based on sequences at sites within normally seven housekeeping genes, MLST uses neutral markers to identify genetically related isolates of *Salmonella* and classifies them into ‘sequence types’ (ST). Clustering algorithms group sequence types at a higher level into evolutionarily related groups. The use of this approach to study over 4,000 isolates from more than 500 serovars showed that the population structure of *S. enterica* is described by many discrete clusters of STs. Within these clusters a central node can represent the most recent common ancestor of the groups of isolates which radiate from it.

The relationships between clusters of STs and traditional serovars are largely but not entirely consistent. For some serovars, such as *S. Typhi*, all members of a particular serovar are found within one cluster, and the cluster contains isolates of this serovar exclusively. In contrast *S. Typhimurium* isolates are positioned within multiple clusters, and the major cluster containing the central sequence type ST19 also contains isolates of other serovars. In some cases there is very little match between serotype and ST cluster, for example *S. Paratyphi B* isolates are located within multiple distinct clusters [80]. MLST is relatively amenable to high-throughput techniques; requiring PCR to amplify regions within housekeeping genes, followed by sequencing. Despite its greater efficiency and the more comprehensive information it provides, the change from established serological typing methods to MLST will require a significant shift in global thinking.

MLST provides information for an outline population structure, which can be further investigated with whole genome sequencing (WGS). In light of the advances in sequencing technology the application of WGS to study *Salmonella* is increasing. As mentioned earlier, sequence analysis of *S. Typhi* has shown that all isolates form a single tree radiating from a common ancestor, indicating a recently (several thousand years) expanded clonal population [96]. The global population of *S. Typhi* has been estimated to be relatively small; unusually most of the links between sequential haplotypes consist of single SNPs and recombination with other salmonellae is uncommon. 85 haplotypes currently describe the genetic variation within *S. Typhi* but new sequencing efforts are likely to expand this number. The population generally shows a weak evidence of strong selection, potentially linked to the existence of the asymptomatic carrier state [96, 97]. However recent work has identified that antibiotic usage is contributing to the global selection of a single multi-drug resistant haplotype, H58, frequently found with resistance to quinolones through mutations in *gyrA* [98]. WGS studies such as this one are important to identify emerging strains and investigate causes of selection.

A



B

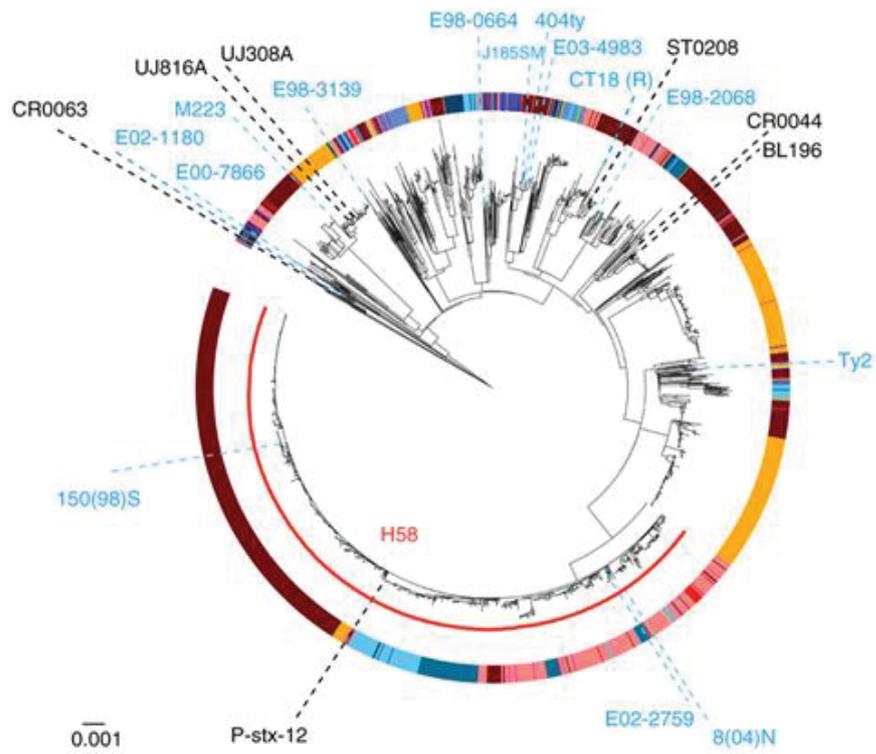


Figure 1.4. Phylogenetic relationships within serovars *S. Typhimurium* and *S. Typhi*. (A) Minimal spanning tree of *S. Typhimurium* and its serological variants based upon MLST for a collection of > 4,000 *Salmonella* isolates. Each circle represents one sequence type and segments represent individual isolates. Connecting lines indicate relationships between sequence types. White boxes indicate related clusters. The tree demonstrates that the serovar *S. Typhimurium* contains groups of relatively phylogenetically distant *Salmonella* and that isolates of other *Salmonella* serovars cluster with *S. Typhimurium*. (B) Rooted maximum-likelihood tree of > 1,800 isolates of *S. Typhi* inferred from 22,145 SNPs. The coloured ring indicates the geographical source of the isolate, and the red arc isolates of the haplotype H58. The tree demonstrates that the *S. Typhi* serovar is clonal. The H58 haplotype accounts for almost half of this global collection of isolates. A and B taken from [80] and [98] respectively.

1.3.5.2 The genome of *S. enterica*

As members of the *Enterobacteriaceae*, *Salmonella* genomes approach 5 Mb in size and contain sets of core genes encoding functions conserved throughout the family, such as intestinal persistence. The core genes are largely syntenic between individual members of the family, and are interspersed with species- and strain-specific regions, some encoding pathogenicity determinants. In addition to the chromosome many strains of *Salmonella* contain one or more plasmids, mobile genetic elements that typically encode virulence genes and genes which confer antibiotic resistance.

1.3.5.2.1 The genome of *S. Typhimurium*

The first *S. Typhimurium* to be sequenced completely was the common laboratory strain LT2, with a 4,857 kb chromosome and 94 kb virulence plasmid. The genome of LT2 serves to represent that of a promiscuous *Salmonella* gastroenteritis isolate. Analysis of the genome determined ~ 4,500 coding sequences, with just over 100 genes encoded on the plasmid. Comparisons were made between the genome of LT2, other *Salmonella* within subspecies I (*S. Typhi*, *S. Paratyphi* A, and *S. Paratyphi* B), and *Salmonella* within other subspecies. ~ 8% of genes in LT2 were found to have a homologue in at least one other member of subspecies I, but were not in *Salmonella* of other subspecies. As the only subspecies specialised to colonise warm blooded hosts, such genes may be important in conferring this adaptation [99].

1.3.5.2.2 The genome of *S. Typhi*

The complete sequencing of two isolates of *S. Typhi* facilitated a comparison between the genomes of the *S. Typhimurium* and *S. Typhi* serovars [100]. A major finding was the

extensive genome degradation which has occurred during the evolution of *S. Typhi*. Whilst the *S. Typhimurium* genome contains around 40 predicted pseudogenes, *S. Typhi* has ~ 200, corresponding to the inactivation of ~ 3 - 5% of all genes [100]. The nature of these pseudogenes has been investigated in order to understand how the functional loss in *S. Typhi* might relate to the characteristics of this serovar. 7 of the 12 fimbrial systems and other fimbrial-like genes are inactivated, and the loss of these structures important for directing attachment to surfaces such as host cells may contribute to the restriction of this serovar to human hosts. *ShdA* and *ratB*, genes, associated with intestinal persistence in *S. Typhimurium*, are also inactivated in *S. Typhi*. The loss of function might help to explain the lower numbers of *S. Typhi* in the gastrointestinal tract in the period following initial exposure, instead favouring rapid systemic dissemination [97].

Loss of gene function in the invasive serotype *S. Typhi* extends beyond this serovar to other salmonellae responsible for systemic disease, including *S. Paratyphi A*. The emerging sequence type of *S. Typhimurium* responsible for iNTS in sub-Saharan Africa, ST313, has accumulated inactivating mutations in some genes also inactivated in *S. Typhi*. In particular many metabolic genes are inactivated, for example Tartrate dehydratase (*ttdA*) is also inactivated or absent in *S. Typhi*, *S. Paratyphi A* and *S. Paratyphi B*. It has been suggested that differences in metabolic capacity might influence *Salmonella* fitness in different environments, and thereby impact both upon the niche within the body, and the host range of the bacteria [89]. Restriction of host range by the loss of gene capacity in an ancestor with a broad host range has been reported in other human-restricted pathogens such as *Mycobacterium leprae* and *Yersinia pestis* [101, 102].

As well as functional gene loss, during its evolution *S. Typhi* has gained genes absent in gastroenteritis-causing *Salmonella*. For example a major characteristic of *S. Typhi* is the expression of the Vi capsular polysaccharide, encoded in a genetic region designated SPI-7 containing many features associated with the horizontal acquisition of DNA. *S. Paratyphi C* and some isolates of *S. Dublin* have also acquired a SPI-7-like region [103]. The Vi capsule has been shown to modulate the recruitment of neutrophils to the gut, thought to reduce diarrhoea thereby adapting *S. Typhi* to a systemic lifestyle [104].

1.3.6 Virulence mechanisms of *S. Typhimurium*

While numerous *Salmonella* serotypes contribute to the burden of human disease, the majority of research has been carried out in just two: *S. Typhi* and *S. Typhimurium*. In specific experimental animal models *S. Typhimurium* behaves as a prototypical intestinal pathogen. Both serotypes are models for the study of host-pathogen interactions in the context of infection with intracellular pathogens, the cause of many major human diseases [105].

S. Typhimurium utilises a large assortment of virulence strategies in order to take advantage of the host environment, both in the intestinal lumen and intracellularly. Many of these highly effective strategies are shared by other *S. enterica* and in some instances by other bacterial pathogens, though the specific mechanisms behind their execution are often distinct. Chemotaxis towards and attachment to the host epithelial surface bring bacteria to a location which facilitates their entry into host cells [106]. Once inside virulence strategies are employed to avoid onslaught of the host defence mechanisms, and to control and replicate within the host cell.

The many individual pathogenic adaptations of *S. Typhimurium* described in the following section combine to implement a relatively simple, yet highly effective overall strategy for survival and transmission. *S. Typhimurium* virulence factors induce the host innate immune system to trigger rapid inflammation of the intestinal tissue. Dramatic changes in the environment within in the intestinal lumen result, and alter the ability of the bacterial species residing here to survive [107]. For many key beneficial species of the microbiota these changes convert the intestinal lumen from a hospitable to a hostile environment, and they fail to survive in the inflammatory environment. Conversely the pathogenic adaptations of *S. Typhimurium* enable it to thrive in the altered environment and its population ‘blooms’ [108]. Neutrophil chemotaxis to the gut and extravasation induced by inflammation lead to an increase in vascular permeability and fluid build-up in the tissue, and the passage of neutrophils and dendritic cells into the intestinal lumen damages the integrity of the epithelial barrier. The combination of these effects contributes to diarrhoea, which aids the faecal-oral transmission of *S. Typhimurium* to fresh hosts [109]. Induction of inflammation makes an important contribution to lowering colonisation resistance in other intestinal infections, such as *E. coli*, also [110].

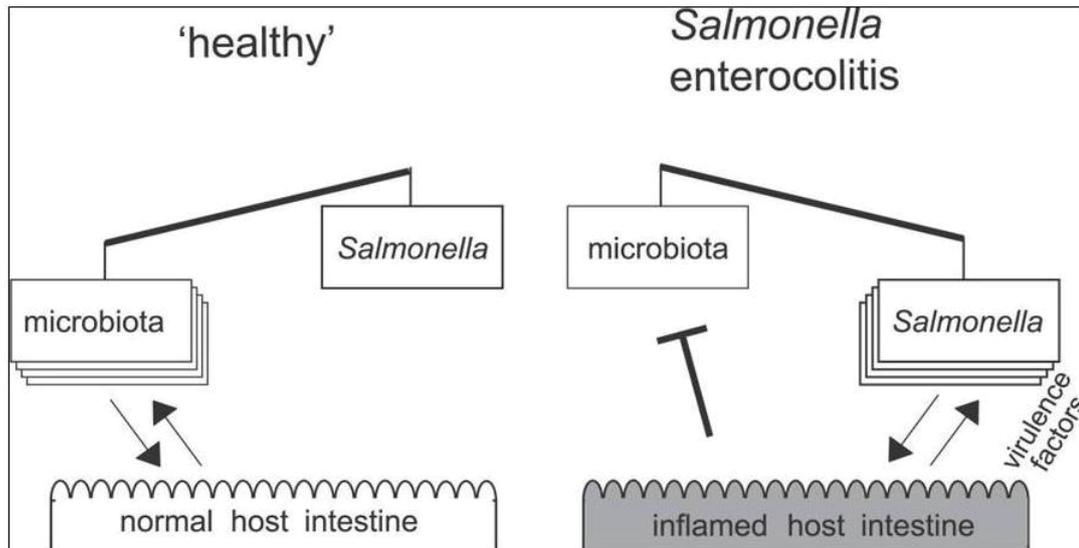


Figure 1.5. Interactions between *S. Typhimurium*, the microbiota and the intestinal mucosa. In the healthy gut mutually beneficial interactions between the microbiota and the host support the microbiota, and a diverse microbiota holds back *Salmonella* through colonisation resistance. *S. Typhimurium* targets the host with virulence factors resulting in inflammation. Inflammation alters the gut environment to favour *S. Typhimurium* growth and triggers host response pathways which directly inhibit the microbiota. Modified from [108].

1.3.6.1 Type 3 secretion systems (T3SS)

Key virulence-associated determinants of *Salmonella* are two horizontally acquired regions absent in related commensal *E. coli* species, SPI-1 and SPI-2 (see section 1.3.1) [111]. T3SS are complex bacterial membrane-associated protein export systems assembled from > 20 proteins. A needle-like structure spanning the envelope delivers specific bacterial proteins, known as 'effectors', into the cytosol of a host cell [112]. The SPI-1 and SPI-2 T3SS deliver different collections of effectors and function consecutively during bacterial interaction with a host cell. SPI-1 is activated in extracellular bacteria by stimuli present in the lumen whilst SPI-2 is actively transcribed when the bacteria are within a host cell, although there may be temporal overlap in the activities and differences between serotypes [113].

Salmonella SPI-1 encodes a T3SS pivotal for the directed invasion of non-phagocytic host cells [114]. At least 13 effectors encoded both within and separately to SPI-1 are translocated into the host cytoplasm by this system. Although the catalogue of functions these effectors perform *in vivo* is incomplete, their key role lies in remodelling of the cytoskeleton. Host actin binding proteins maintain a fine balance between monomeric and polymeric

filamentous forms of actin to control the structure of the cytoskeleton, and carry out specific functions such as vesicular transport. By interfering with signalling pathways controlling this dynamic process, and by directly binding actin, *Salmonella* effectors hijack and remodel the cytoskeleton [115, 116]. For example, actin binding proteins SipA and SipC decrease the critical concentration needed for actin polymerisation and nucleate and bundle actin [117, 118]. It is thought that these processes contribute to drive the growth of membrane ruffles and filopodia; protrusions of the host cell membrane which allow the cell to engulf and internalise bacteria [119]. Many more of the effectors transported by the SPI-1 T3SS contribute to modify the cytoskeleton for uptake and actively recover the normal state after internalisation. Co-ordination of these opposing processes requires strict temporal control [120].

The T3SS encoded by SPI-2 is essential for the intracellular survival and replication of *Salmonella* following uptake by phagocytes such as macrophages and dendritic cells, and after invasion of epithelial cells [121, 122]. In the absence of adaptations for intracellular survival, phagosome-lysosome fusion occurs after bacterial uptake and the compartment becomes unfit for survival. The vacuole is acidified, and antimicrobial peptides and hydrolytic enzymes are released. Under such conditions bacteria can be digested within 30 minutes of fusion [123]. Phagocytosis of *Salmonella* proceeds differently; as a result of T3SS effectors both the uptake mechanism and the morphology of the phagosome formed are unconventional. Uptake occurs by a process likened to macropinocytosis, creating a large compartment referred to as the ‘spacious phagosome’. This contrasts with the close-fitting phagosomes generated by the zipper-like uptake mechanism for other bacteria [124, 125]. Effector proteins delivered to the host cytoplasm by the SPI-2 T3SS act to change the interaction of the *Salmonella*-containing compartment with the endosomal system, avoiding the fusion of lysosomes and creating a specialised ‘*Salmonella*-containing vacuole’ (SCV) [122, 126]. SPI-2 effector proteins act to control the cellular location of the SCV and vesicle fusion through interference with microtubule-dependent trafficking events [127]. *Salmonella* exit from infected host cells has been shown to require SPI-2 effectors suggesting this T3SS is important for bacteria to spread to new infection foci [128]. They are implicated in a wide variety of processes, for example disruption of dendritic cell antigen presenting, though the biological functions of many remain unclear [129].

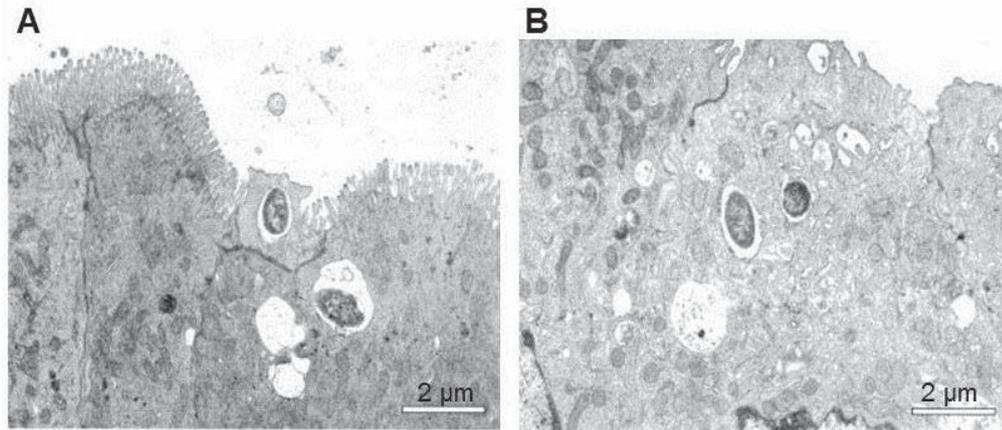


Figure 1.6. Epithelial cell invasion by *S. Typhimurium*. Transmission electron micrographs showing invasion of epithelial cells in the calf intestinal mucosa by *S. Typhimurium*. (A) Membrane ruffling is induced at the apical surface of the epithelial cell during invasion. (B) *S. Typhimurium* within vacuoles inside an M cell. Modified from [83].

The SPI-1 and SPI-2 T3SS can trigger pathogen-favourable host inflammatory responses by independent mechanisms. The host signalling pathways which mediate these responses are covered in detail in section 1.5.1. Briefly, SPI-1-induced colitis occurs independently of a key host signalling adaptor molecule, Myd88. SPI-1 pathways involve direct injection of proinflammatory molecules into epithelial cells, facilitating *Salmonella* invasion, which exposes receptors in the epithelial cell cytosol to PAMPs. The SPI-2 T3SS-mediated mechanism of triggering inflammation involves enhanced intracellular bacterial replication, increasing the load of bacteria to which the innate immune system is exposed [126, 130].

1.3.6.2 Nutrition and energy sources

A significant selective advantage of *S. Typhimurium* residing in the inflamed intestine is thought to arise from the ability to utilise the limited available nutrients [131]. Nutrients which support the growth of the resident microbiota are depleted by loss of luminal content in diarrhoea. *S. Typhimurium* can exploit chemotaxis up D-galactose gradients of the mucus layer and attachment to terminal mucus carbohydrates to position it to make use of mucus as an energy source. The absence of a selective advantage for flagella when alternative nutrient sources are available implicates chemotaxis as an adaptation to the inflamed intestinal environment [132].

Further, *S. Typhimurium* produces energy more efficiently than competing microbes. Most species of the microbiota are strictly anaerobic and rely on fermentation to supply energy for survival and growth. Conversely *S. Typhimurium* is able to support growth by anaerobic respiration using a molecule abundant in the inflamed gut, tetrathionate, as a terminal electron acceptor. Tetrathionate ($S_4O_6^{2-}$) is produced from oxidation of thiosulphate ($S_2O_3^{2-}$) on release of reactive oxygen species (ROS) by infiltrating neutrophils in inflammation [133].

1.3.6.3 Resistance to host antimicrobial peptides

S. Typhimurium is equipped with protective defences against the antimicrobial peptides produced constitutively by epithelial cells, and induced in the inflammatory response. Many bacterial species secrete a low molecular weight iron chelating molecule (siderophore), enterobactin, to enable acquisition of the essential element iron. The antimicrobial peptide lipocalin-2, released in the lumen in response to IL22 signalling, binds to enterobactin to inhibit growth of species which rely on this siderophore [134]. By producing a glycosylated derivative of enterobactin not bound by lipocalin-2, salmochelin, *S. Typhimurium* ensures a plentiful supply of iron [135]. In another strategy for protection against antimicrobial peptides *S. Typhimurium* modifies the lipid A moiety of its LPS, avoiding cationic peptides which target bacterial surfaces through electrostatic interactions. To increase resistance to the antimicrobial peptide polymixin the lipid A of LPS is modified by addition of aminoarabinose [136]. The response requires activation of a crucial two-component signalling pathway called PhoQ-PhoP by antimicrobial peptides, mediating the expression or repression of over 40 genes. In addition to modification of LPS these transcriptional changes generate a variety of responses including survival inside macrophages and altered antigen presentation [137, 138].

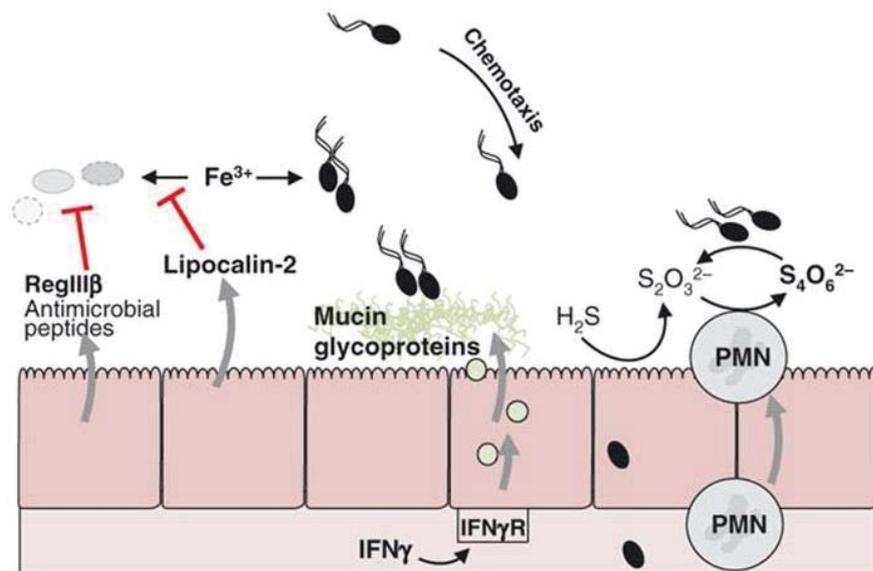


Figure 1.7. Adaptations of *S. Typhimurium* to the inflamed intestinal mucosa. Black cells represent *S. Typhimurium* and grey cells species of the microbiota. *S. Typhimurium* is resistant to harmful effects of many host antimicrobial peptides, for example it avoids shortages in iron induced by the host peptide lipocalin-2 by production of an iron chelating-protein absent in species of the microbiota. It uses chemotaxis to move toward the glycocalyx to take advantage of mucus glycoproteins as a nutrient source. Tetrathionate produced by infiltrating neutrophils in the inflammatory environment is adopted by *S. Typhimurium* as an electron acceptor giving a metabolic advantage over species of the microbiota which rely on fermentation. Taken from [139]

1.3.7 Animal models of human *S. Typhimurium* infection

Animal models are invaluable for studying the pathogenesis of *Salmonella* infection. Cell-based studies have helped identify many virulence factors and elucidate their mechanism of action but the relevance of these factors in the progression of a human infection is challenging to study in such a simplified system. Unlike cell systems, animal models of *S. Typhimurium* infection replicate aspects of the complex network of interactions occurring in human infection between the gut tissue, immune system, microbiota and pathogen. As discussed below calves are a model for human NTS infection, but due to considerations such as cost are used to a lesser extent than murine models. As humans and some species of primate are the only organisms naturally susceptible to *S. Typhi* this presents a barrier to the study of this serotype *in vivo*. However in mice *S. Typhimurium* infection gives rise to typhoid-like symptoms, and this mouse typhoid model has been studied extensively to gain understanding of typhoid fever pathogenesis. A primate model of *Salmonella* in rhesus macaques infected with Simian Immunodeficiency Virus (SIV) has been used to study the interaction between HIV and *Salmonella* [90].

Murine models of infection are attractive for several reasons, including relatively low costs of housing animals and ease of handling compared to larger organisms, the availability of a large number of genetically modified inbred lines, and well established techniques for genetic and immunological manipulation. *S. Typhimurium* is a natural pathogen of rodents and was discovered as the causative agent of murine typhoid in 1892. In contrast to the self-limiting gastroenteritis of human *S. Typhimurium* infection, in mice bacteria fail to invade epithelial cells, and *Salmonella* pass through the mucosa via M cells without triggering extensive inflammation. Consequently, translocated bacteria disseminate systemically causing a typhoid-like infection in susceptible mouse strains. Upon oral delivery of *S. Typhimurium* mice develop signs of infection at 4 - 8 days post-infection (PI) including fever, enlarged Peyer's patches and diffuse enteritis, similar to human typhoid patients. Bacterial replication in the spleen and liver results in enlargement of these organs, and formation of granulomas. Ultimately infected mice die, likely due to the liver lesions caused by lipid A-elicited cytokine and inducible nitric oxide synthase (iNOS) responses [140].

The failure of *S. Typhimurium* to thrive in the mouse gut indicates the microbiota here is more effective than that of humans in mediating colonisation resistance against invading *Salmonella*. However pre-treatment of mice with antibiotic prior to infection has proved successful in depleting the microbiota so that *Salmonella* may colonise the intestinal niche and trigger inflammation, creating a mouse model of *Salmonella* gastroenteritis. The pathological and histological characteristics of infection following antibiotic treatment of mice closely mirror infection in humans and calves.

Of importance in the study of *Salmonella* infection in mice is the differing susceptibility of mouse strains to this bacterium. Strains are divided into distinct groups of resistant, e.g. 129SvEv, and susceptible strains such as C57BL/6 and BALB/c. Following antibiotic pre-treatment, resistant and susceptible strains both develop acute intestinal inflammation; however while systemic dissemination in susceptible strains leads to death, resistant strains develop chronic colitis including crypt abscesses, ulceration, overshooting regeneration of the epithelium, and also inflammation of the gall duct epithelium (cholangitis) [141]. In the typhoid model of infection susceptible strains succumb to infection as described earlier in this section, whereas resistant strains control infection to become chronic carriers with *Salmonella* persisting in mLN over a year after infection [142]. A gene identified as responsible for the difference in susceptibility to *Salmonella*, and also other intracellular pathogens is Natural Resistance-Associated Macrophage Protein 1 (*Nramp1*), also called

Slc11a1 [143, 144]. NRAMP1 is produced selectively in macrophages and monocytes and targets the phagosome membrane, altering the environment within to assist in control of microbes here [145].

1.3.7.1 Bovine models

Cattle are natural hosts of *S. Typhimurium* and they can display clinical and histological manifestations strikingly similar to those of infection in humans. The outcome of infection depends on the dose of bacteria; an oral dose of 10^4 - 10^7 colony forming units (CFU) leading to transient diarrhoea followed by recovery, and a dose of $> 10^8$ resulting in lethality, though as in humans severity of infection reduces with age in early life. Clinical symptoms appearing within 12 - 48 h of infection include diarrhoea leading to dehydration and fever. Inflammation of the ileum is severe, with polymorphonuclear leukocyte (PMN) influx into the inflamed mucosa resulting in an initial drop in circulating immune cells and necrosis of the upper mucosa. Calf ligated ileal loops inoculated with *Salmonella* have provided a model to study the early stages of infection. With oral dosing, tissue along the intestine becomes occupied by *Salmonella* with different timings and densities of bacteria, whereas inoculation of ligated ileal loops resolves difficulties for studying temporal aspects of infection [146, 147]. Despite the closeness of the disease caused by *S. Typhimurium* in calves and humans, the costs of bovine models and lack of genetic homogeneity prevent their extensive use.

1.3.7.2 The streptomycin mouse model

Following the delivery of a single 20 mg dose of the antibiotic streptomycin, the natural colonisation resistance against *S. Typhimurium* afforded by the microbiota is transiently lost. An estimated > 10 -fold reduction in the density and complexity of the microbiota has been observed [108]. Recovery of protection takes a few days and may be accelerated by co-housing with untreated mice [148]. Though antibiotic delivery has reported effects on intestinal tissue homeostasis beyond depletion of the microbiota, it is thought that these effects are of minor if any significance, relative to the effect on colonisation resistance.

Oral delivery of *S. Typhimurium* bacteria as few as 100 results in high levels of colonisation; 10^8 - 10^{10} CFU/g intestinal tissue, though larger doses are used experimentally. Inflammation occurs rapidly; within 6 - 8 h the mucosa is invaded by bacteria and pronounced

mucosal inflammation can be detected. Inflammation is localised to the proximal colon and in particular the caecum, the blind-ended sac located at the junction between the ileum and the colon which in herbivorous animals plays an important role in the digestion of their cellulose-rich diet. In humans the small intestine is generally regarded as the worst affected region and is the predominant site of inflammation in calves also. However in mice small intestine is rarely inflamed. Despite the difference in the tissue affected the histopathology of the inflammation is comparable [149]. Similarly studies of *Salmonella* virulence mechanisms suggest that similar factors are required for the colonisation of the mouse caecum and human/bovine small intestine.

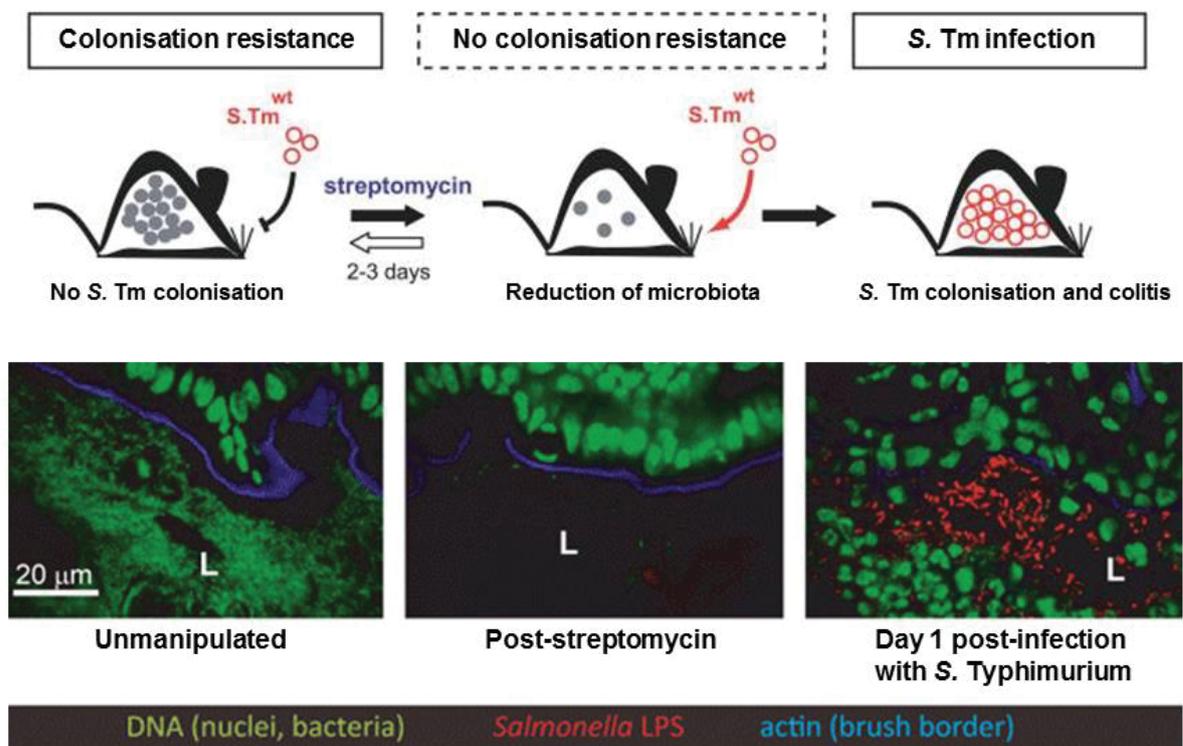


Figure 1.8. Antibiotic treatment in the streptomycin mouse model overcomes colonisation resistance. Upper - Schematic showing the relationship between the microbiota and colonisation of the mouse intestinal tract by *S. Typhimurium*. A substantial reduction in the microbiota occurs within 24 h of an oral dose of streptomycin enabling orally delivered *S. Typhimurium* to take hold in the caecum and colon. Without infection recovery of the microbiota is sufficient 2 - 3 days after streptomycin treatment to restore colonisation resistance. Lower - Confocal microscopy images of fluorescently stained caecum from an untreated mouse, 24 h after streptomycin treatment, and day 1 PI after streptomycin treatment. ‘L’ indicates the gut lumen. Luminal bacteria are dramatically reduced following streptomycin treatment. Within 24 h of infection *S. Typhimurium* is clearly visible in the lumen. Modified from [108] and [139].

Invasion of the mucosa occurs through multiple routes. As in the murine typhoid model *Salmonella* invade the intestinal patches of the GALT, but in the streptomycin model bacteria also invade the absorptive intestinal epithelial cells. Acute intestinal inflammation occurs regardless of the absence of the patch structures, ILFs and lymph nodes in mice lacking a receptor for immune system development factors called lymphotoxins (*LTβR*^{-/-}), indicating that the induction of inflammation occurs through the interactions of *Salmonella* with the enterocytes [149].

A major advantage of the model is that it is highly robust; mice are consistently colonised with high numbers of bacteria and quickly develop acute intestinal inflammation. Regardless of the numbers of bacteria in the inoculum the final density and the pathological characteristics are similar.

The pronounced inflammation of the caecum and proximal colon has been described in detail. In streptomycin-treated infected mice the caecum is shrunken, pale and pus-filled, and the proximal colon pale and swollen. At the microscopic level changes in the caecum include pronounced oedema in the submucosa and oedematous changes in the lamina propria. Crypts are elongated and display disrupted architecture, there are fewer goblet cells, and the epithelium is eroded and/or ulcerated. There is pronounced PMN infiltration of the submucosa, lamina propria and epithelial layer, and PMN migration into the intestinal lumen. Changes in the colon are similar but less severe, as described in [149].

In most aspects *S. Typhimurium* infection in streptomycin pre-treated mice closely mirrors calves and humans however a notable difference is the absence of diarrhoea in mice. Murine infection results in increased stool water content and mucus, but such effects are incomparable with the extensive diarrhoea which defines human NTS. The difference is thought to result from the adaptation of the mouse gastrointestinal tract to maximise fluid absorption. At the microscopic level the changes that occur in the infected mucosa are similar in calves and in the streptomycin mouse model.

The use of genetically modified mice and bacteria in the streptomycin model has helped to uncover the pathways by which *S. Typhimurium* elicits inflammation in the intestinal tract. Two independent inflammation triggering pathways, one of which requires the T3SS of SPI-1 and the other requiring SPI-2 contribute in parallel, and through infections with *S. Typhimurium* derivatives harbouring mutations in these loci, the mechanisms have

been disentangled. Bacteria which are deficient in SipA, SopE and SopE2 effector proteins delivered into host cells by the T3SS of SPI-1 display attenuated inflammation in the streptomycin mouse model [150]. The induction of inflammation by active injection of inflammatory effectors is referred to as the ‘classical pathway’. *Salmonella* mutants lacking a functional SPI-1 T3SS are able to induce inflammation via the ‘alternative pathway’ [149]. Penetration of the epithelium by routes other than through invasion of epithelial cells, such as transport via dendritic cell extensions and in M cells, bypasses the need for SPI-1 activity and the SPI-2-dependent occupation of macrophages leads to inflammation. The exact mechanism is not understood but it is thought detection of high bacterial loads through MyD88 signalling might be responsible.

Knockout and other genetically manipulated mice have produced insight into host response pathways, both protective and detrimental. The cytokine IFN γ induced early in response to infection was shown to be required for full inflammation and control of bacterial loads through knockout studies [151]. Infection of *IL22*^{-/-} mice showed that despite the important role of IL22 signalling for maintenance of the mucosal barrier, IL22 increases the severity of infection with *S. Typhimurium* by induction of responses at the epithelial surface and in the lumen which favour survival of *Salmonella* over commensals [152].

1.4 Murine models of intestinal inflammation

S. Typhimurium infection in streptomycin pre-treated mice is one of many mouse models of intestinal inflammation. While diseases involving intestinal inflammation have disease-specific mechanisms they also share common pathways, and findings from one disease model can generate greater understanding of the pathogenesis of multiple conditions. Two clinically distinct diseases characterised by intestinal inflammation, ulcerative colitis and Crohn’s disease (known collectively as IBD), are relatively common autoimmune disorders [153]. Symptoms of IBD include diarrhoea, abdominal pain, rectal bleeding and weight loss. A combination of environmental and genetic factors give rise to disease and the causes are not well understood. An involvement of bacteria has been demonstrated but there are multiple theories as to how, for example whether dysbiosis is a cause or a consequence of IBD, and whether intestinal inflammation induced by pathogens plays a role [73]. Studying systems such as the streptomycin mouse model can help to uncover mechanisms by which bacteria induce inflammation that may be relevant to IBD. Enteric bacterial pathogens share many

common virulence strategies and through the study of multiple infection models these conserved approaches can be understood and used to identify targets for drugs and vaccines.

Gut inflammation models include chemical, genetic, immunological and bacterial systems, each suited to the study of different aspects or stages of inflammation. As mentioned earlier mutant mice are important for the study of the innate and adaptive immune system in disease development. Immunological models which involve the adoptive transfer of specific immune cell subsets also help to determine the contributions of specific cell types. The following two examples are chosen to demonstrate the utility of chemical models for study of inflammatory processes, and to introduce a model of bacterial diarrhoeal disease with similarities to *S. Typhimurium* infection.

1.4.1 *Citrobacter rodentium*

C. rodentium is a Gram-negative bacterial pathogen of mice which colonises the intestine through specialised attaching and effacing (A/E) lesions. Closely related human A/E pathogens Enteropathogenic and Enterohaemorrhagic *E. coli* (EPEC and EHEC) colonise mice poorly. Consequently *C. rodentium* provides a useful model to study pathogenic mechanisms specific to A/E bacteria, and also those widespread amongst intestinal bacteria. Oral delivery of *C. rodentium* to mice results in colonisation 3 - 4 weeks in duration, initially occurring in the caecal patch and spreading to the distal colon. Concomitant with the peak of colonisation at 5 - 14 days PI is acute intestinal inflammation. *C. rodentium*-induced colitis is mild relative to *S. Typhimurium*-induced colitis in the streptomycin pre-treatment model, and characterised by crypt hyperplasia, loss of goblet cells and infiltration of mononuclear cells. In most strains of mice bacteria are cleared by 21 - 28 days PI, although very young mice are more susceptible and may die [154].

In contrast to *Salmonella*, *C. rodentium* is largely non-invasive, residing in the lumen attached to the surface of epithelial cells. Despite their different niches they share common virulence strategies. In common with the SPI-1 and SPI-2 T3SS of *S. enterica* the locus of enterocyte effacement (LEE) in *C. rodentium* encodes a T3SS for translocation of effectors, encoded both within the LEE genomic region and separately on prophages and elsewhere in the genome. The varied functions of the *C. rodentium* effector proteins have been explored through systematic mutagenesis and include permeabilisation of the mitochondrial membranes which triggers induction of cell death (EspF), and disruption of epithelial barrier

function (Map). In common with the SPI-1 T3SS effectors, which commandeer the host cell cytoskeleton to induce ruffling and ultimately bacterial uptake, several *C. rodentium* effectors interact with the host cytoskeleton. These direct cytoskeletal rearrangements leading to the formation of pedestal-like protruding structures in the host membrane with localised destruction of the microvilli [155].

The importance of the innate immune system in overcoming *C. rodentium* has been investigated. Epithelial cell secreted defensins may play a role, as has been indicated by increased susceptibility in *IL12*^{-/-} mice which show reduced expression of β -defensin 3. iNOS is increased in epithelial cells during infection although this appears to have limited functional significance as *iNOS*^{-/-} mice clear bacteria normally [156]. Mice lacking T and B cells are chronically infected, demonstrating adaptive immunity is essential for efficient *C. rodentium* clearance [157]. The restoration of protection in CD4⁺ T cell deficient mice by transfer of serum IgG and IgM from previously recovered mice indicates that the T helper-dependent serum antibody response is key [158]. The elimination of infection appears to require IgG in particular over IgM or IgA, which may access the gut lumen due to the increased epithelial permeability in the infected gut [159].

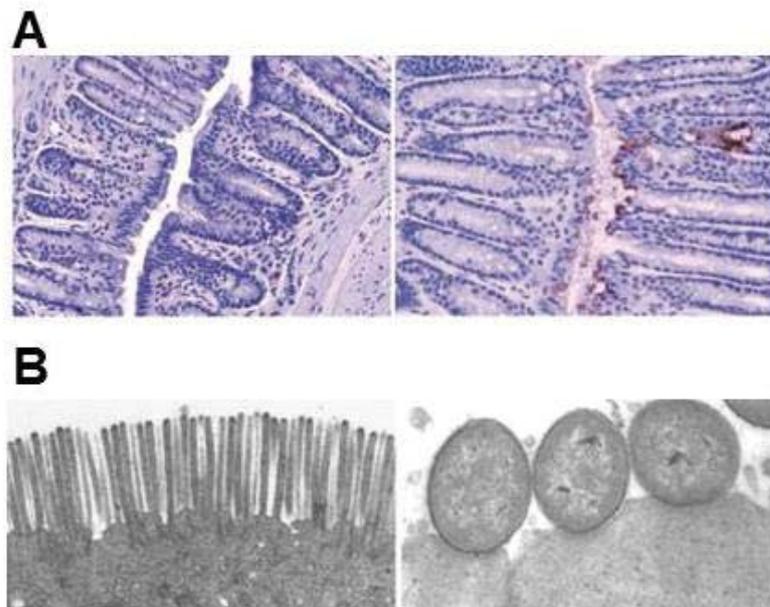


Figure 1.9. Microscopic changes in the mouse colon during infection with *C. rodentium*. (A) Hematoxylin and eosin staining combined with staining for the *C. rodentium* adhesin intimin in colonic tissue from an uninfected mouse (left) and a mouse colonised with *C. rodentium* (right). In infected tissue crypt hyperplasia transforms the mucosal architecture. (B) Transmission electron micrographs showing the normal brush border microvilli (left) and erosion of microvilli in A/E lesions associated with *C. rodentium* (right). Taken from [160].

1.4.2 Dextran sodium sulphate (DSS)-induced colitis

One model of chemically-induced inflammation involves the addition of the sulphated polysaccharide dextran sodium sulphate to drinking water. DSS is highly variable in molecular weight, ranging from 5 - 1400 kDa, and inflammatory responses induced vary according to size of the DSS and a range of other factors including concentration and frequency of administration, as well as mouse strain, intestinal flora composition and animal stress. DSS is toxic to epithelial cells and triggers cell death allowing DSS to penetrate the mucosal membrane.

The similarity between the clinical and histopathological features of DSS colitis and human IBD make this a valuable model. Addition of 2 - 5% DSS to drinking water for a period of around a week results in acute colitis and can involve weight loss, diarrhoea, blood in stools, anaemia, and eventually death. Histological changes include the depletion of the mucus barrier, and infiltration of neutrophils into the lamina propria and submucosa with their further migration to the lumen causing crypt abscesses. With continuous treatment of low doses mice develop chronic colitis within a few weeks. The changes resulting from longer term exposure include infiltration of mononuclear leukocytes, extensive disruption of crypt architecture, and widening of the gap between the bases of the crypts and the muscularis mucosa [161].

The microbiota has been shown to play an important role in the development of DSS colitis, as with human IBD. The delivery of antibiotics in acute DSS colitis results in improved histological parameters in treated mice compared with untreated controls showing that bacteria or their products are important to trigger inflammation [162]. Similarly germ-free mice fail to display the changes in intestinal morphology associated with DSS treatment in conventionally housed mice [163]. Microbes are not required for the destruction of the mucosa, however the loss of tight junction proteins and barrier integrity allows the luminal antigens and microorganisms entry to the mucosa triggering an overwhelming immune response. As early as one day after the commencement of DSS treatment inflammatory cytokines including TNF α , IL1 β , IFN γ , IL10 and IL12 are induced, increasing further with continued treatment. IRF1 is a transcription factor downstream of TNF α and IFN γ signalling which induces major immune pathways including interferon signalling and iNOS production. *IRF1*^{-/-} mice show increased susceptibility to DSS colitis indicating these responses are protective [164]. Similarly *TLR4*^{-/-} and *MyD88*^{-/-} mice display increased bacterial translocation

to lymph nodes and decreased epithelial proliferation showing the importance of TLR signalling for protection [165, 166].

1.5 Immune response to *S. Typhimurium* infection in the intestinal mucosa

Two arms of an immune response combine to confer protection against infectious microorganisms; an innate component is triggered initially and mediates rapid protection via antigen-independent mechanisms, and an adaptive response requiring several days to become active targets specific antigens. The innate component serves to highlight potential danger via immune receptors and downstream signalling pathways, resulting in production of cytokines. Cytokines mediate a plethora of functions acting to keep the threat under control, including recruitment and activation of phagocytes. The adaptive arm of the response mediates targeted killing through the action of specialised cells, and production of antibodies which bind and opsonize microbes to facilitate clearance by phagocytes. Upon re-exposure, the immunological memory conferred by the adaptive system controls the pathogen more rapidly and typically prevents re-infection.

1.5.1 Innate immune surveillance and signalling

The earliest interaction of the immune system with a microbe occurs via PRR, specialised receptors for interaction with conserved microbial molecules and structures (PAMPs). Several classes of PRR exist and together they possess a large number of binding specificities, across multiple locations. Three of the major families are the Toll-like receptors (TLRs), nucleotide-binding and oligomerization domain-like receptors (NLRs) and complement. Different TLRs are located at distinct sites within the cell membrane and membranes of intracellular compartments, both in epithelial cells and sentinel cells directly below, the compartmentalisation allowing a close watch over the microbes which invade the epithelium. NLRs are located in the cytosol where they typically detect the presence of virulence-associated factors such as T3SS components. In contrast to the two cellular receptors, complement is a humoral component of the PRR repertoire and is activated by cleavage after exposure to bacteria which have crossed the epithelial barrier. It has been argued that a distinction exists between pattern recognition and ‘pathogen-induced process recognition’, the former simply alerting the immune system to the presence of a microbe while the latter provides some indication of the pathogenic potential or level of threat [167].

Downstream of receptor activation, signalling pathways coordinate an immune response to the assault. Collections of TLRs recruit common adaptor proteins to signal by shared pathways to induce a similar response. For example TLRs 1, 2, 4 and 5 recruit the common adapter MyD88 and activate mitogen-activated protein (MAP) kinase signal transduction pathways. These activate transcription factor activator protein-1 (AP-1) and nuclear factor kappa-light chain enhancer of activated B cells (NF κ B), which induce a collection of proinflammatory cytokines including IL23 α , TNF α , IL12 α , IL1 β and IL18 in mononuclear and epithelial cells. Typically during infection a relatively small fraction of host cells contain bacteria, and therefore amplification pathways are required in order to generate a signal strong enough to induce effector functions. Two cytokines of the primary proinflammatory signal, IL18 and IL23 initiate major amplification pathways. T cells are important for amplification by the production of cytokines such as IFN γ , IL17 and IL22 [168]. This combined collection of cytokines mediates major antibacterial responses: macrophage activation, neutrophil recruitment, and epithelial release of antimicrobials.

Activation of C3 results in the production of the anaphylotoxins C3a and C5a, potent inducers of inflammation. These stimulate release of histamine and TNF α from basophils and mast cells, resulting in vasodilation and recruitment of neutrophils [169]. Reactive thioester groups of C3b produced in the proteolytic cleavage of C3 form esters with hydroxyl groups of bacterial surface carbohydrates, acting to opsonize bacteria for phagocytosis by neutrophils, macrophages and dendritic cells. Finally, activation of the alternative pathway for complement activation leads to the formation of the membrane attack complex (MAC), which kills gram negative bacteria by the formation of pores in the bacterial outer membrane.

S. Typhimurium activates a multitude of PRRs, including both humoral proteins and cellular receptors. The O-antigen of its LPS activates complement component 3 (C3) triggering the alternative pathway of complement activation. The lipid A moiety of LPS activates TLR4 in host cell vesicles, TLRs 1 & 2 are activated by the biofilm protein CsgA and TLR5 on the basolateral membrane of intestinal epithelial cells is activated by the major protein subunit of flagellin, FliC [170, 171]. Each of these TLR recruit Myd88 and activate signalling to induce a range of proinflammatory genes as earlier described. *S. Typhimurium* is responsible for triggering pathogen-induced process recognition as a result of T3SS-dependent delivery of proteins into the host cytosol. The needle complex protein PrgJ and the flagellar component flagellin are recognised by NOD-like receptor NLRC4. Further, cell wall fragments released into the cytosol as a result of host cell invasion are detected by the

receptors NOD1 and NOD2, which via protein kinase receptor interacting protein-2 (RIP2), mediate NF κ B activation.

1.5.2 Three major innate effector responses to *S. Typhimurium*

1.5.2.1 Macrophage activation

The intracellular niche preferred by *S. Typhimurium* for survival and replication is the SCV inside tissue resident mononuclear phagocytes within the reticuloendothelial system. The host defends against *Salmonella* residing here by enhancing the ability of macrophages to kill bacteria within. In response to activation macrophages use iNOS to generate reactive nitrogen species for bacterial killing. Macrophage activation is enhanced by stimulation with the cytokine IFN γ , produced by CD $\alpha\beta$ T cells, CD4⁺ $\alpha\beta$ memory type -1 T helper cells (Th1), and natural killer (NK) cells in response to IL12 and IL18. The importance of the IFN γ axis in maintaining barrier function in the mucosa is illustrated by increased susceptibility to disseminated infections with NTS serotypes in humans with defective components of the IFN γ axis [172].

In order to stay ahead of the increased killing action of the macrophage, *Salmonella* induces pathways which lead to host cell death and release of bacteria to infect new host cells [173]. After activation of the NOD-like receptor NLRC4 by the T3SS needle complex components PrgJ and flagellin, NLRC4 forms a complex with caspase-1 and ASC resulting in inflammasome formation and macrophage death by an inflammatory process called pyroptosis. Alternatively a T3SS-independent pathway via TRIF and receptor interacting proteins 1 & 3 (RIP1 & 3) induces programmed necrosis [174, 175].

Later in the course of the infection *Salmonella*-specific antibodies opsonize the bacteria to further assist macrophage killing. T cell-dependent processes contribute also [176].

1.5.2.2 Neutrophil recruitment

Compared with the macrophage, a second innate phagocytic cell type is a highly effective destroyer of *Salmonella*. Circulating neutrophils are recruited to the intestinal mucosa by chemoattractants produced by epithelial cells. Despite the largely intracellular lifestyle of *Salmonella* there exist small windows of opportunity for neutrophil attack in the short extracellular transition from an infected cell to a new host. The cytokines IL23 and

IL1 β , produced in the initial inflammatory response, act upon Th17 cells, $\gamma\delta$ T cells, and NKT cells to induce the production of IL17 [177]. Stimulation by this cytokine along with IL1 β and TNF α induces epithelial cell production of neutrophil-recruiting CXC chemokines. The anaphylotoxin C5a produced during the complement cascade also functions to recruit neutrophils.

Lessons from infections in which neutrophil function is somehow absent or impaired demonstrate the importance of this cell type in resolving infection. Neutropenic individuals fail to control *S. Typhimurium* in the gastrointestinal mucosa and develop life-threatening bacteremia [178]. In *S. Typhi* infection the presence of the capsular polysaccharide Vi has been shown to negatively impact upon C3b fixation and neutrophil killing by opsonophagocytosis. Greater resistance to killing by neutrophils likely aids the ability of *S. Typhi* to cross the intestinal barrier of the GI tract and disseminate systemically. Though essential, neutrophil recruitment comes with a cost; extravasation of neutrophils and passage into the gut lumen causes tissue damage [179].

1.5.2.3 Antimicrobial peptides

IL23 produced by *Salmonella*-infected macrophages and dendritic cells induces production of IL22 by CD4⁺ $\alpha\beta$ memory type-17 cells (Th17 cells), $\gamma\delta$ T cells, and NKT cells. IL22 acts upon intestinal epithelial cells to stimulate the production of antimicrobial molecules which are secreted into the intestinal lumen to attack the bacteria. These molecules include the enterobactin-binding protein lipocalin-2 and the bactericidal C-type lectin, regenerating islet-derived 3 gamma (Reg3 γ) [180].

Depletion of the cells which initiate signalling to induce antimicrobial peptide production results in a compromised epithelial barrier as demonstrated in a mouse model and in SIV infection in rhesus macaques [90, 181].

Unlike the activation of macrophages and recruitment of neutrophils which target bacteria following their departure from the lumen into mucosal tissue, the release of antimicrobial peptides acts in the lumen where many beneficial bacterial species reside. Consequently these can have severe impacts on both enteric pathogens and normal gut bacteria, though as described in (section 1.3.6.3) pathogenic bacteria including *S. Typhimurium* have evolved mechanisms to overcome the effects of the antimicrobial peptides.

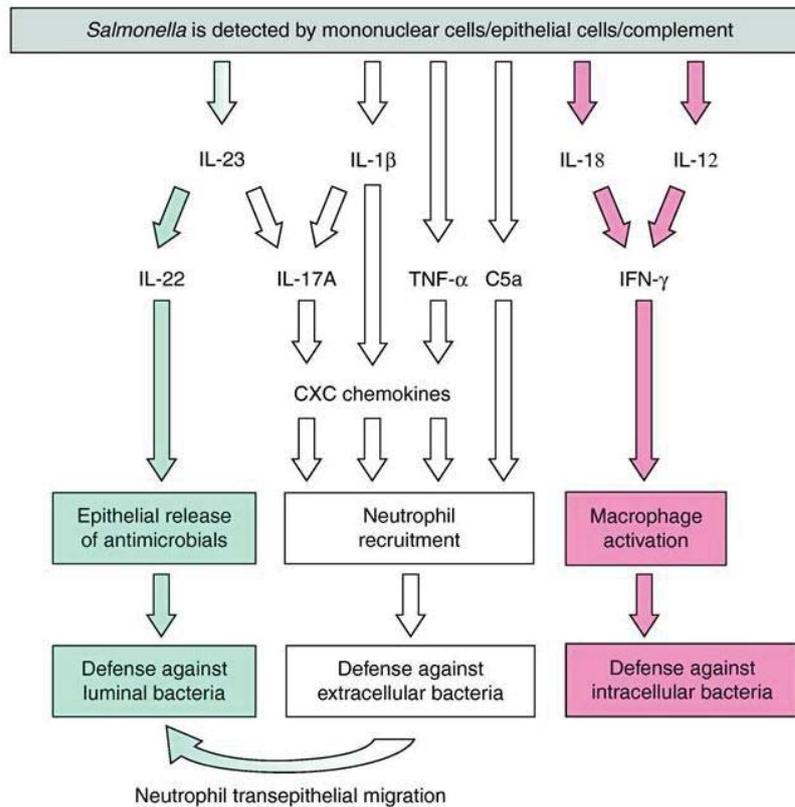


Figure 1.10. The major pathways in the mucosal response to *S. Typhimurium*. Following detection of *Salmonella* by epithelial or mononuclear cells, cytokine production is induced to bring about immune effector functions. Epithelial release of antimicrobial molecules occurs in response to activation by IL22. Multiple cytokines trigger the production of CXC cytokines which recruit circulating neutrophils to the site of infection, further aided by the anaphylotoxin C5a produced in complement activation. Th1 cytokines IL-18 and IL-12 act to induce IFN γ which stimulates macrophage killing of ingested bacteria. Through these three effector responses the immune system targets bacteria in multiple sites; in the lumen, and both extracellular and intracellular bacteria. Taken from [182].

1.5.3 The adaptive immune response

The role of the adaptive immune system in resolving infection and preventing re-infection has been relatively well studied in the murine *S. Typhimurium* model of typhoid fever. However, comparatively little research has investigated the involvement of the adaptive immune system in the gastrointestinal mucosa during gastroenteritis. Studies in the murine typhoid model have demonstrated that the importance of specific lymphocyte subsets is highly dependent on the infection method employed, and the particular *S. Typhimurium* and mouse strains involved [176]. However the consensus describes a significant role for T cells in both infection clearance and resistance to re-infection, with a role for B cells largely

restricted to the latter of these. T cell depletion studies and adoptive transfer of T cell-enriched cell fractions have shown that CD4⁺ T cells play a more significant role than CD8⁺ T cells in the murine model [183]. CD4⁺ T cells undergo polarisation to Th1 cells in the cytokine environment created by infection [184]. The Th1 effector function of IFN γ production is key to the host response. Experiments show that in the absence of certain IFN γ -producing cell subsets responsibility for production of the cytokine is assumed by other cell types [185].

Systemic infection with *S. Typhimurium* results in a strong antibody response to both protein and non-protein antigens, though these may not be essential for resolving the primary infection [186]. However the transfer of serum from vaccinated to naïve mice can confer protection, demonstrating the importance of antibodies in the response to re-challenge [187]. The mechanism by which protection is achieved is only partially understood. Antibodies may bind *Salmonella* in the brief extracellular period following exit from an apoptotic phagocyte, preventing entry to a new cell. Opsonization of *Salmonella* for Fc-receptor mediated uptake by phagocytes and triggering the classical pathway of complement activation leading to complement fixation on the bacterial surface may also be important. Also *Salmonella*-specific IgA antibodies in the mucosa have been shown to play an important role; cells producing sIgA specific to the O-antigen are one of the best correlates of protection following vaccination in humans [188].

To study the role of the adaptive immune system in non-typhoidal *Salmonella* infection the streptomycin mouse model described in section 1.3.7.2 has been studied [50]. *S. Typhimurium* with a mutation in the SPI-2 secreted effector protein *sseD* generates infection which is confined to the intestinal mucosa and GALT and cleared naturally. In this model mice lacking either T cells (TCR β ^{-/-} δ ^{-/-}) or B cells (J_H^{-/-}) demonstrate normal clearance and recovery from infection, indicating T or B cells are not essential for these processes.

Salmonella-specific sIgA are detected in the intestinal lumen within two weeks of infection with *S. Typhimurium sseD*, and are proposed to play an important role. sIgA was shown to agglutinate *S. Typhimurium* in the gut lumen, reducing access to the mucosal surface for invasion into tissue. However the presence of sIgA did not affect the process of clearance during an initial infection with *S. Typhimurium* in this model, instead *Salmonella*-specific luminal IgA was found to protect against re-infection. The use of an invasion deficient strain of *S. Typhimurium* confirmed previous work finding invasion is essential to

trigger production of *Salmonella*-specific sIgA [189]. Many of the sIgA antibodies produced during infection were specific for the O-antigen; the highly exposed and polymeric nature of this molecule making it a strong target. The O-antigen-specific nature of sIgA protection was demonstrated by following an initial infection with *S. Typhimurium* with a secondary *S. Enteritidis* challenge; mice were not protected against infection with the heterologous serovar. Although sIgA protects against re-infection, clearance of *S. Typhimurium* in the primary infection was shown to rely on the presence of a diverse microbiota. Future research is needed to identify how the adaptive immune system and microbiota may interact to deliver protection against intestinal pathogens [50].

1.5.4 Genetic susceptibility to intestinal inflammation in humans

Experimental animal models have provided many insights into pathways involved in inflammatory responses to enteric pathogens, however human studies are also an important approach to understanding inflammatory pathways and host defence mechanisms. Rare genetic deficiencies resulting in increased susceptibility to infection inform about the roles of the proteins they encode in a normal protective response.

1.5.4.1 Infection

Several genetic defects in individuals highly susceptible to severe or even fatal infections with otherwise weakly pathogenic strains of *Mycobacteria* or *Salmonella* have been identified. Mutations were discovered in genes of the pathway by which Th1 cells stimulate macrophage killing of intracellular bacteria, showing the central importance of this pathway in controlling intracellular bacteria. Mutations in IL12-p40 and IL12R β 1 affect the stimulation of Th1 and NK cells while mutations in IFN γ R1 and IFN γ R2 impact upon IFN γ detection by macrophages. Mutations in signalling molecules downstream of the IFN γ receptor such as STAT1 similarly prevent the translation of activating signals to effector functions in the macrophage [190, 191]. Such defects in the host immune response have been seen to provide a niche for *Salmonella* within the host and facilitate evolution to adapt the bacteria to the systemic environment (Klemm, unpublished).

1.5.4.2 Inflammatory bowel disease

Genome-wide association studies (GWAS) involving sequencing or genotyping many individuals with or without a particular disease have been used to identify genetic loci associated with the disease condition. In inflammatory bowel disease, GWAS has identified over 160 loci, many of which are associated with both Crohn's disease and ulcerative colitis [192]. A potentially close relationship between the autoimmune disease-associated loci and bacterial defences was clear for some of these genes. Genes important for processing intracellular bacteria are highly overrepresented in the GWAS loci for Crohn's disease, for example NOD2, IRGM and ATG16L1 [193]. NOD2 is an intracellular sensor for bacteria which together with NOD1 recruits the autophagy protein ATG16L1 to the membrane at the site of bacterial entry [194]. The failure of this autophagy pathway leads to altered bacterial handling and ultimately inflammation. Many of the associations identified by GWAS require further study to determine their potential role in bacterial infection and inflammation.

1.6 Aims of the thesis

The Wellcome Trust Sanger Institute (WTSI) has established a phenotypic screening platform using novel mutant mice that incorporates a pathogen challenge component (<http://www.immunophenotyping.org/>). This screen includes a systemic but not an oral *Salmonella* challenge. Therefore we decided to explore the potential of the murine *Salmonella* oral streptomycin treatment model as a secondary phenotyping component of the screen. To this end we used combined functional genomic approaches, including RNAseq and proteomics, to analyse wild type and selected mutant mice in depth. We anticipated this work would lead to the identification of key signatures and drive hypotheses for investigation in further experiments. Examples of such signatures are included in the work described herein.

2 Materials and methods

2.1 Materials

2.1.1 Bacterial strains

Salmonella enterica serovar Typhimurium SL1344 (naturally streptomycin resistant) was the wild type strain used in this study. Also used was an SL1344 $\Delta InvA$ derivative lacking the ability to invade epithelial cells (streptomycin and kanamycin resistant) [114].

2.1.2 Oligonucleotides

2.1.2.1 Sequencing of 16S rRNA genes

The following oligonucleotides were used in PCR of bacterial V1-V2 16S rRNA gene regions for microbial community analysis.

27F = AATGATACGGCGACCACCGAGATCTACAC TATGGTAATT CC
AGMGTTYGATYMTGGCTCAG

1 = 5' Illumina adapter, 2 = Forward primer pad, 3 = Forward primer linker, 4 = Forward primer.

338R = CAAGCAGAAGACGGCATAACGAGAT nnnnnnnnnnnn AGTCAGTCAG AA
GCTGCCTCCCGTAGGAGT

1 = Reverse complement of 3' Illumina adapter, 2 = Golay barcode, 3 = Reverse primer pad, 4 = Reverse primer linker, 5 = Reverse primer.

The Golay barcode is unique for each sample, as described previously [195].

2.1.2.2 Primers used in quantitative polymerase chain reaction (qPCR)

IL10rb spanning exons 1 and 2: Applied biosystems TaqMan® Gene Expression Assay
Mm00434153_m1

IL10rb spanning exons 3 and 4: Applied biosystems TaqMan® Gene Expression Assay
Mm00434155_m1

IL22ral spanning exons 4 and 5: Applied biosystems TaqMan® Gene Expression Assay
Mm01192943_m1

2.1.3 Antibodies

Target	Clone	Host	Fluorochrome/ conjugate	Source	Experiment(s) and dilution(s)
C3c	-	Rabbit	FITC	Dako	Western blotting (1 in 2000)
C3d	-	Goat	-	R & D Systems	Western blotting (1 in 5000) Fluorescence microscopy (1 in 10)
C3	-	Rabbit	-	LSBio	Fluorescence microscopy (1 in 10)
CD34	MEC 14.7	Rat	-	Abcam	Fluorescence microscopy (1 in 50)
β -tubulin	-	Rabbit	-	Abcam	Western blotting (1 in 50)
<i>Salmonella</i> common structural antigens (CSA)	-	Goat	FITC	KPL	Fluorescence microscopy (1 in 50)
CD4	RM4-5	Rat	BV510	Biologend	Flow cytometry (1 in 3000)
CD8a	53-6.7	Rat	APC-H7	BD	Flow cytometry (1 in 200)
CD19	1D3	Rat	PE-CF594	BD	Flow cytometry (1 in 2000)
Ly6G	1A8	Rat	V450	BD	Flow cytometry (1 in 600)
Ly6B	7/4	Rat	FITC	Serotec	Flow cytometry (1 in 1000)
CD115	AFS98	Rat	APC	Biologend	Flow cytometry (1 in 500)
CD3	145- 2C11	Armenian Hamster	PerCP-Cy5.5	Biologend	Flow cytometry (1 in 400)
CD11b	M1/70	Rat	PE-Cy7	Biologend	Flow cytometry (1 in 2000)
CD45	30-F11	Rat	Alexa Fluor 700	Biologend	Flow cytometry (1 in 600)
Goat IgG H & L	-	Donkey	Alexa Fluor 594	Abcam	Fluorescence microscopy (1 in 1000)
Rat IgG H & L	-	Goat	FITC	Abcam	Fluorescence microscopy (1 in 1000)
Rabbit IgG H & L	-	Goat	Cy3	Abcam	Fluorescence microscopy (1 in 1000)
Rabbit immunoglobulins	-	Goat	HRP	Dako	Western blotting (1 in 2000)
Goat IgG	-	Chicken	HRP	R & D systems	Western blotting (1 in 5000)

Table 2.1. Antibodies used in this study

2.1.4 Mice

All experiments were performed using 6 - 10 week old mice maintained in specific pathogen-free conditions, matched by age and sex within experiments. All mice were of C57BL/6 genetic background. Mice were housed in sterilised cages with food and water available *ad libitum*. All mice were killed by cervical dislocation. The care and use of all mice was in accordance with UK Home Office regulations (United Kingdom Animals Scientific

Procedures Act 1986). Mice were monitored throughout experiments for clinical symptoms including piloerection, hunched gait and lethargy, and mice displaying signs of severe disease were killed to prevent further suffering.

2.2 Methods

2.2.1 *S. Typhimurium* culture and preparation of inoculums

Bacteria were picked from Luria-Bertani (LB) agar plates containing appropriate antibiotic (30 µg/ml kanamycin and/or 50 µg/ml streptomycin sulphate) and grown overnight in 20 ml LB broth in a non-baffled flask at 37 °C with constant shaking at 190 rpm with appropriate antibiotics (concentrations as for plates). For infection of mice overnight cultures were serially diluted in sterile phosphate-buffered saline (PBS) to achieve a final dilution of 1 in 100,000. Following infection inoculum was serially diluted in PBS and plated on LB agar to confirm the CFU administered.

2.2.2 *In vivo* experiments and tissue harvesting

2.2.2.1 Streptomycin pre-treatment and infection with *S. Typhimurium*

Mice were treated 24 h prior to infection with 50 mg streptomycin: 200 µl of 250 mg/ml streptomycin sulphate in sterile water delivered via oral gavage. Mice were infected via oral gavage with between 8×10^3 and 2×10^5 CFU of *S. Typhimurium* in a 200 µl inoculum while naïve controls received 200 µl PBS. Oral gavage was performed with a sterilised blunt-tipped gavage needle under anaesthesia with isoflurane.

2.2.2.2 Enumeration of *S. Typhimurium* counts in tissue

Caecum, colon, small intestine, liver and spleen were removed from mice following confirmation of death for analysis of tissue-associated *S. Typhimurium* CFU. Caecum was cut open longitudinally and content removed by gentle scraping with tweezers. Colon and small intestinal content were removed by gently dragging tweezers along the length of the intact tissue. Tissue was mechanically disrupted (Steward stomacher 80, 2 min at high speed) in 5 ml of sterile water. Serial dilutions were made in PBS and plated on LB agar plates containing streptomycin (50 µg/ml). Colonies were counted following overnight growth at

37 °C. Implements used for the collection of material were cleaned thoroughly with 70% ethanol between samples.

2.2.2.3 Harvesting of tissue for RNAseq

Following confirmation of death caecum was removed from mice and divided into sections for further processing including RNA extraction for RNAseq analysis. 2 mm x 2 mm pieces of caecal tissue were cut reproducibly and placed in microtubes containing 1 ml RNAlater for RNA stabilisation. Tissue pieces were stored at -20 °C prior to RNA extraction.

2.2.2.4 Harvesting of intestinal content and tissue for microbiota analysis

All samples for microbiota analysis were placed in cryotubes and flash-frozen in liquid nitrogen immediately upon collection. For collection of faeces mice were placed in sterilised beakers and 1 - 2 faecal pellets removed and combined for analysis. Mice were killed at day 4 PI and following confirmation of death caecum and colon were removed. Caecum was cut open to allow gentle scraping of content from the surface of the tissue directly into a cryotube. A piece of caecal tissue approximately 1 cm in length from the region adjoining the colon was removed for analysis. For collection of colon content the colon was cut open longitudinally and content material closest to the proximal end removed for analysis. Care was taken to use sterile implements for the collection of samples for microbiota analysis, and implements were cleaned thoroughly with 70% ethanol between samples.

2.2.2.5 Harvesting and culture of peritoneal macrophages

4% w/v thioglycollate (Sigma) solution was prepared using ddH₂O water, autoclaved and stored at 4 °C. 250 µl thioglycollate was delivered to mice by intraperitoneal (IP) injection. At day 4 post-injection mice were killed and the peritoneal cavity washed with 8 - 10 ml of cold PBS. Isolated cells were centrifuged at 220 g for 15 min at room temperature and resuspended in warmed media (RPMI-1640 (Sigma), 10% heat-inactivated foetal calf serum (FCS) (Biosera), 2 mM L-glutamine (Invitrogen), 10,000 U/ml penicillin-streptomycin (Invitrogen)). Peritoneal cells were counted and seeded in 24-well plastic plates with two wells per mouse. Cells were allowed to fix to the culture dish overnight at 37 °C, 5% CO₂. The following day cells were washed with warmed PBS (37 °C) to remove non-adherent cells and peritoneal macrophages were used for RNA or protein extraction. For RNA extraction

buffer RLT from an RNeasy Mini kit (Qiagen) was added directly to cells in the 24-well plate. Removal and lysis of cells was performed using a plastic cell scraper followed by transfer of lysate to a 1.5 ml microtube and vortexing. RNA was extracted according to the directions provided in the RNeasy kit accompanying handbook. For protein extraction chilled Radio-Immunoprecipitation Assay (RIPA) buffer (150 mM sodium chloride, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate (SDS), 50 mM Tris pH 8, cOmplete protease inhibitor cocktail (Roche), in ddH₂O) was added directly to cells in the 24-well plate at 1 ml/10⁷ cells. Adherent cells were removed and lysed using a plastic cell scraper and lysate transferred to a pre-cooled 1.5 ml microtube. Lysate was maintained under constant agitation for 30 min at 4 °C followed by centrifugation at 16,000 g for 20 min. Supernatant was transferred to chilled microtubes for protein concentration determination and further analysis.

2.2.2.6 Harvesting and culture of bone marrow-derived macrophages (BMDM)

Bone marrow cells were obtained by flushing marrow from bones of the hind legs with cold PBS using a 25-gauge needle. Cells were dissociated by gently passing the bone marrow suspension through an 18-gauge needle. Cells were centrifuged at 220 g for 15 min at room temperature and re-suspended in warmed media (RPMI-1640, 10% heat-inactivated FCS, 2 mM L-glutamine, 10,000 U/ml penicillin-streptomycin, 10% filtered supernatant from mouse L929 cell culture as the source of macrophage colony-stimulating factor (M-CSF)). Cells were plated in sterile 10 cm plastic petri dishes, two dishes per mouse, and incubated for 1 week at 37 °C and 5% CO₂. On day 3 post-harvest fresh warmed media was added and on day 5 post-harvest cells were washed with warmed PBS (37 °C) to remove non-adherent cells followed by addition of fresh media. On day 7 post-harvest fully-differentiated BMDM were used for RNA or protein extraction as described for peritoneal macrophages (in each case the solution for cell lysis was added directly to the petri dish).

2.2.3 Peripheral blood leukocyte analysis

Blood was obtained by cardiac puncture performed under anaesthesia with isoflurane and followed immediately by cervical dislocation. Harvested blood was transferred to 100 µl EDTA coated blood tubes and inverted to mix. 25 µl of whole blood was incubated in a 96-well plate for 20 min with a combination of antibodies for cell surface markers (CD4, CD8a, CD19, Ly6G, Ly6B, CD115, CD3, CD11b and CD45) diluted in 25 µl FACS buffer

(5% heat-inactivated FCS, 0.1% sodium azide in PBS) as listed in section 2.1.3. Cells were fixed for 5 min by addition of three volumes of CellFIX solution (BD Biosciences) followed by centrifugation at 400 g for 3 min, and the cell pellet resuspended in Pharm Lyse solution (BD Biosciences) for red blood cell lysis. Subsequently cells were incubated for 5 min, centrifuged again at 400 g for 3 min, and again resuspended in Pharm Lyse solution and incubated for 5 min. Finally cells were washed once in FACS buffer prior to resuspension in FACS buffer for flow cytometry analysis. All incubations were at room temperature. Compensation controls were prepared with UltraComp eBeads (eBioscience). Samples were analysed on a FACS Aria II (BD) and subsequent data analysis performed using the software FlowJo v7.6.5.

2.2.4 Microscopic analysis

2.2.4.1 Histopathological analysis

Tissue Processor and Stainer and Coverslipper machines were operated by Yvette Hooks (WTSI).

Following confirmation of death a 0.5 cm tubular section of caecum, close to the blind end and adjacent to the caecal patch, was excised, caecal content gently removed, and the tissue placed in 4% formaldehyde for 24 h. A 0.5 cm section of proximal colon and pieces of liver and spleen were placed in formaldehyde also. Permanent fixation was performed using a Sakura Vacuum Infiltration Processor 5 which moves samples through formaldehyde, a series of mixtures of ethanol and water for dehydration, and xylene to clear samples before immersion in molten paraffin wax. Samples were embedded in paraffin wax using an embedding station (Leica), and 5 µm sections cut using a RM2125 rotary microtome (Leica) and floated on a 37 °C water bath for 3 - 5 min. Sections were then transferred to Superfrost plus glass slides (VWR International).

Paraffin sections were deparaffinised, rehydrated, stained with Gill's 2 hematoxylin (Leica) and eosin (Leica), dehydrated and mounted using a Leica Stainer and Coverslipper machine. Briefly, the machine moves slides through a series of solutions; xylene to deparaffinise, a series of mixtures of ethanol and water for rehydration and dehydration, and finally sections were mounted with and Leica mountant. Slides were visualised using a

LSM510 confocal microscope (Carl Zeiss Ltd.). Histopathological analysis was performed by Professor Mark Arends (University of Edinburgh Division of Pathology).

2.2.4.2 Immunofluorescence staining

For immunofluorescence staining caecal tissue was removed as described in section 2.2.4.1 and placed in 4% paraformaldehyde (PFA) for 1 h. Tissue underwent three washes of 10 min in PBS before transfer to 0.075 M glycine in PBS for 30 min for quenching of unreacted aldehyde groups. Fixed caecum pieces were embedded in optimal cutting temperature compound and frozen with cryofreeze aerosol (Agar scientific) before flash-freezing in liquid nitrogen. 5 µm cross sections of OCT-embedded caecum were cut using a cryostat and air-dried on poly-l-lysine-coated glass microscope slides for 2 h. Sections were outlined with a wax pen and incubated with blocking solution (10% FCS in RPMI 1640 culture medium) for 30 min. Blocking solution was replaced with primary antibody diluted in blocking solution as described in section 2.1.3. Unstained and 'secondary antibody-only' controls were also prepared with blocking solution only. Slides were incubated with primary antibody for 1 h and underwent three washes of 5 min in PBS. Secondary antibody diluted in blocking solution as described in section 2.1.3 was added to sections for 1 h after which sections again underwent three washes of 5 min in PBS. For staining of multiple antigens in a single section antibodies were added to consecutively as follows: primary 1 → wash → secondary 1 → wash → primary 2 → wash → secondary 2 → wash. Prolong Gold antifade reagent with DAPI (Life technologies) was applied before mounting and visualisation of sections with a Zeiss LSM510 confocal microscope.

2.2.4.3 Three-dimensional confocal imaging

Following removal caecum was cut open longitudinally and flushed with cold Hank's balanced salt solution containing 0.4 M N-acetyl-L-cysteine. Tissue was placed in 4% PFA in PBS for 2 h at room temperature for fixation and stored in 1% PFA in PBS at 4 °C prior to further processing. Tissue was washed three times in PBS before transfer to blocking and permeabilisation solution for an incubation of 4 h at room temperature (2% Triton X-100, 10% FCS, in PBS). Tissue was again washed three times in PBS before incubation with primary antibodies in 0.2% Triton X-100, 5% FCS, in PBS at 16 °C with gentle agitation overnight. Tissue was again washed three times in PBS before incubation with secondary antibodies in 0.2% Triton X-100, 5% FCS, in PBS at 4 °C with gentle agitation overnight.

Tissue was again washed three times in PBS followed by nuclear counterstaining with 10 nM DAPI dilactate in PBS for 1h at room temperature with gentle rocking. Tissue was washed six times in PBS and immersed in FocusClear (CelExplorer) and covered for 4 hours. Tissue was mounted in Prolong Gold prior to imaging with a Zeiss LSM510 confocal microscope.

2.2.5 RNA methods

2.2.5.1 RNA extraction

RNA was extracted from tissue using an RNeasy Mini kit (Qiagen) according to the directions provided in the accompanying handbook. Tissue disruption and cell lysis were performed in buffer RLT in a GentleMACS Dissociator (Miltenyi Biotec) using the pre-set program for total RNA extraction. The concentration and quality of extracted RNA were assessed using a Bioanalyzer (Agilent). Samples with an RNA Integrity Number RIN < 8 were excluded from further processing and analysis.

2.2.5.2 Reverse transcription and qPCR

Complementary DNA (cDNA) was synthesized from extracted RNA using the QuantiTect Reverse Transcription Kit (Qiagen). qPCR performed for the validation of RNAseq (Section 3.3.6.1) was carried out using Absolute Blue QPCR SYBR Green ROX mix (Thermo Scientific), and SYBR green primers. qPCR performed for analysis of gene expression in *tmla* mutant mouse lines (Section 6.3.1) was carried out using Absolute Blue QPCR ROX mix (Thermo Scientific) and TaqMan probe-based gene expression analysis.

2.2.5.3 RNAseq

2.2.5.3.1 Library preparation and sequencing

Multiplexed RNA libraries were prepared using the Truseq RNA sample prep kit (Illumina). Briefly, the kit uses oligodT beads to enrich for mRNA prior to fragmentation using divalent cations under elevated temperature. Random primers were used for reverse transcription of mRNA fragments, producing cDNA to which Illumina adapter sequences with indexing barcodes; allowing multiple samples to be pooled for sequencing; were ligated. Samples were multiplexed such that all samples within an experiment were equally divided between flow cell lanes to avoid lane bias. Libraries were quantified on a Bioanalyzer high

sensitivity DNA chip (Agilent) and sequenced on a HiSeq 2000 sequencer (Illumina) to produce 100 bp paired end reads. The target sequencing output was 5 Gb/sample.

2.2.5.3.2 RNAseq data analysis

Sequencing reads were aligned with Tophat [196] version 2.0.8 to the mouse reference genome version MM10/GRCm38. The read counts per gene were generated with featureCounts version 1.4.5-p1; the annotation for which was from ENSEMBL 77. Read counts were used to represent gene expression levels. R version 3.1.1 was used to import count data, and the DESeq2 package used to normalise read counts between samples based on the sample sequencing depth, and detect differentially expressed genes [197]. DESeq2 uses the Benjamini-Hochberg adjustment for multiple testing. Regularised logarithm (rlog)-transformed data was used for construction of principal component analysis (PCA) plots and heatmaps using the R package ggplot2.

2.2.6 16S rRNA gene sequencing for microbiota analysis

2.2.6.1 DNA extraction, library preparation and sequencing

DNA was extracted from faecal pellets, colon content, caecal content and caecal tissue using the FastDNA Spin Kit for Soil and FastPrep Instrument (both MP Biomedicals). 16S rRNA genes were amplified by PCR with Q5 High-Fidelity DNA Polymerase (New England Biolabs) using primers as described in section 2.1.2. Four separate PCR reactions were performed per sample and successful amplification was verified by gel electrophoresis of PCR products. Products of the four PCR reactions were pooled and DNA precipitated by ethanol precipitation as follows. 0.3 volumes (relative to the pooled sample volume) of 1 M sodium chloride were added to the sample. 2 volumes (relative to the combined volume of sample and sodium chloride) of chilled ethanol were added and mixed by inversion, and samples placed at -20 °C overnight. Samples were spun at 16,000 g, 20 min, 4 °C, and the DNA pellet washed with 600 µl of chilled 70% ethanol, then re-spun and dried using an Eppendorf Concentrator before resuspension in water. DNA was quantified by Qubit 2.0 Fluorometer using a dsDNA HS Assay Kit (both Life Technologies) and an equimolar mix of the DNA samples prepared. Equimolar mix was run on a 1% agarose gel and gel purification performed using a Wizard SV Gel and PCR Clean Up Kit (Promega). DNA was sequenced by

paired-end sequencing on a MiSeq sequencer (Illumina). The number of clusters sequenced per sample was $48,068 \pm 20,647$ (mean \pm standard deviation).

2.2.6.2 16S rRNA sequencing data analysis

FASTQ files were processed using the MOTHUR software [198]. The software platform R was used to remove low abundance bacterial groups (density < 0.5%) and generate plots. The web-based tool Interactive Tree Of Life (iTOL) was used to display the tree of sample relatedness alongside microbiome composition data [199]. MOTHUR was used to generate the within-sample α -diversity estimates.

2.2.7 Proteins

2.2.7.1 Analysis of tissue extracts by MS

2.2.7.1.1 Extraction of protein from caecal tissue

350 μ l of lysis buffer (4% SDS, 150 mM sodium chloride, 50 mM Tris buffer pH 7.6, 2 mM EDTA, 40 mM TCEP, in HPLC-purified H₂O) was chilled on ice in GentleMACS M-tubes (Miltenyi Biotec). 40 - 60 mg pieces of flash-frozen caecal tissue were added to tubes and homogenisation performed in a GentleMACS Dissociator using the pre-set program for protein extraction. Tubes were spun briefly and content transferred to 1.5 ml microtubes. Samples were heated at 95 °C for 10 min before sonication with an ultrasonic probe (Fischer Scientific); 40% energy, 30 cycles (or until cleared) of 1 s pulses at 1 s intervals. Sonicated samples were centrifuged at 16,000 g for 15 min at 18 °C and supernatants transferred to clean microtubes.

2.2.7.1.2 Protein concentration determination

Protein concentration determination was performed by 660 nm protein assay (Pierce) with addition of Ionic Detergent Compatibility Reagent (Pierce).

2.2.7.1.3 Preparation for MS

400 mM iodoacetamide was added to samples at 200 μ l/ml protein extract and samples incubated in the dark for 1 h at room temperature. Protein concentrations of IAA-treated extracts were normalised to that of the least concentrated sample by addition of lysis

buffer, then diluted 10-fold in 8 M urea in 0.1M Tris/HCl pH 8.5 for filter-aided sample preparation (FASP) [200]. All FASP centrifugation steps were performed at 20 °C. Urea-diluted extract volumes containing 80 µg total protein were transferred to Amicon Ultra-0.5 Centrifugal Filter Units, nominal molecular weight limit 30 kDa, (Merck Millipore) and centrifuged at 14,000 g until the dead volume was below the lowest marked level (~ 15 min). Flow-through was collected for the analysis of < 30 kDa proteins. Loading and centrifugation were repeated as required for transfer of the entire volumes containing 80 µg of protein. 200 µl of 8 M urea in 0.1 M Tris/HCl, pH 8.5, was pipetted into the filter units followed by centrifugation as before, three times consecutively. As required, waste liquid was pipetted carefully out of the collection tube for disposal. 200 µl of 0.05 M ammonium bicarbonate in HPLC-purified H₂O was pipetted into filter units with centrifugation as before, three times consecutively. Filters were then transferred to fresh collection tubes containing 20 µl of 0.05 M ammonium bicarbonate to aid chamber humidification, and 350 µl of 4.6 ng/µl Trypsin Gold (Promega) in 0.05 M ammonium bicarbonate was pipetted into filter compartment for protein digestion. Filter units were placed in a thermomixer at 600 rpm for 1 min before overnight incubation at 37 °C. Following incubation filter contents were transferred to a clean microtube, acidified to pH 3 - 4 with 25% trifluoroacetic acid, and diluted in HPLC-purified H₂O to a final protein concentration of 0.2 µg/µl. Yeast enolase (Waters) was added at a final concentration of 1 pmol/µl to provide an internal standard.

The retained flow-through fractions from centrifugation of extracts in Amicon Ultra-0.5 Centrifugal Filter Units were combined in naïve and infected sample pools. Pooled samples were applied to Vivaspin centrifugal concentrators, nominal molecular weight limit 5 kDa (Sigma) and washed with 8 M urea in 0.1 M Tris/HCl, pH 8.5 and 0.05 M ammonium bicarbonate as for the > 30 kDa protein fractions. Samples were collected and acidified for MS analysis without prior trypsin digestion.

2.2.7.1.4 MS analysis

MS analysis and database searching was performed by Dr Lu Yu (WTSI).

Quantitative liquid chromatography – tandem mass spectrometry (LC-MS/MS) was analysed on a nano-Acquity UPLC system (Waters Corp) coupled to a Synapt G2-S HDMS mass spectrometer (Waters Corp) with a nanoelectrospray source. Briefly, 300 ng of each sample was loaded and desalted at 5 µl/min for 5 min onto a trap column at 5 µm Symmetry

C18 180 μm \times 20 mm column (Waters). Peptides were separated on a nanoACQUITY UPLC HSS T3 column at 75 μm id \times 150 mm (Waters) using a linear gradient of 1 to 40% MeCN with 0.1% formic acid over 90 min and cycle time 120 min, at a flow rate of 0.3 $\mu\text{l}/\text{min}$ and column temperature of 35 $^{\circ}\text{C}$. Data collection was performed in ion mobility-assisted data independent acquisition (HDMS^E) modes, used a 0.6 s alternating cycle time between low (4 V) and high (15 - 45 V) collision energy (CE). m/z at 785.8426 from [Glu1]-Fibrinopeptide (Sigma) at 100 fm/ μl was used as lockmass and acquired at every 60 s. Each sample was analysed in four runs.

2.2.7.1.5 MS data analysis

MS data analysis was performed by Dr James Wright (WTSI).

For the > 30 kDa protein fractions raw spectrum files were processed and identified using Protein Lynx Global Server 3.0 (Waters), against a combined mouse, *Salmonella* and contaminate database (UniProt Proteomes April 2014). The Ion Accounting (IA) results were exported from PLGS and further processed using in house software. Peptides were filtered at a 4% false discovery rate (FDR) before being clustering into a definitive list of proteins using an Occam's razor approach. Protein quantification was performed using a High3-MSMS intensity based approach, where for each protein the summed fragment ion intensities of the most intense three peptides were summed together. Only unique PepFrag1 peptides were allowed in the quantification. These quantification values were normalised using a multistage approach. Firstly, the protein quantifications were normalised by a spiked in protein standard (Yeast Enolase), the %CV was then calculated for each normalised protein across all samples. Using this %CV the 10 least variable proteins across the samples were selected and used in a secondary normalisation. The median normalised quantification values for each sample types were then used to calculate log₂ fold changes between the different sample types. Fold changes were assessed for significance using a moderated t-test.

The data from analysis of < 30 kDa protein fractions was processed with MaxQuant. Potential contaminants, protein groups with a Q-value of > 0.015, and protein groups with an MS/MS count of ≤ 1 were excluded. The ratio of protein group intensities in naïve and infected pooled samples were used to reflect the difference in protein abundance between these conditions.

2.2.7.2 Western blotting

2.2.7.2.1 Preparation of intestinal content and faecal samples

Faeces and intestinal content was immediately homogenized upon collection in chilled RIPA buffer using a sterile pipette tip, followed by vortexing for 30 s. Samples were centrifuged at 16,000 g for 20 min at 4 °C and supernatants removed for protein concentration determination.

2.2.7.2.2 Preparation of tissue samples

RIPA buffer was added to flash-frozen tissue at 40 µl/mg and tissue homogenized in a GentleMACS Dissociator (Miltenyi Biotec) using the pre-set program for protein extraction. Homogenised tissue was maintained under constant agitation for 2 h at 4 °C followed by centrifugation at 16,000 g for 20 min. Supernatants were removed for protein concentration determination.

2.2.7.2.3 Preparation of plasma

Whole blood in EDTA-coated collection tubes was centrifuged at 2,000 g for 15 min and supernatant removed for protein concentration determination.

2.2.7.2.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein concentration of plasma and protein extracts was determined by bicinchoninic acid (BCA) assay (Pierce). Plasma and protein extracts were diluted for blotting in RIPA buffer and prepared for gel electrophoresis by addition of 4X laemmli buffer (8% SDS, 20% β-mercaptoethanol, 40% glycerol, 0.008% bromophenol blue in 0.125 M Tris HCl pH 6.8). Samples were heated to 95 °C for 5 min and allowed to cool before electrophoresis by SDS-PAGE. Samples were loaded onto 12% Mini-PROTEAN TGX precast gels (Biorad) and run at 175 V for 45 min with Tris/glycine-running buffer.

2.2.7.2.5 Blotting and protein visualisation

Proteins were blotted onto ethanol-activated polyvinylidene difluoride (PDVF) membrane using a semi-dry transfer system (70 mA, 75 min) in transfer buffer (48 mM Tris, 39 mM glycine, 0.04% SDS, 20% methanol in ddH₂O). Membranes were blocked by

incubation in 5% milk in PBS-T (0.1% Tween-20 in PBS) for 1 h. Primary antibody in 2% milk in PBS-T was added for 1 h followed by three washes of 5 min in PBS-T. Appropriate HRP-conjugated secondary antibody was added for 45 min in 2% milk in PBS-T, followed by three washes of 15 min in PBS-T. Amersham Enhanced Chemiluminescence (ECL) Western Blotting Detection reagent was used for detection of protein bands with Amersham Hyperfilm ECL (both GE Healthcare) as per the manufacturer's instructions. All membrane incubations were carried out at room temperature on a rocking platform. ImageJ software was used for densitometry analysis of protein bands.

2.2.8 Statistical tests

Testing for differences in mouse weight loss was performed using one-way ANOVA. Mann Whitney U-tests were performed for analysis of organ *Salmonella* CFU and blood leukocyte populations. Pathway analysis was performed using the analysis tools available on the InnateDB website (<http://www.innatedb.com/index.jsp>). The hypergeometric analysis algorithm and Benjamini Hochberg p-value correction methods were used in all pathway analysis. Densitometry analysis of signal intensity in Western blotting was performed with a Student T-test.

Venn diagrams were produced using Venny and BioVenn [201, 202].

In all cases, a p-value of ≤ 0.05 was considered to be significant. Tests were performed using GraphPad Prism 6 graphing and statistical software (GraphPad Software, Inc.).

Star symbols were used to indicate statistical significance as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

3 Pathological and transcriptional changes in the streptomycin mouse model of human *S. Typhimurium* gastroenteritis

3.1 Introduction

3.1.1 Characterisation of the streptomycin mouse model

The streptomycin mouse model described in 1.3.7.2 was developed as a surrogate system to investigate the pathology and pathogenesis of *S. Typhimurium*-mediated gastroenteritis. Oral treatment with streptomycin prior to challenge with *S. Typhimurium* greatly facilitates colonisation of intestinal tissues, in particular the caecum, giving rise to a 10^5 - 10^6 -fold reduction in the oral 50% infectious dose [203-205]. Whilst the murine typhoid model has been studied extensively, the streptomycin model described herein is a comparatively recent development. Consequently, fewer studies have addressed the processes and pathways involved in *S. Typhimurium*-induced intestinal inflammation compared with the systemic aspects of *Salmonella* infection.

Others have described features of the streptomycin model, such as effects on the microbiota and histological changes in inflammation. However, a more detailed characterisation, particularly at the molecular and functional genomic level, could further our understanding. Additionally, in order to establish the streptomycin model as a pathogen challenge for mutant mice in higher throughput functional genetic studies, the model must be thoroughly validated and be reproducible over multiple challenges. A thorough investigation of the model would be beneficial to the research community interested in open access data.

Previous studies have detected no significant difference in caecal histopathology between streptomycin-treated and untreated mice, excluding obvious toxic effects of the antibiotic upon caecal tissue [149, 204]. Increased susceptibility to infection is thought a consequence of the disruption of natural microbial species present in the mouse intestine that can competitively exclude or outcompete *Salmonella* [148]. Existing work has investigated changes in intestinal microflora in response to streptomycin and *S. Typhimurium*, however new sequencing approaches and cheaper costs permit much deeper profiling of the microbiota than was performed in these studies [108, 206]. Some of the inflammatory changes that occur in the caecum during *S. Typhimurium* infection in the streptomycin model have been

described. They include oedema in the submucosa and lamina propria, structural disruption and lengthening of crypts, epithelial cell erosion, and infiltration of leukocytes into the lamina propria and intestinal lumen [149]. Here, we sought to confirm and extend previous investigations using a combination of histological and functional genomic approaches.

3.1.2 Transcriptional profiling

Measurement of RNA in transcriptional profiling produces a snapshot of the activity of individual genes in cells or tissue. Though less closely correlated with biological activities than proteins, RNA molecules are arguably significantly more easily quantified at a genome-wide level and therefore their patterns of expression serve as an important indicator of gene activity. Comparison of transcriptional profiles from cells or tissue under different conditions and treatments produces lists of differentially expressed (DE) genes, which can provide insight into the processes and activities defining these different states.

Microarray technology has been widely used to study changes in gene expression. Briefly, cDNA produced by reverse transcription of RNA in the sample is hybridised to a set of probes representing the selection of genes under investigation, and fluorescent signals generated by bound cDNA are measured for cDNA quantification. Although a valuable source of transcriptomic data for many years, currently microarrays are undergoing replacement by next generation sequencing. RNAseq has many advantages over microarrays; principally the unbiased nature of the approach allows direct quantification of multiple RNA species present in the cell, potentially including both coding and non-coding RNAs such as microRNAs and siRNAs. In contrast to microarrays, RNAseq is well suited to the quantification of alternatively spliced transcripts and detection of genetic variation such as SNPs. RNAseq largely avoids errors inherent to microarray technology, such as false positives from cross-hybridisation of probes with cDNA molecules related to the intended targets. Furthermore RNAseq has been shown to achieve a larger dynamic range and sensitivity in measurements. Several studies have reported that overlap between DE genes determined by microarray and RNAseq is incomplete [207, 208]. For example, a study of gene expression in bovine macrophages in response to *M. bovis* infection determined that the percentage of overlapping genes was just 48% for microarray and 37% for the RNAseq data [209]. In support of RNAseq, qPCR estimates of fold changes in DE genes more closely mirror those determined by RNAseq than microarray [210].

Transcriptional profiling has been used to study the changes in response to infection with *Salmonella* in a number of settings. A study of zebrafish embryos experimentally infected with *S. Typhimurium* used transcriptional profiling to investigate the lower vertebrate innate response to *Salmonella* infection [211]. Blood samples from HIV patients with iNTS and other invasive bacterial infections were studied by microarray to identify features of the transcriptomic response which distinguish iNTS [212]. A study of the streptomycin mouse model using microarrays provided early insight into the transcriptomic response to infection in the colon [213]. In our study we expand on this work, choosing the caecum for our investigations as this tissue is more consistently and severely affected than the colon in the streptomycin model.

3.1.2.1 Pathway analysis

Studying a list of DE genes can identify candidate genes, which may be involved in the response to an infection, however individual genes do not act alone. When examining such a list it is difficult to appreciate how single genes contribute to activities involving the concerted activity of multiple genes. Pathway analysis can provide a route into the interpretation of the data produced in transcriptomics and other large-scale profiling experiments such as proteomics and metabolomics, moving from single genes to whole pathways. The approach of pathway analysis is to compare lists of DE genes or proteins linked to established pathways to identify pathways potentially activated or repressed by the treatment condition. ‘Pathway’ is a broad term used with multiple meanings, from describing a set of physically interacting molecules, to a group of enzymes which respond similarly to stimulation, or a group of genes whose expression shows related patterns of regulation. A variety of pathway databases have been generated, many of which combine the types of pathways described. These databases differ in the sources of information used to define pathways, the general categories of pathways they contain, and the stringency applied to when assigning genes to pathways [214].

InnateDB is a publically available database of pathways and molecular interactions, and collection of systems biology tools for the analysis of datasets of DE genes [215]. Originally created to contain interactions involved in innate immunity, InnateDB now incorporates interaction data from a wider collection of databases. In the most recent report the database contained over 196,000 experimentally validated interactions including more than 18,000 involved in innate immunity, and 3,000 pathway annotations [216]. Up to date

figures are available at (<http://www.innatedb.com/statistics.do?s=>). Interactions are manually added to InnateDB based upon information on molecular interactions from published articles [217]. Manual curation results in fewer erroneous entries compared with automated data mining approaches and allows inclusion of associated contextual information. In addition interactions sourced from published articles are often validated by numerous experimental techniques and hence likely contain fewer false interactions compared with use of interaction screening approaches such as yeast two-hybrid systems. As a member of the International Molecular Exchange (IMEx) consortium, InnateDB is required to abide by strict curation rules and capture deep interaction data. Pathways involved in innate immunity are non-linear and highly complex, therefore system-oriented analysis tools such as InnateDB are highly valuable.

3.2 Aims of the work described in this chapter

Changes occurring in mice infected with *S. Typhimurium* following streptomycin treatment were investigated using a number of established experimental techniques. These include histopathology and RNAseq analysis of caecal tissue in order to assess microscopic and molecular changes respectively, flow cytometry for the analysis of immune cell populations in blood, and 16S rRNA gene sequencing of faecal material and mucosal tissue to investigate changes in the microbiota. These data provide a detailed description of the streptomycin mouse model and data sets against which infections in mutant mice can be compared.

3.3 Results

3.3.1 Infection with *S. Typhimurium* results in weight loss and colonisation of gastrointestinal and systemic organs

Previous studies have shown that C57BL/6 mice orally treated with streptomycin and infected with *S. Typhimurium* exhibit weight loss, and that *Salmonella* effectively colonises intestinal tissues. In addition *Salmonella* disseminates to systemic organs, although counts here are lower than in the gut. We sought to replicate these findings in our colony of specific pathogen free (SPF) C57BL/6N mice with wild type *S. Typhimurium* SL1344. Mice were assigned to one of two groups; a naïve control group which received streptomycin followed at 24 h by PBS, and an infection group which received streptomycin followed at 24 h by

S. Typhimurium SL1344. Upon delivery of streptomycin mice were weighed daily and culled at day 4 PI, upon which organs were removed for determination of *Salmonella* CFU.

The effects of inactivating mutations in SPI-1 have been investigated in the streptomycin mouse model and shown to attenuate the severity of colitis [149]. We aimed to confirm this by infection with a $\Delta InvA$ SL1344 mutant derivative. In this mutant derivative the absence of an essential component of the SPI-1 T3SS secretory apparatus results in defective host cell invasion. In a rabbit ileal loop model of *S. Typhimurium* infection SL1344 $\Delta InvA$ is relatively attenuated [218].

We observed mice infected with wild type SL1344 experienced highly significant weight loss, whereas naïve controls and mice infected with SL1344 $\Delta InvA$ did not significantly lose weight following treatment (Figure 3.1A). We observed high numbers of *Salmonella* CFU in the caecum and colon of wild type SL1344-infected mice ($\sim 10^8 - 10^9$ CFU/g tissue) and high to moderate infection in the liver and spleen ($\sim 10^5 - 10^6$ CFU/g tissue). The small intestine was relatively poorly colonised ($\sim 10^4$ CFU/g). The main *Salmonella*-containing tissues in the intestinal tract showed significantly lower levels of colonisation by the SL1344 $\Delta InvA$ derivative compared with wild type SL1344 (Mann Whitney U test, caecum $p = 0.002$, colon $p = 0.026$), whereas dissemination to the spleen and liver was relatively less impacted by the SPI-1 mutation (Figure 3.1B).

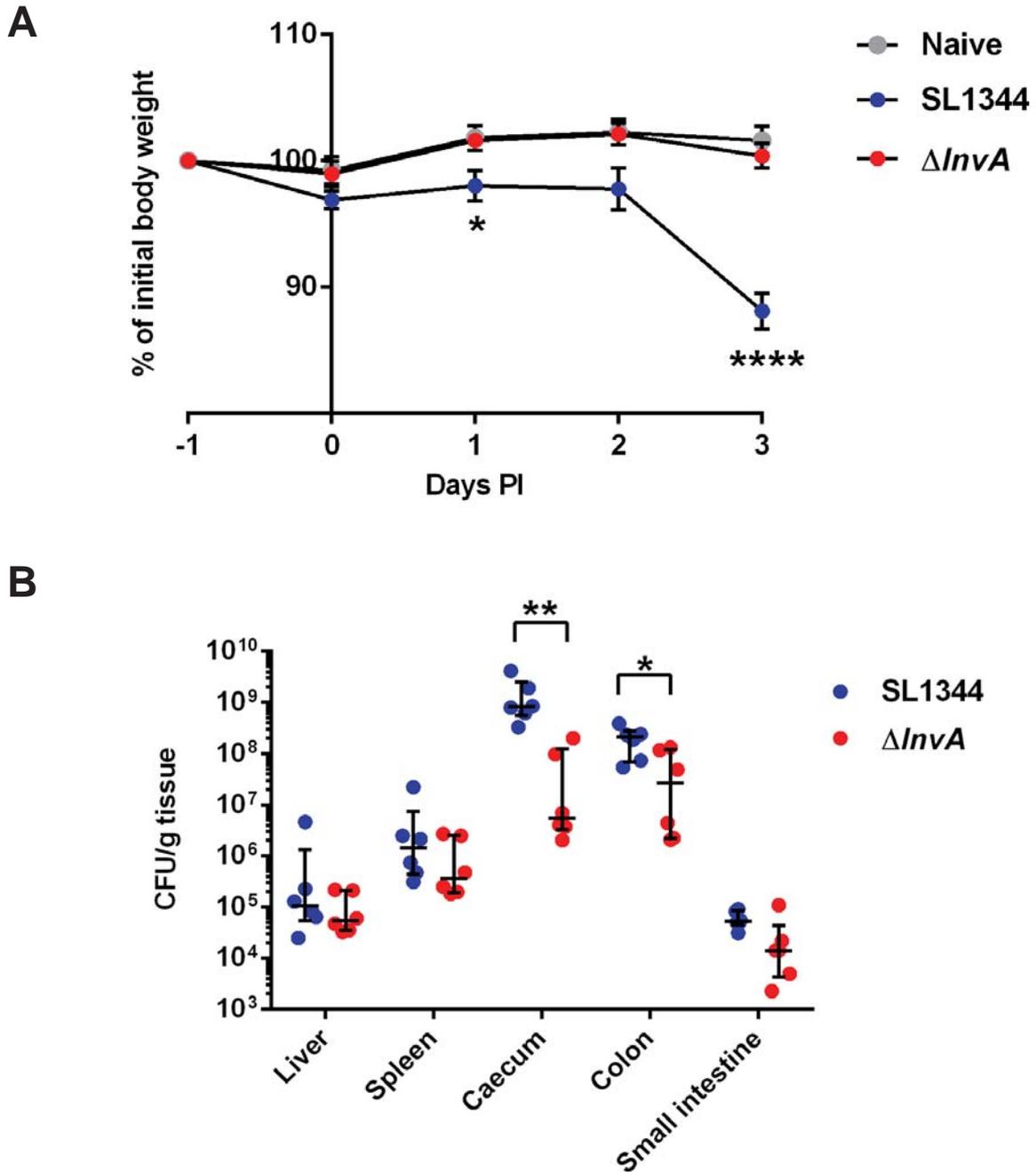


Figure 3.1. Weight loss and tissue colonisation in the streptomycin mouse model. (A) Representative weight curve of C57BL/6N mice treated with streptomycin prior to oral delivery of PBS (naïve controls), wild type *S. Typhimurium* SL1344 (9×10^4 CFU), or *S. Typhimurium* SL1344 $\Delta InvA$ (1.6×10^5 CFU). Data shown are mean \pm SEM, minimum 6 mice per group. One way analysis of variance for each time point was performed for statistical analysis. (B) Organ CFU from wild type SL1344 and SL1344 $\Delta InvA$ infected mice at day 4 PI. Bars indicate median and interquartile range. Significance was determined by Mann Whitney U test.

3.3.2 *S. Typhimurium* infection results in major changes in caecal morphology

Marked changes have been shown to occur in the mouse caecum upon *S. Typhimurium* infection following streptomycin treatment [149]. We confirmed similar changes take place in our study. No difference was observed in the gross pathology of caeca between untreated and naïve control mice receiving 50 mg streptomycin. At day 4 PI the caeca of mice infected with *S. Typhimurium* SL1344 were markedly smaller than those of naïve controls with thickened tissue walls (Figure 3.2). Normal dark brown caecal content was replaced in infected mice by thin yellowish fluid. The colon also appeared thickened and content material was reduced and softened. The caeca of mice infected with SL1344 $\Delta InvA$ appeared intermediate to those from naïve and SL1344 infected mice in terms of general pathology (data not shown).



Figure 3.2. Changes in caecum morphology in *S. Typhimurium* infection. Photographs of representative caeca at day 4 PI from naïve control (upper) and *S. Typhimurium* SL1344-infected (lower) mice. Tissue is oriented with colon to the left and small intestine to the right of the caecum.

3.3.3 Severe inflammation and leukocyte infiltration in infected caecal tissue

For microscopic examination of infected tissue, caeca were paraffin embedded, sectioned, and stained with hematoxylin and eosin. Inspection revealed that the severely inflamed caeca of wild type *S. Typhimurium*-infected mice displayed many of the histopathological hallmarks of the streptomycin mouse model (Figure 3.3). In naïve mice the

epithelial surface was smooth and continuous with clearly defined crypts of depth 100 - 200 μm . The lamina propria formed a relatively thin layer between the epithelial cells and surrounding muscle and contained few cells. Dramatic changes in the tissue architecture were visible at day 2 PI. Crypt length was increased to $> 500 \mu\text{m}$ and crypt architecture disorderly and uneven. The epithelial surface was rough and at the luminal interface damaged cells were breaking away from the tissue. The lamina propria was enlarged as a result of fluid accumulation and contained infiltrating cells. At day 4 PI the described changes were further developed with more extensive damage to epithelial cells, greater loss of crypt organisation, and more leukocytes within the lamina propria and crypts. Caeca of mice infected with SL1344 ΔInvA were highly similar in appearance to those infected with wild type SL1344.

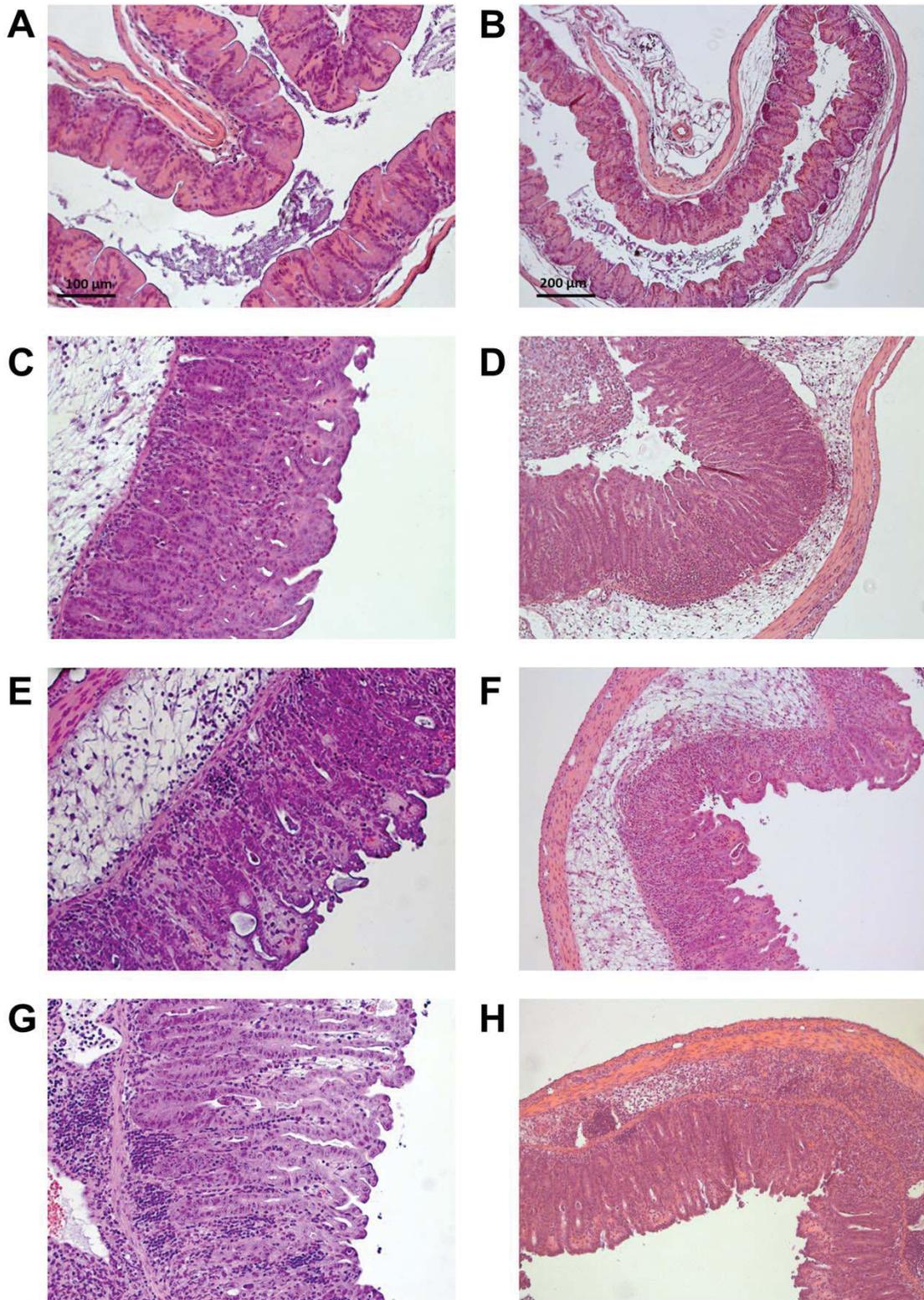


Figure 3.3. Inflammatory changes in *S. Typhimurium*-infected caecal tissue. Representative hematoxylin and eosin stained caecal tissue sections from: (A) and (B) naïve control mice at day 4 PI, (C) and (D) *S. Typhimurium* SL1344-infected mice at day 2 PI, (E) and (F) *S. Typhimurium* SL1344-infected mice at day 4 PI, (G) and (H) *S. Typhimurium* Δ *InvA*-infected mice at day 4 PI.

3.3.4 Changes in the indigenous microflora in response to streptomycin and *S. Typhimurium* infection

The potential effects of streptomycin treatment and infection with *S. Typhimurium* on the composition of the murine intestinal microbiota were investigated. Faecal samples were collected from cohoused mice immediately prior to streptomycin treatment and 24 h after treatment. At day 4 PI, colon content, caecal content, and caecal tissue were collected from PBS-treated naïve control and wild type SL1344-infected mice. In all sample groups n = 6. Figure 3.4A summarises the schedule for sample collection. 16S rRNA gene-based sequencing was performed on extracted DNA for quantitative comparison of microbial composition.

3.3.4.1 Microbiota samples cluster according to treatment group in microbial composition analysis

The PCA plot in Figure 3.4B displays the relationship between samples according to the first and second principal components, which account for 24.5% and 12.3% of the variance between the microbial communities of individual samples. In Figure 3.4C the relationship between samples according to their microbial compositions is displayed in a cluster dendrogram. Samples were divided between two major groups. One group contained faecal samples from untreated mice and intestinal content and tissue from naïve control mice at day 4. The other group contained faecal samples from mice 24 h post-streptomycin treatment and intestinal content and tissue from SL1344-infected mice at day 4. Within the first of these groups faecal samples from untreated mice clustered together (Figure 3.4B & C). Figure 3.4C illustrates that these samples are positioned between intestinal content and tissue samples from individual naïve control mice at day 4, indicating a highly similar microbial community structure in untreated mice and after a five day period of recovery following streptomycin treatment. In the second of the two major groups the 24 h post-streptomycin faecal samples formed a cluster separate from intestinal content and tissue from SL1344-infected mice at day 4. The bar plot of microbial community composition shows clear differences between the post-streptomycin and infected sample clusters.

3.3.4.2 Species of the indigenous mouse microflora are displaced by *Salmonella*

In material from naïve controls at day 4, colon content samples were interspersed between caecal tissue and caecal content. Conversely within SL1344-infected mice at day 4 colon content samples clustered apart from caecal content and tissue. In infected caecal content and tissue, *Salmonella* formed the greatest proportion of the microbial community (caecal content: $70.7\% \pm 34.3\%$, caecal tissue: $81.2\% \pm 5.7\%$), whereas in the colon *Salmonella* accounted for less than half of bacteria (colon content: $43.4\% \pm 11.9\%$) (mean \pm standard deviation). These findings indicate that the microbial community in the caecum and colon is not equally affected in infection.

3.3.4.3 Response to streptomycin and *S. Typhimurium* infection at the level of phyla

Much work has described the effects of intestinal infection and inflammation on the composition of the microbiota at the level of phyla. Previous studies of the streptomycin mouse model show common signatures of infection despite the use of different mouse backgrounds, streptomycin dosage and delivery, and carrying out work at different locations. Upon infection Proteobacteria are seen to be greatly expanded, both as a result of the introduction of *Salmonella* which belongs to this phylum and due to the expansion of other genera within the phylum. Conversely the phyla Bacteroidetes and Firmicutes are repressed by the presence of *Salmonella* [108, 139, 206]. Previously streptomycin treatment was shown to increase the proportion of Bacteroidetes and decrease the proportion of Firmicutes and other bacteria [206]. Five days was shown to be sufficient for recovery of the microbiota at the phylum level following a 20 mg dose of streptomycin [108].

Figure 3.4D shows the proportional abundance of the phyla Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria in the samples sequenced in this study. Contrary to the findings in [206] and [108] bacteria assigned to the phylum Bacteroidetes were reduced following streptomycin treatment, and the proportion of Proteobacteria was unchanged. In addition a previously unreported expansion of Actinobacteria was observed due to the genus *Olsenella*. In line with the work referred to above the vast expansion of Proteobacteria upon infection displaces the phyla Firmicutes and Bacteroidetes. We found that clear differences distinguish samples taken at five days following streptomycin treatment from those of untreated mice; Proteobacteria and bacteria outside the named phyla were elevated in the treated mice.

3.3.4.4 Changes in proportional abundance of specific genera in response to infection

Infection is associated with the appearance of bacterial genera below the limits of detection in untreated mice. For example the genus *Ralstonia* of the family *Burkholderiaceae* is below the detection threshold in untreated mice whilst in infection becomes a low frequency component of the microflora (caecal content: $1.8\% \pm 2.3\%$, caecal tissue: $2.9\% \pm 2.8\%$, colon content: $< 0.1\%$). Similarly the genus *Enterococcus* is present at very low frequency in untreated mice ($< 0.1\%$) and dramatically increased in infected samples (caecal content: $4.3\% \pm 3.8\%$, caecal tissue: $3.5\% \pm 3.9\%$, colon content: $4.9\% \pm 3.8\%$). A similar pattern was observed with the genus *Undibacterium* although the presence of this genus in infected samples displayed greater variability than *Enterococcus*.

The genera *Shigella* and *Escherichia* are closely related; based on genetic similarity they would form a single genus. However the historical division created to describe epidemiological and clinical differences persists. As a result of their close genetic similarity these genera are difficult to distinguish by 16S rRNA gene based sequencing methods [219]. Indeed, *Shigella* species originate from within the genus *Escherichia*. The genus *Escherichia* contains many natural commensal species as well as pathogenic strains. In contrast, *Shigella* is not reported to contain commensal species, although they can persist in asymptomatic children in developing countries [19]. It is therefore likely that the 16S rRNA gene sequences assigned here to *Escherichia*_ *Shigella* are *Escherichia*. Similar to *Ralstonia* and *Undibacterium* the group *Escherichia*_ *Shigella* expands in infection, however in contrast to these groups *Escherichia*_ *Shigella* were also detected in the flora of naïve control mice at day 4. These genera were undetected in untreated mice but in all other groups they were recorded at detectable levels, comprising $< 1\%$ of bacteria in 5 out of 6 mice at 24 h post-streptomycin, increasing to $4.1\% \pm 6.4\%$ in samples from naïve control mice at day 4, and $7.6\% \pm 11.4\%$ in samples from infected mice.

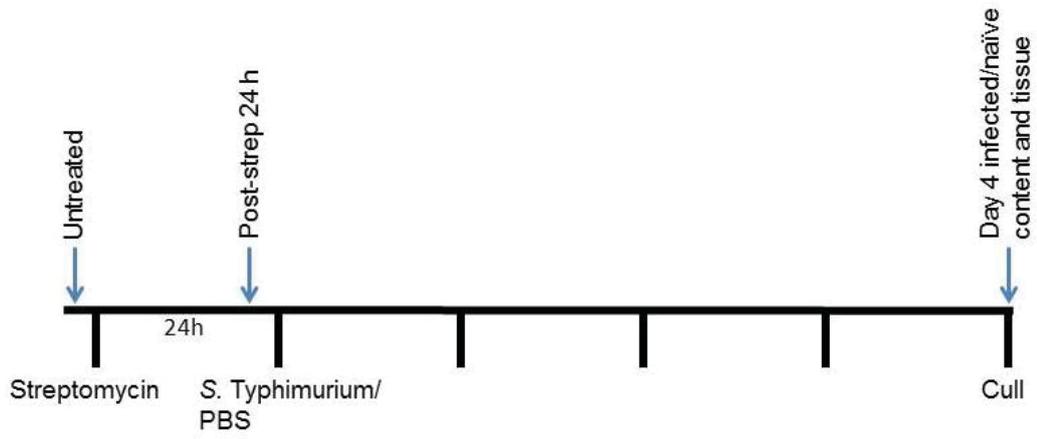
The genus *Olsenella* of the family *Coriobacteriaceae* displays an unusual sample distribution. In all untreated faecal samples and samples taken at day 4 from both naïve control and infected mice, *Olsenella* comprise $< 1\%$ of the sample microbiota. However in samples taken at 24 h post-streptomycin treatment *Olsenella* comprise a significant proportion of the microbiota ($32.7\% \pm 15.7\%$). The dramatic proportional increase in *Olsenella* shortly after delivery of streptomycin might be explained by a degree of antibiotic resistance in members of this genus, creating a strong selective advantage over susceptible members of the

microbiota. By day 4 PI the introduction of *Salmonella* and the regrowth of other commensals returns the proportional abundance of *Olsenella* to pre-treatment levels.

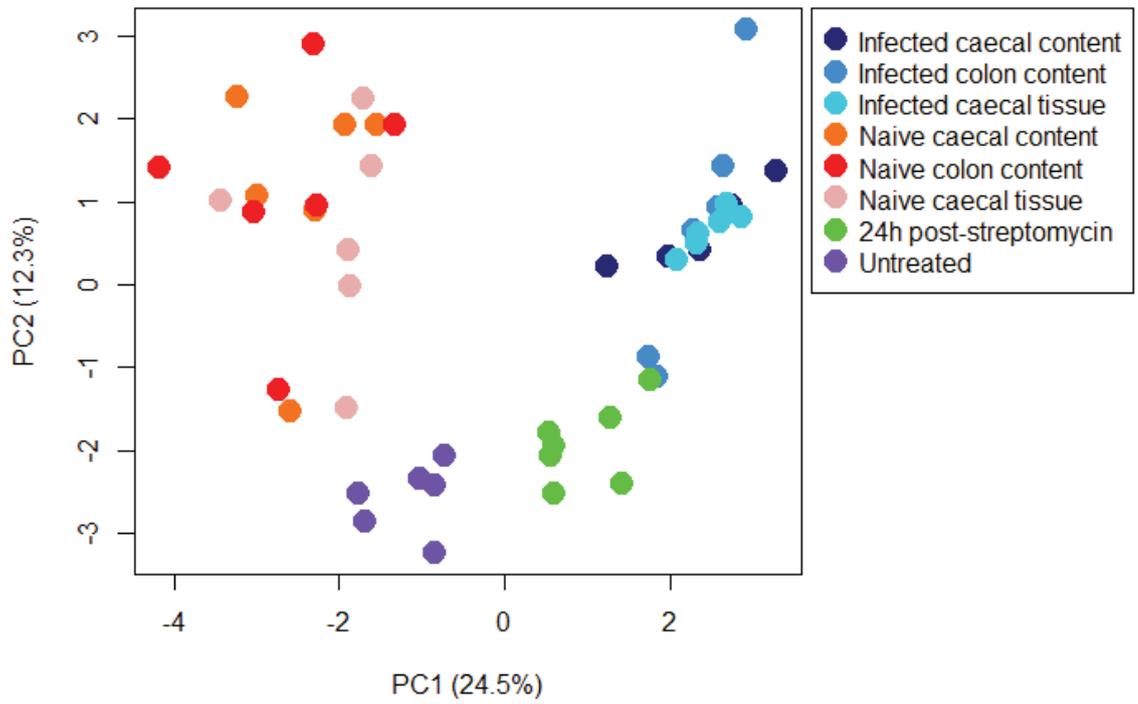
3.3.4.5 Effects of streptomycin treatment and *S. Typhimurium* infection upon microbial diversity

Previous work has shown that streptomycin treatment reduces the microbial diversity of the intestinal microflora [220]. The ‘within sample diversity’ (α -diversity) of microbial composition can be represented by the inverse Simpson diversity index. The inverse Simpson diversity index for the samples described throughout section 3.3.4 is displayed in Figure 3.4E. Untreated samples possessed the highest within sample diversity. Naïve control samples at day 4 were less diverse, followed by faecal material at 24 h post-streptomycin and finally with the lowest diversity the infected samples. The reduction of diversity upon streptomycin treatment is in agreement with [220]. Following five days of recovery post-streptomycin treatment, microbial diversity was significantly increased but remained lower than the untreated condition. In infected mice at day 4 PI the expansion of *S. Typhimurium* reduced the microbial diversity beyond that of streptomycin treatment alone.

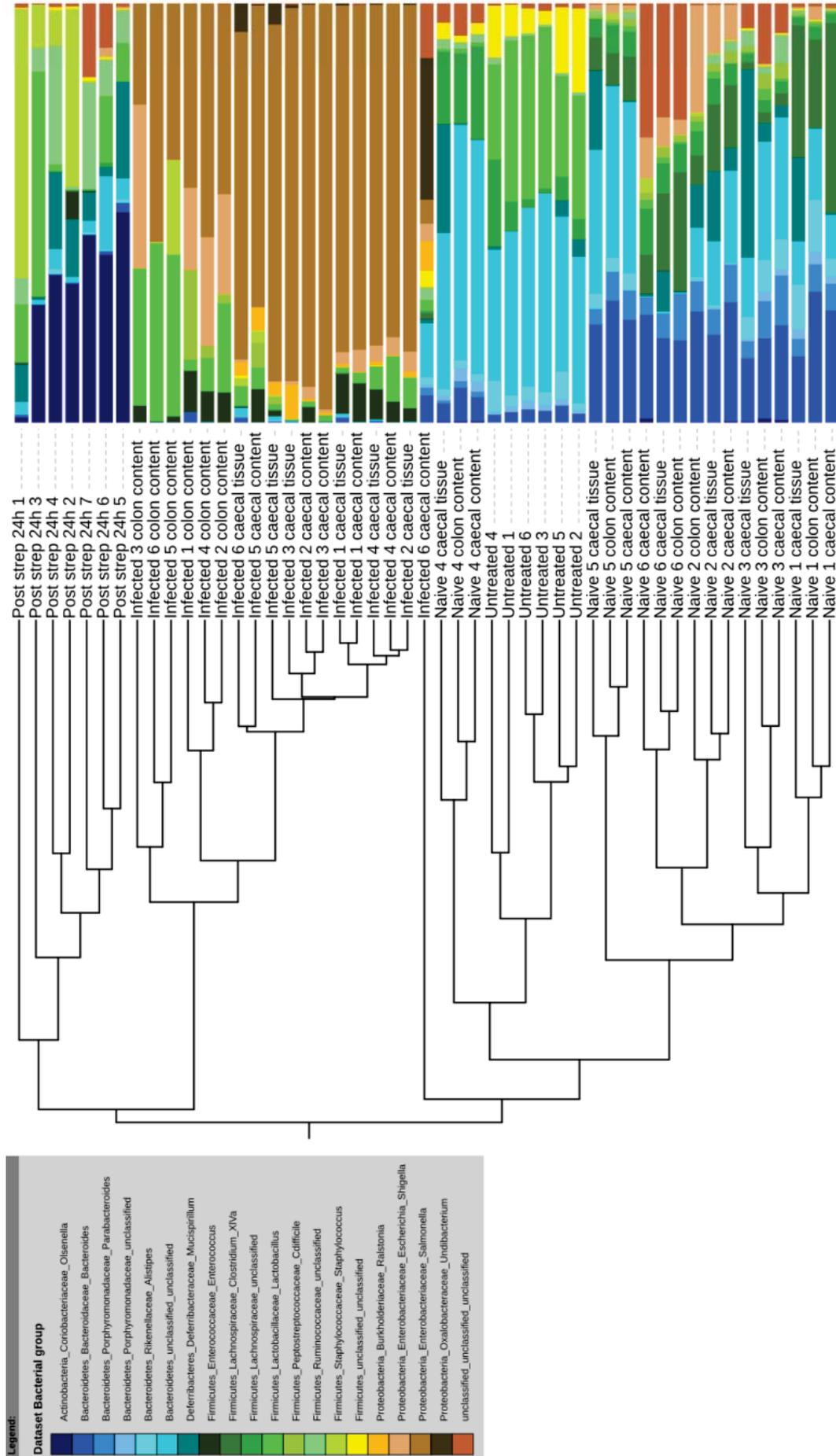
A



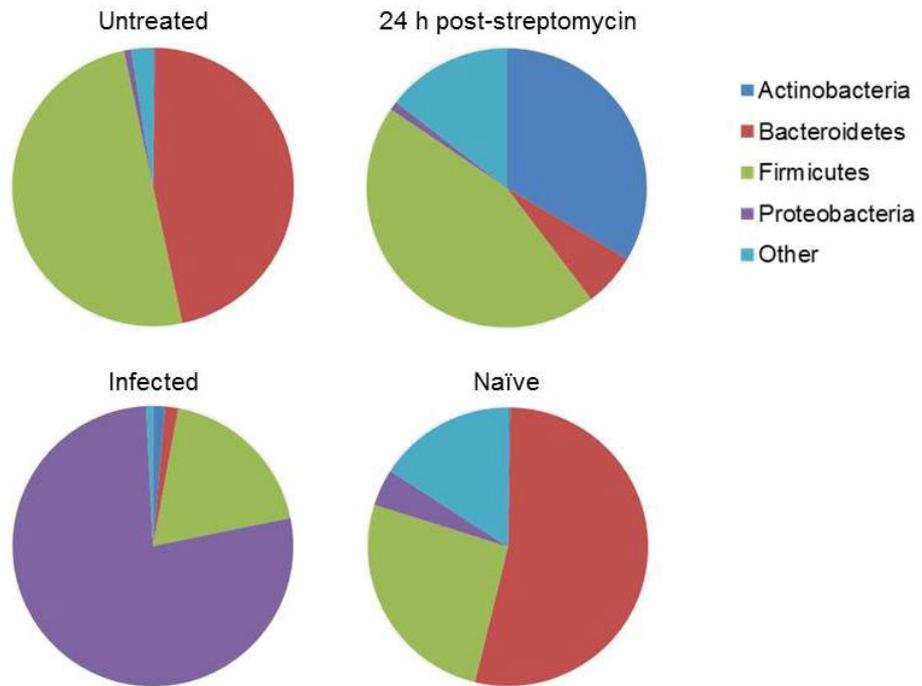
B



C



D



E

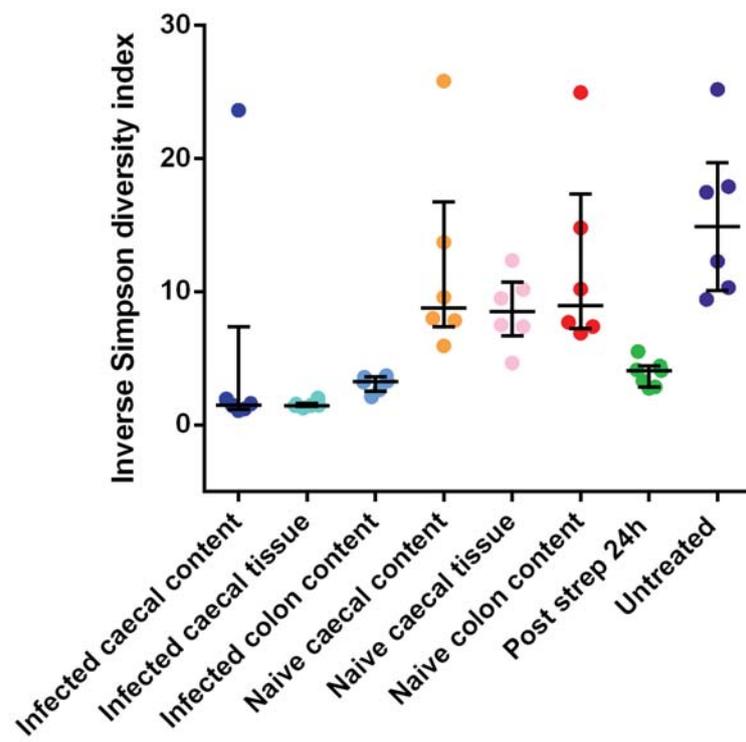
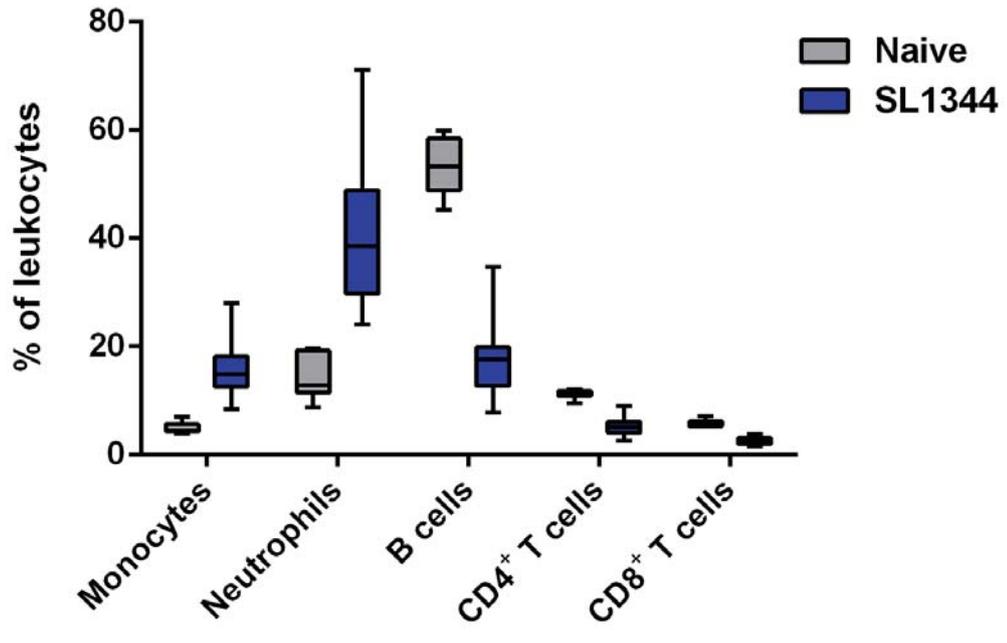


Figure 3.4. Effects of streptomycin treatment and infection with *S. Typhimurium* SL1344 on the community structure of the microbiota. (A) Timeline outlining the sample collection schedule for microbial community analysis. Faecal samples were collected from untreated mice and the group ‘24 h post-streptomycin’. At day 4 PI (or PBS treatment in naïve controls), mice were culled and samples of colon content, caecum content and caecal tissue taken for analysis. (B) PCA plot. Points indicate individual samples from the sample groups described in (A). (C) Cluster dendrogram of samples based upon bacterial community composition at the level of operational taxonomic units (OTU), with bar graph representing OTU proportional abundance. (D) Average proportional abundance of major bacterial phyla. ‘Infected’ and ‘Naïve’ pie-charts display averages for all samples in the ‘Day 4 infected’ and ‘Day 4 naïve (streptomycin-only)’ groups (colon content, caecal content and caecal tissue). (E) Inverse Simpson within sample diversity index.

3.3.5 Blood leukocyte populations respond to infection in the streptomycin mouse model

Blood was taken from SL1344-infected mice and naïve controls at day 4 PI for analysis of peripheral blood leukocyte composition. Proportional abundance of monocytes, neutrophils, B cells, CD4⁺ T cells and CD8⁺ T cells, in total leukocytes (CD45⁺ cells) was determined by staining for cell surface markers followed by flow cytometry (Figure 3.5). At day 4 PI the abundance of all five leukocyte populations tested were significantly changed in SL1344-infected mice compared with naïve controls, Mann Whiney U test ($p < 0.0001$). Proportional abundances of adaptive immune cells types (B cells, CD4⁺ T cells, and CD8⁺ T cells) were reduced in infected mice whilst innate immune cell types (neutrophils and monocytes) were proportionally increased. Unfortunately in the absence of total cell counts the relationship between the observed changes in cell proportions and changes in the absolute numbers of specific immune cell types is unknown.

A



B

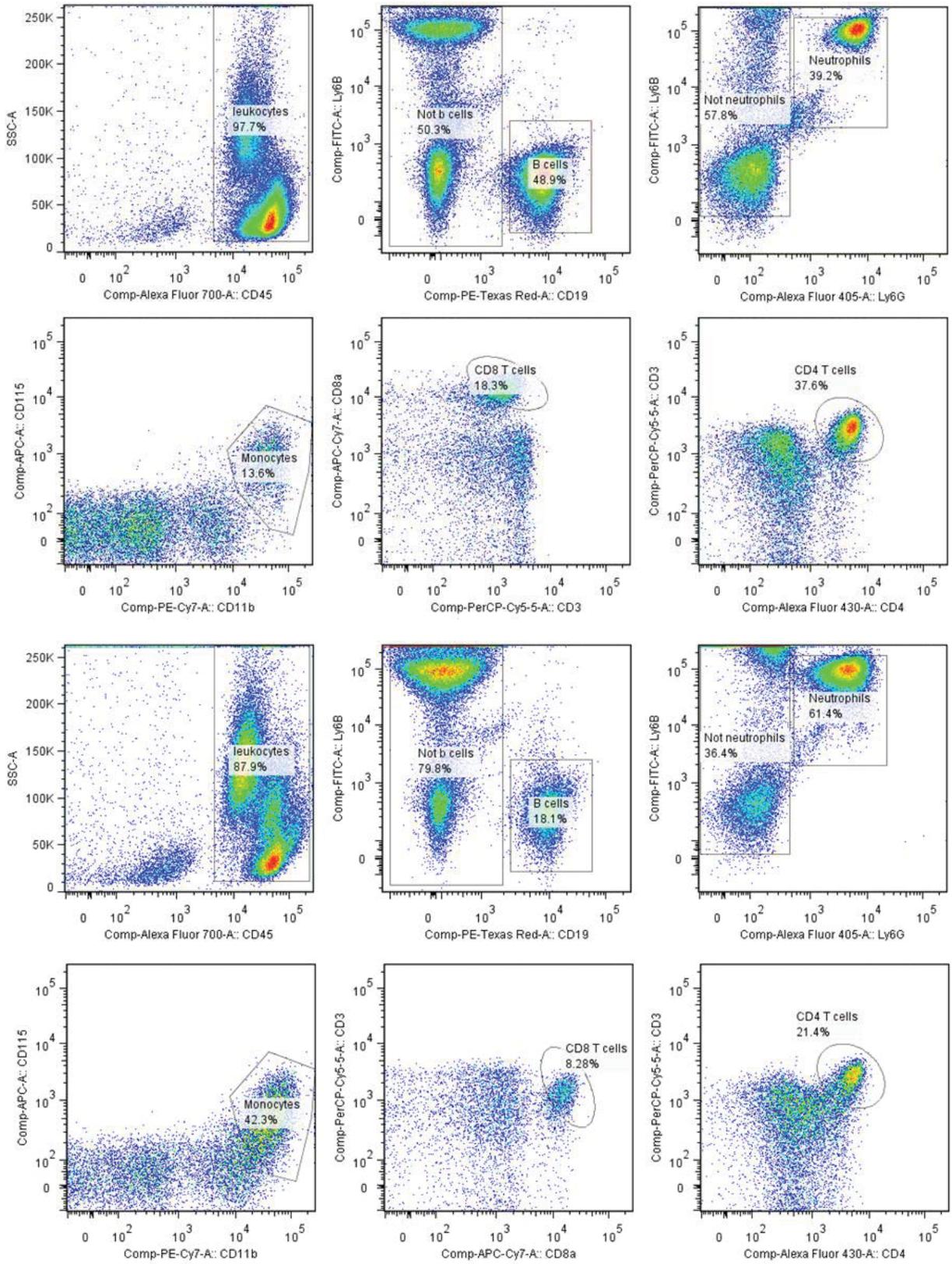


Figure 3.5. Flow cytometry analysis of leukocyte populations in peripheral blood during *S. Typhimurium* infection. (A) Box and whisker plot showing the median, interquartile range and range in proportional abundance of leukocyte subsets for naïve control and *S. Typhimurium*-infected mice at day 4 PI, minimum of 7 mice per group. (B) Representative density plots for a naïve control sample (upper six plots) and SL1344-infected sample (lower six plots). A sequential gating strategy was used as follows: CD45⁺ singlet events detected in red cell-depleted blood samples were gated as leukocytes. Of events in the leukocyte gate, CD19⁺Ly6B⁻ events were gated as B cells, and remaining events assigned to a non-B cell gate. Non-B cell events were assigned to either a neutrophil gate (Ly6B⁺Ly6G⁺) or non-neutrophil gate (all other events). Within the non-neutrophil gate three gates were defined: monocytes (CD11b⁺CD115⁺), CD8⁺ T cells (CD3⁺CD8a⁺) and CD4⁺ T cells (CD3⁺CD4⁺). The percentages shown for each gate are proportions of events in the parent gate (B and non-B cells as a proportion of lymphocytes, neutrophils and non-neutrophils as a proportion of non-B cells, and monocytes, CD8⁺ and CD4⁺ T cells as a proportion of non-neutrophils. Proportions are based on a minimum of 17,000 events/sample (mean = 47,576 events).

3.3.6 Transcriptional changes in caecal tissue during *S. Typhimurium*-induced inflammation

3.3.6.1 Naïve and infected caecal tissue produce distinct transcriptional profiles

To gain insight into transcriptional changes in the intestinal mucosa upon *S. Typhimurium* infection RNA was extracted from caecal tissue and sequenced. RNA profiles were generated for *S. Typhimurium* SL1344-infected, SL1344 $\Delta InvA$ -infected mice and naïve controls (n = 5). Principal component analysis demonstrates tight clustering of samples according to treatment group, the first principal component accounting for 78% of the variance between samples (Figure 3.6).

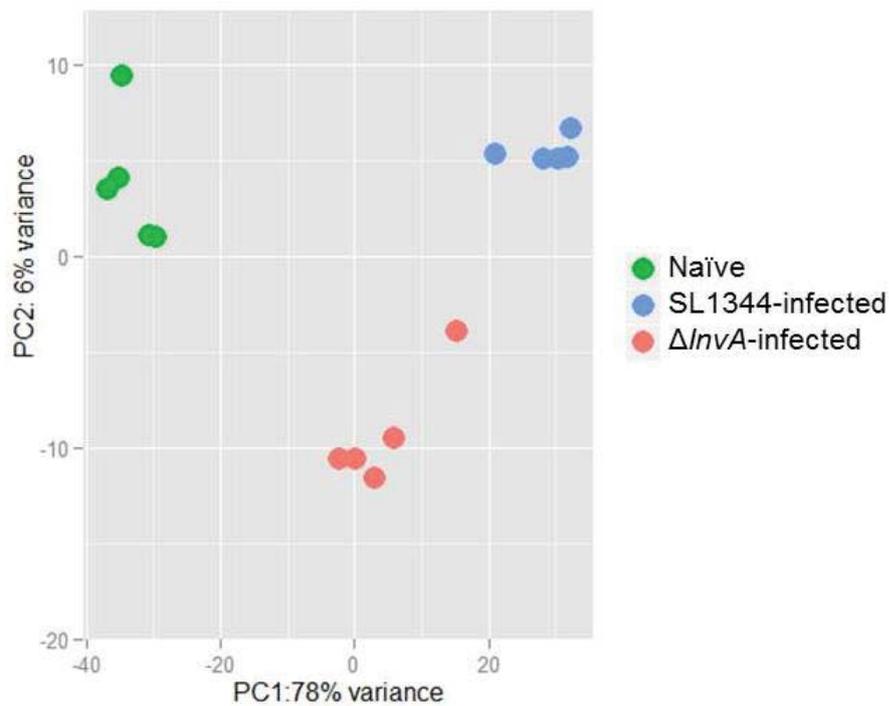


Figure 3.6. Principal component analysis of caecal transcriptional profiles. Points represent profiles from individual mice in the three treatment groups; naïve controls, *S. Typhimurium* SL1344-infected mice, and SL1344 $\Delta InvA$ *S. Typhimurium*-infected mice. Principal component analysis was performed using the 500 genes most variable across all samples.

Differential expression analysis was performed and genes selected for further analysis according to fold change and p-value thresholds: genes with a log₂ fold change of < -1 or > 1 and an adjusted p-value of < 0.05 were used. We detected a large number of differentially regulated genes which conformed to the described thresholds for effect size and significance: 3,599 genes were found to be upregulated and 2,764 downregulated in infection with *S. Typhimurium* SL1344 compared to uninfected controls. As might be expected fewer genes exceeded the thresholds in SL1344 $\Delta InvA$ infected mouse tissues; 1,997 were upregulated and 1,614 downregulated in caecal tissue taken from mice infected with this attenuated SL1344 derivative. Comparison of the transcriptome of SL1344-infected samples with SL1344 $\Delta InvA$ -infected samples gave rise to the smallest number of differentially regulated genes; 1,070 upregulated and 889 downregulated genes. A visual representation of the comparisons between transcriptomes is provided in Figure 3.7A, and the distribution in gene expression changes between treatment groups are also summarised in Figure 3.7B. The distribution in fold changes is similar for each of the three comparisons made, the largest differences in

expression between naïve and SL1344-infected samples. The most highly regulated genes overall were downregulated upon infection.

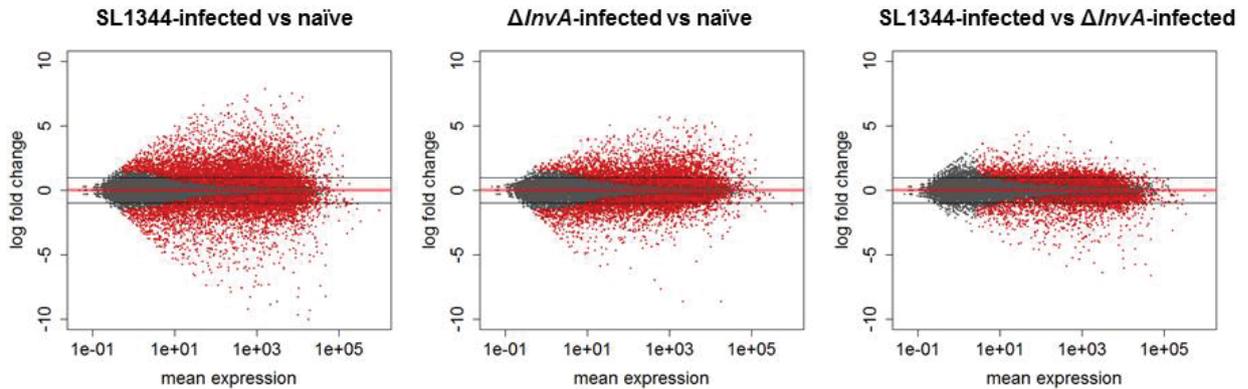
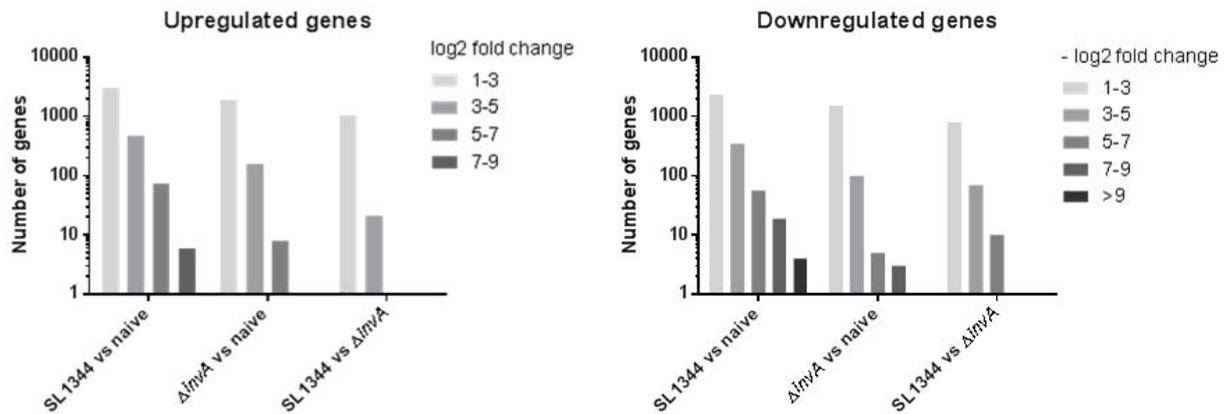
A**B**

Figure 3.7. Infection with *S. Typhimurium* remodels the transcriptome in the mouse caecum. (A) MA plots providing an overview of transcriptome-scale differences in gene expression. Log₂ fold change is plotted against normalised transcript abundance; each point represents a single gene. Points coloured red correspond to genes significantly regulated at a 10% FDR and grey points represent genes which show no significant change between treatment groups. Horizontal lines indicate the log₂ fold change thresholds of ≤ -1 and ≥ 1 applied for selection of genes for further analysis. The DESeq2 package used to produce these plots moderates log₂ fold changes from genes with low or highly variable counts as seen by the narrowing spread of points toward the y axis. (B) Distribution in the fold changes of DE genes (log₂ fold change of < -1 or > 1 , adjusted p-value of < 0.05).

3.3.6.2 Biological functions of genes most highly regulated in *S. Typhimurium* infection

The top 30 most differentially regulated genes are shown in Figure 3.8. Upregulated genes in the top 30 were lipocalin-2, CXCL5 (also known as ‘neutrophil activating protein’), S100a8 and S100a9, matrix metalloproteinase 8, cystatin A1 and neutrophilic granule protein.

The remainder of genes were downregulated. Many of the most highly upregulated genes are widely reported to be involved in response to infection, for example the role of lipocalin 2 was described in section 1.3.6.3, and proteins of the matrix metalloproteinase family play numerous roles in infection including remodelling of the extracellular matrix, cytokine processing and leukocyte recruitment [221, 222]. However it is interesting to note that of the seven upregulated transcripts listed here, none appear in a catalogue of 511 genes upregulated in the ‘common host response’ generated by comparison of transcriptomic data from multiple cell types and infections [223]. Therefore our study identifies transcriptional responses in intestinal tissue directed more specifically to respond to *S. Typhimurium*, in addition to more general signatures of infection. qPCR analysis of a selection of eight genes observed by RNAseq to be upregulated in infection confirmed the upregulation of these genes. However the order of the eight genes by fold change magnitude was only moderately well conserved between the two techniques, with some considerable differences in actual fold change values.

Amongst the most downregulated genes in infection were many enzymes without apparent evidence of previous links to infection, involved in processes as diverse as carbohydrate modification (*A4gnt*), biosynthesis of steroid hormones (*Hsd3b2*) and fatty acid metabolism (*Cyp2c55*). Aquaporin 8, a member of a larger family of aquaporin proteins, downregulated in our dataset, was found previously to undergo downregulation in DSS colitis [224]. Other members of the aquaporin family have been implicated in diarrhoeal disease; cellular localisation of aquaporins 2 & 3 is reported to change during infection with *C. rodentium*, potentially contributing to diarrhoea [225]. Four genes of the gasdermin C family and related gene gasdermin C-like 2 are amongst the most downregulated genes in infection. The gasdermin superfamily was discovered relatively recently [226]. Expressed predominantly in epithelial cells, the functions of this family are poorly understood. The human genome encodes a single gasdermin C gene whilst mice have four paralogues. Expression of human gasdermin C has been shown in the middle to upper region of the epithelium of the oesophagus, small intestine and colon, and it has been suggested to perform functions related to cell mobility, but the distribution and functions of the mouse paralogues are not known [227]. Also among the most downregulated transcripts were those transcribed from two predicted genes; *Gm6086* and *Gm1123*. Such dramatic regulation in infection suggests these genes are part of pathways highly relevant to intestinal infection and inflammation; our lack of knowledge of these genes demonstrates the need for more work in this area.

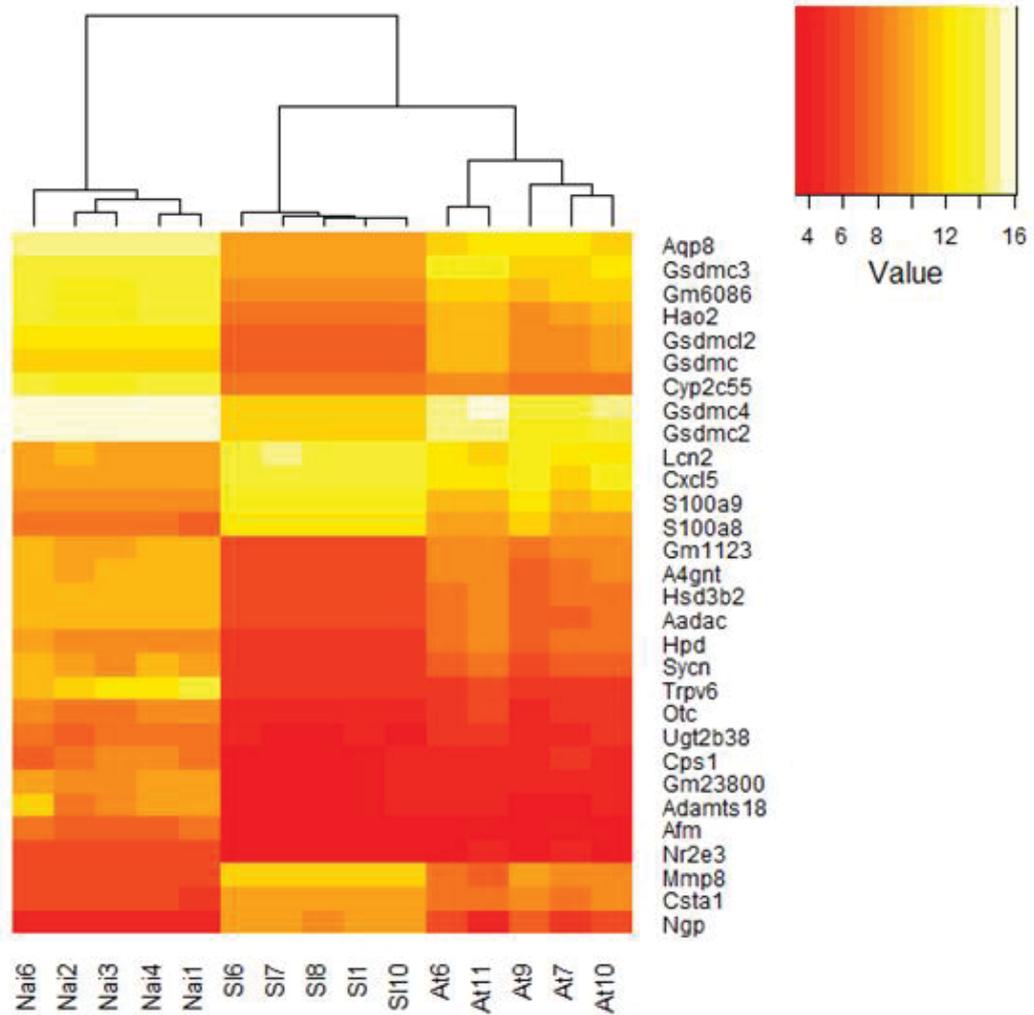


Figure 3.8. Expression profiles of the 30 most differentially expressed genes during *S. Typhimurium* infection of caecum. Heatmap of normalised transcript abundance in caeca of naïve control mice (samples with prefix ‘Nai’), *S. Typhimurium* SL1344-infected mice (prefix ‘S1’) and mice infected with the attenuated SL1344 $\Delta InvA$ derivative (prefix ‘At’), as determined by RNAseq.

3.3.6.3 Pathway analysis of transcripts regulated in infection with *S. Typhimurium*

The InnateDB online analysis tool was used to identify pathways significantly enriched for genes regulated in infection with wild type *S. Typhimurium*. Transcriptionally upregulated genes (\log_2 fold change > 1 , adjusted p-value < 0.05) and downregulated genes (\log_2 fold change < -1 , adjusted p-value < 0.05) were uploaded and pathway overrepresentation analysis (ORA) performed. Pathways in the Reactome database significantly associated with regulated genes are summarised in Figure 3.9 and listed in full with accompanying p-values and numbers of genes in Appendices 1 & 2. The top 10 gene

ontology (GO) terms associated with regulated genes in each of the three categories (biological process, molecular function, and cellular compartment) are listed in Appendix 3. Large numbers of pathways are statistically associated with the genes dysregulated in *S. Typhimurium* infection; with a p-value of < 0.05 after correction for multiple testing are 92 Reactome pathways associated with upregulated genes and 88 associated with downregulated genes.

Many pathways associated with upregulated genes have an established role in immunity, and of the 884 genes in InnateDB annotated to the pathway ‘immune system’ over one quarter are upregulated in the RNAseq dataset for *S. Typhimurium*-infected caecal tissue. Signalling pathways involving cytokines, chemokines, G protein alpha-i, G protein coupled receptors (GPCR), programmed cell death protein 1 (PD-1), TLRs, interleukins and the T cell receptor (TCR), are among pathways highly associated with upregulated genes. More general signalling-related pathways; including ‘generation of second messenger molecules’ and ‘signal transduction’; are also highly associated with upregulated genes. Interspersed between pathways involving immune cells and signalling are pathways of broader function such as ‘platelet activation, signalling and aggregation’, and ‘extracellular matrix organisation’. These represent ‘secondary processes’ occurring in infection; while many immune pathways are targeted to directly attack bacteria, other pathways remodelling damaged tissue and preventing blood loss through damaged vessels are also important.

GO terms significantly associated with genes upregulated in *S. Typhimurium* infection are similar and complementary to the Reactome pathways. Associated molecular functions include many terms related to cytokines and receptor binding whilst the most highly associated biological processes include activities involved in recruiting effector cells to the site of infection. Cellular component terms associated with regulated genes indicate sites where activities are most affected by infection. Locations most associated with upregulated genes are outside the cell in the extracellular space and the cell surface at the plasma membrane. These sites are in accordance with the described association of upregulated genes with receptor signalling, much of which occurs at the cell surface.

Downregulated genes are associated with large numbers of pathways involved in metabolism; 275 genes involved in metabolism are downregulated in the RNAseq dataset for *S. Typhimurium*-infected caecal tissue. Downregulated metabolic pathways involve diverse metabolites including lipids and components of lipids, lipoproteins, amino acids, ketone

bodies and pyruvate. Together metabolite transport and metabolic pathways comprise the majority of downregulated pathways.

The results of ORA for gene ontology (GO) terms support the dramatic downregulation of transcripts encoding proteins involved in metabolism; ‘metabolic processes’, ‘oxidation-reduction processes’ and ‘lipid glycosylation’ are the three most highly associated biological processes. In addition to metabolic and transport processes the top 10 includes the ‘steroid hormone mediated signalling pathway’, previously reported to be an important pathway downregulated in a metabolomic study of the murine typhoid model [228]. Presence of ‘response to starvation’ amongst the downregulated metabolic pathways indicates some of the metabolic changes occurring in *S. Typhimurium* infection may be similar to changes associated with nutrient restriction. However the process ‘response to starvation’ contains only a relatively small number of genes (35, of which 14 are downregulated) and therefore support for this hypothesis is limited.

Molecular functions significantly associated with downregulated genes include the general term ‘catalytic activity’ and many more specific enzymatic processes. The cellular component most highly associated with downregulated genes is the mitochondrion; 247 downregulated genes are annotated to this term identifying mitochondria as a site where normal cellular activities are altered extensively during infection. Other associated compartments inside cells include peroxisomes and the endoplasmic reticulum membrane. Peroxisomes are important metabolic organelles with most notable functions in β -oxidation of fatty acids, one of the pathways significantly associated with downregulated transcripts. Also associated with downregulated genes is the brush border; the array of microvilli which protrude from the apical surface of intestinal epithelial cells. As for upregulated genes the plasma membrane is associated with downregulated genes also, representing a site of complex changes during infection.

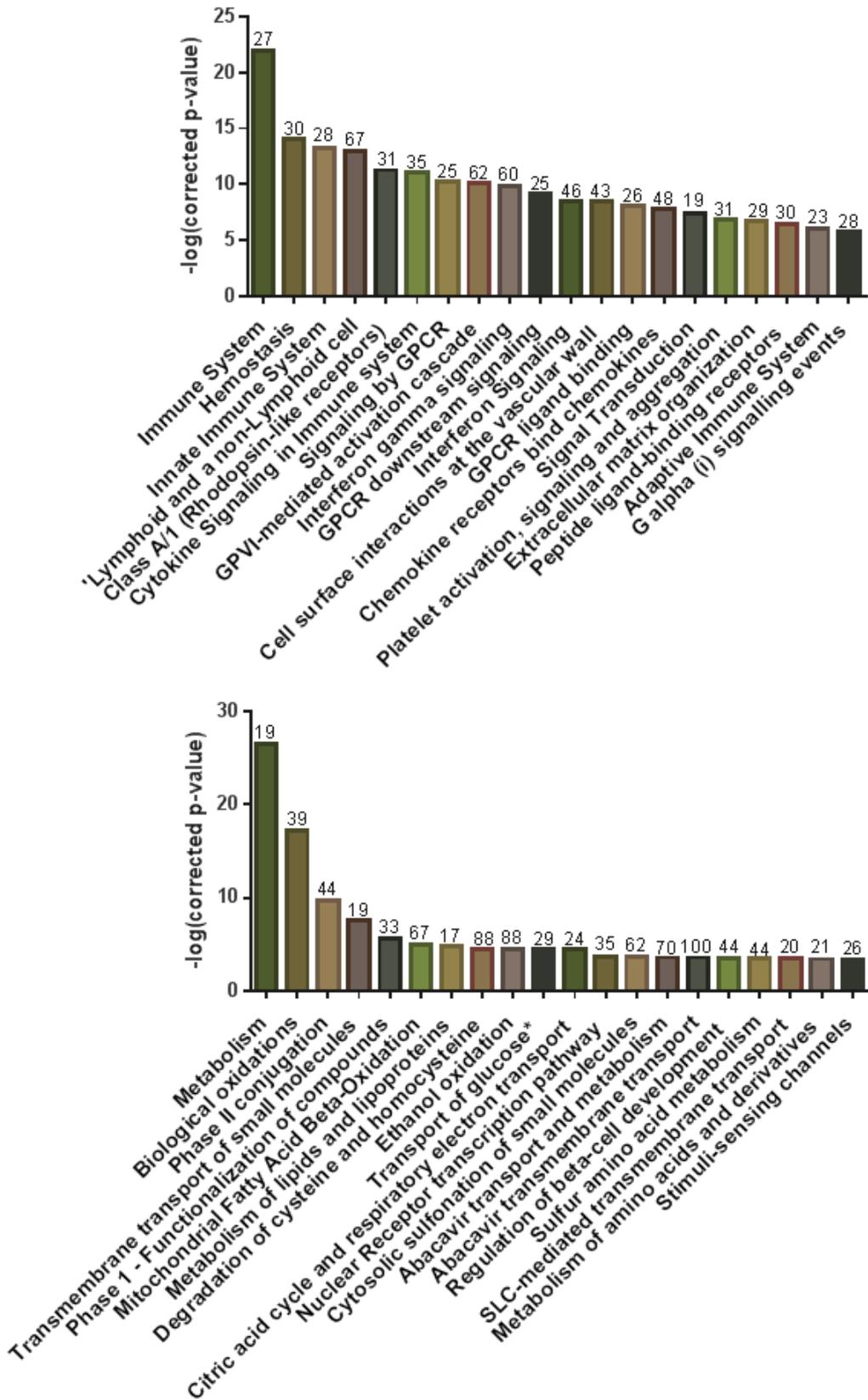


Figure 3.9. Pathway ORA of genes regulated in *S. Typhimurium* infection. The InnateDB pathway analysis tool was used to identify pathways annotated with genes regulated in infection, and to perform statistical

analysis to determine pathways significantly enriched with regulated genes. Pathways in this figure are from the Reactome database. Upper - The 20 most significantly associated pathways for genes upregulated in wild type *S. Typhimurium* infection. Lower - The 20 most significantly associated pathways for genes downregulated in wild type *S. Typhimurium* infection. Bar height is the $-\log$ p-value for the pathway association with the regulated gene set. Numbers above bars denote the percentage of genes annotated to the pathway which were up- or down-regulated in the dataset. Two pathway names have been shortened: 'Lymphoid and a non-Lymphoid cell = Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell, and Transport of Glucose* = Transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds.

3.4 Discussion

In the work described in this chapter we demonstrate replication of reported weight loss, and *S. Typhimurium* colonisation of gastrointestinal tissues and dissemination to systemic organs in C57BL/6 mice following oral treatment with streptomycin and infection with wild type *S. Typhimurium*. We observed that a mutant derivative of *S. Typhimurium* defective in the SPI-1 T3SS machinery, SL1344 $\Delta InvA$, was able to disseminate to extra-intestinal sites with efficiency similar to the wild type SL1344. However, mice infected with SL1344 $\Delta InvA$ displayed reduced intestinal organ counts with minimal weight loss. Histopathological analysis determined that SL1344 $\Delta InvA$ infection resulted in obvious intestinal pathology, although more extensive analysis would have determined if the severity is statistically indistinguishable from infections with wild type *S. Typhimurium*.

A previous study to investigate the effect of inactivating mutations in the SPI-1 T3SS using SL1344 $\Delta invG$ in the streptomycin mouse model produced contrasting results [149]. Specifically at 48 h PI bacterial loads in intestinal content and liver did not differ significantly between SL1344 $\Delta invG$ and wild type infection, however lower loads of SL1344 $\Delta invG$ were detected in spleen, and colitis induced by the attenuated strain was less severe. The use of an earlier time point in this study (48 h PI) is noteworthy. These results led the authors to conclude SPI-1 is important to establish systemic infection and elicit profound inflammation, but not for intestinal colonisation. In contrast with the findings in [149] our results suggest the capacity for invasion is non-essential for induction of intestinal inflammation and systemic dissemination, yet intestinal colonisation is enhanced for *Salmonella* with ability to invade intestinal cells. The relative importance of entry of enterocytes and M cells as the initial step in the transition of *Salmonella* from a luminal to an intracellular existence has been debated, however both routes have been found to require SPI-1 [116]. Therefore our finding of

comparable numbers of wild type and $\Delta InvA$ *Salmonella* at systemic sites is interesting. At least two possibilities could contribute to explain our results. The first is an important role for phagocytic cells in the initial uptake of *Salmonella* in the streptomycin-treated intestinal environment; many previous studies were performed in mice with an intact microbiota [114]. Second, the later time point used in our study compared with [149] might provide time for extensive replication of bacteria at systemic sites, swamping any initial differences in translocation of bacteria from the gut.

We detected lower numbers of *S. Typhimurium* SL1344 $\Delta InvA$ in the colon and caecum. Organ bacterial counts were determined by plating tissue homogenates following removal of intestinal content by physically scraping intestinal content from the tissue surface. Due to the large numbers of *Salmonella* residing in the intestinal lumen compared to those invading the tissue it is possible our results do not truly reflect the difference in tissue-associated bacteria, and counts representing the combination of tissue-associated bacteria with those in the mucus layer and luminal content may be valuable. The microbiota analysis in section 3.3.4 shows that in naïve control mice at day 4 PI the microbiota has begun to recover from streptomycin treatment. We suggest a small reduction in the severity of inflammation caused by SL1344 $\Delta InvA$ results in an environment less toxic, or more favourable to the natural microflora compared with wild type-infected mice. The slightly lessened inflammation may tip the balance between *Salmonella* and the microbiota a little more in favour of the microbiota relative to mice infected with wild type SL1344. At five days post-streptomycin treatment there may be some recovery or even reintroduction of natural flora in SL1344 $\Delta InvA$ -infected mice resulting in relatively lower numbers of *Salmonella* in the lumen. 16S rRNA gene sequencing of SL1344 $\Delta InvA$ infected mice would be required to investigate this hypothesis.

The findings described in this chapter advance knowledge gained from several earlier efforts to describe the effects of streptomycin and *S. Typhimurium* infection on the microbiota. A previously described reduction in microbial diversity upon streptomycin treatment was reproduced in this work. Further we showed that whilst microbial diversity is partially restored after a period of recovery, delivery of *S. Typhimurium* to streptomycin pre-treated mice reduces diversity below that in mice 24 h post-streptomycin treatment.

In terms of microbial composition, we observed faecal matter sampled at 24 h post-streptomycin treatment is more closely related to tissue and intestinal content from infected

mice than faeces from untreated mice and naïve controls, indicating that five days post-streptomycin treatment the microbiota is making significant progress on the path to recovery of the pre-treatment state.

Whilst in naïve control mice the microflora of the colon and caecum appear indistinguishable, the separation between content of infected colon and caecum shown in Figure 3.4C indicates these tissues do not respond equally to infection. Indeed the counts data presented here suggest the infectious burden of the caecum is greater, and previous work found that *S. Typhimurium*-associated inflammation is more severe in caecum [149].

In a previous study using a similar methodology to investigate the effects of streptomycin and *S. Typhimurium* infection on the microbiota approximately 100 sequences per animal were generated [108]. In this study we sequence on average $48,068 \pm 20,647$ clusters generating an average of $1,969 \pm 36.4$ counts per sample following alignment to bacterial rRNA V1-V2 gene regions. Consequently this vast increase in depth produces a more comprehensive description of the microbiota composition and allows us to detect changes in low frequency microbes such as *Ralstonia* and *Undibacterium*. Previous work identified that along with *Salmonella* other members of the Proteobacteria are increased, and our work uncovers the identity of some of these genera. However, there is currently an absence of literature exploring the functional importance of genera such as *Olsenella*, *Ralstonia* and *Undibacterium* in the microbiota, and their relation, if any, to disease.

The analysis of blood leukocyte composition by flow cytometry showed the relative proportions of circulating innate and adaptive immune cells are changed upon *S. Typhimurium* infection following streptomycin pre-treatment. Neutrophils and monocytes were increased in proportion following infection while B and T cells were reduced. Total leukocytes in blood were not quantified therefore absolute numbers of cells cannot be compared, however it is likely that expansion of monocytes and neutrophils in response to activating cytokine signalling is responsible for a decrease in the relative proportion of adaptive cells, rather than absolute numbers. It would be interesting to repeat this experiment with immune cells isolated from caecal tissue to investigate the relationship between tissue and blood leukocyte profiles.

RNAseq was used to determine the effect of *S. Typhimurium* infection on RNA species present in caecal tissue. The largest number of significantly regulated genes was

detected between naïve control and wild type *S. Typhimurium*-infected sample groups. Fewer genes were differentially regulated between the naïve control and *S. Typhimurium* $\Delta InvA$ -infected sample groups, and least of all between the two infection groups. The smaller effect of SL1344 $\Delta InvA$ infection on caecal tissue at the transcriptional level is in line with the attenuated virulence of this derivative demonstrated in section 3.3.1. However the greater number of genes differentially regulated between naïve control and *S. Typhimurium* $\Delta InvA$ -infected sample groups compared with differentially expressed transcripts between the two infection groups indicates that the wild type- and SL1344 $\Delta InvA$ -infected tissues do share many transcriptional responses to *Salmonella*. This is in line with the observation that SL1344 $\Delta InvA$ infection induces substantial inflammation in histopathological analysis although suggests the severity of inflammation is less than that induced by the wild type at the molecular level.

Examination of the most differentially regulated genes in infection provided some insight into major aspects of the host response, particularly for upregulated genes. For example the central importance of neutrophils in attacking *Salmonella* is supported by the presence of CXCL5 for neutrophil recruitment, and neutrophilic granule protein. Examination of individual downregulated genes was less immediately informative as many of the genes were not previously reported to be linked with infection and/or encode enzymes, which function in multiple interconnected complex pathways. Pathway analysis was used to condense lists of dysregulated genes into smaller numbers of significantly associated pathways, with pathways providing clearer links with biological functions. Associated with the 3,599 upregulated and 2,764 downregulated genes were 92 up- and 88 down-regulated pathways respectively.

Metabolic pathways were identified as a major group associated with genes downregulated in infection. Knowledge of the effects upon host metabolism during infection has existed for decades. Early studies of metabolic changes upon infection of cells in culture and altered metabolism in patients with highly progressed systemic infections were superseded by controlled studies of human infections, and most recently studies utilising technology for metabolomic profiling of samples from infected animals or patients [228, 229].

However despite the dramatic transcriptional regulation of genes annotated to metabolic pathways detected here in *S. Typhimurium*-infected caecum there appears little previous effort focused on investigating tissue-localised changes in metabolic pathways

during infection, or little attempt to explain why these changes occur, what signals induce them, and how they benefit or harm both the host and invading bacteria [213]. The nutrient constitution of the luminal content in infection is likely to be affected by changes in food intake, changes in the rate of passage of ingested material through the intestinal tract, and changes in the microbial species present in the gut and their metabolic activities. Nutrient absorption by the host is likely to be decreased as a result of damage to the gut occurring in infection. On this basis the vast regulation of genes encoding metabolic functions might be considered unsurprising, though the pathways involved and the long term effects of such changes merit investigation. Some metabolic processes described previously as altered in systemic infection are amongst those associated with downregulated genes in caecal tissue. For example studies in animal models of infection and inflammation using endotoxin or proinflammatory cytokines have led to the description of major changes in lipid and lipoprotein metabolism as part of the acute phase response (APR), with adipose tissue lipolysis and *de novo* fatty acid synthesis generating an increase in plasma triglycerides. On the one hand these processes have been suggested to provide fuel to meet the high energy demand of responding to infection, whilst on the other there is evidence that lipoproteins are a component of innate immunity; it is possible both may be important. Reported innate immune functions of lipoproteins include directly binding LPS, and prevention of uncontrolled cytokine activation by inhibition of LPS-stimulated NF κ B activation by oxidised lipids [230, 231]. It will be important to examine how previously described systemic effects and tissue-localised pathways are related. 90 genes downregulated in *S. Typhimurium* infected tissue are annotated to the pathway ‘metabolism of lipids and lipoproteins’; whether these changes are linked with those described in response to systemic LPS requires more detailed investigation.

An interesting finding from examination of pathways associated with genes regulated in *S. Typhimurium* infected tissue is the presence of multiple directly opposing pathways apparently regulated in the same direction. For example pathways associated with upregulated genes include ‘response to IFN γ ’ and ‘positive regulation of IFN γ production’, in addition to ‘negative regulation of IFN γ production’. A similar pattern is seen also for IL12. These cases could reflect activation of pathways specifically targeted to limit the damage incurred by host tissues due to proinflammatory cytokine responses. Alternatively, they may indicate changes in cell populations or activation states.

Whilst pathway analysis has become a popular tool for interpretation of transcriptomic data it is important to note some of its shortcomings. Firstly is the bias toward greater significance for more general pathways. More general pathways have more genes annotated to them than highly specific processes, and consequently achieve statistical significance more easily than pathways of smaller numbers of genes. More specific pathways in which all genes were regulated, such as ‘negative regulation of IL12 production’ with 12 genes at 40th in the list, and ‘positive regulation of T-helper type 1 immune response’ with 8 annotated genes at 41st, are lower in significance than larger pathways with only a small fraction of genes regulated.

Annotation of pathways is an ongoing process and naturally results of pathway analysis are biased by the degree of annotation of the regulated genes to pathways. In addition pathway ORA assumes that both genes and pathways behave independently, however these assumptions are incorrect; transcriptional regulation of genes is highly interconnected, and genes contribute to multiple different pathways [214]. This pattern of genes overlapping multiple pathways explains the presence of pathways such as ‘defence response to protozoan’ and ‘response to virus’ amongst pathways highly associated with upregulated genes.

A limitation of the transcriptional analysis described in this chapter is the lack of spatial detail in the gene expression changes; all fold changes are an average of changes occurring throughout the entire tissue from the epithelium through to the serosa. The tissue contains a vast array of different cell types, ranging in frequency from rare immune cells through to the abundant cells of muscle and vasculature. Some transcriptional changes in infection may be restricted to a single cell type or small family of cells while others are widespread. Due to the averaging of expression across the entire tissue, dramatic changes in rare cells may become hidden by constant background expression in the other cells types, and their functional significance unappreciated. Similarly transcriptional changes occurring in opposite directions in different cell types would not be detected. Consequently averaging transcriptional changes across all regions of the tissue may cause a considerable loss of information.

Transcriptional changes serve as an indicator of changes in cellular activities; however there exist a large number of post-transcriptional regulatory mechanisms which impact the amount and activity of the proteins which they encode. Consequently protein levels are a more reliable indicator of activity than RNA transcript abundance. In the following chapter a

novel mass spectrometry method was used for quantitative profiling of protein in caecal tissue in order to address this.

4 Proteomic analysis of tissues from the streptomycin mouse model and integration with transcriptomic data

4.1 Introduction

4.1.1 Label-free mass spectrometry for large scale tissue proteomics

Advances in the field of proteomics have vastly broadened potential applications in recent years, moving beyond simple protein identification to quantitative profiling of complex protein mixtures [232]. Early quantitative proteomic analysis involved two-dimensional gel separation of protein mixtures, with quantitation performed by comparison of stained protein spot volumes prior to protein identification by MS. Current technology permits quantitation at the MS level, giving rise to vast increases in specificity and accuracy, and allowing rapid analysis of large numbers of proteins. The two major approaches for quantification are stable isotope labelling and label-free analysis. Prior to recent advances stable isotope labelling achieved more accurate quantitation. In this approach separate samples labelled with amino acids containing different isotopes are analysed in a single MS run. However highly reproducible high pressure liquid chromatography (HPLC) systems and mass spectrometers have now been developed which allow highly accurate quantitation between separate runs [233]. Isotope labelling is expensive compared with label-free sample preparation and labelling strategies are unsuited to the analysis of tissue from whole organisms, therefore a label-free approach was employed in the proteomic analysis described herein. Figure 4.1 outlines the approach used for the analysis of the murine caecal proteome.

In shotgun proteomics proteins are digested into peptides which are then identified by MS, and the resultant catalogue of peptides is compared against a reference proteome to allow piecing back together of the original proteins in the sample. Analysis is most commonly performed by ‘data-dependent acquisition’ (DDA) in which precursor ions are selected for fragmentation inside the mass spectrometer on the basis of their abundance, and only a fixed number of precursor ions recorded in a survey scan are selected for fragmentation to determine peptide sequence. In this approach precursor ion selection is stochastic and a large proportion of the peptides present are not sampled. Therefore a DDA approach is not well-suited to the analysis of complex proteomes of whole tissue extracts. Data-independent acquisition (DIA) approaches utilise an unbiased strategy in which precursor ions are fragmented irrespective of intensity or other characteristics, producing a complete analysis of

precursor ions. Two major strategies have been described: SWATH-MS, and MS^E developed by Waters. SWATH-MS has been used to produce quantitative profiles of complex samples such as human colorectal cancer tumours, however this approach is limited by the requirement for *a priori* information about peptide fragment ion patterns and retention time which may not be available for the particular sample of interest [234, 235]. MS^E approaches do not require such a library. MS scans are performed alternating between high and low collision energy for ion generation, and fragment ions are measured in the former while intact peptides are measured in the latter. Advanced data analysis software matches precursor peptides and fragment ions. Overall this new approach results in high sequence coverage relative to DDA approaches [236-238].

Over the past decade proteomic analysis has been applied in the quest for understanding of host-pathogen interactions, as reviewed in [239]. The majority of studies to date focussed on the pathogen proteome, for example numerous studies investigated the impact of growth conditions on the proteome of *S. Typhimurium*. In one such study *Salmonella* were sorted from tissue homogenates to characterise the proteome during infection of a mammalian host [240, 241]. The proteome of the host is comparatively much larger and contains a greater dynamic range in protein abundance, presenting a bigger challenge both in detection and quantitation of proteins, and in the interpretation of the resulting data. A small number of studies have investigated changes in the host proteome upon infection using cultured cell lines. In particular, in proteomic analysis of a macrophage cell line during infection with *S. Typhimurium* 1,006 macrophage proteins were detected, of which 24% were changed significantly during infection [242]. A similar study of an intestinal epithelial cell line during infection with EPEC detected over 2,000 host proteins of which 13% were differentially expressed upon infection [243]. Whilst macrophage proteins found to be altered in *Salmonella* infection were involved in diverse functions, epithelial cell proteins whose levels were affected by EPEC were mostly involved in actin dynamics, cell adhesion, G-protein signalling and ion transport. These studies are limited to investigation of early events in infection and fail to capture secreted proteins, an important functional category and one which experiences dramatic changes in infection [244]. To our knowledge no studies to date have sought to describe the effects of bacterial infection on the global host proteome at the level of whole tissue in an *in vivo* infection.

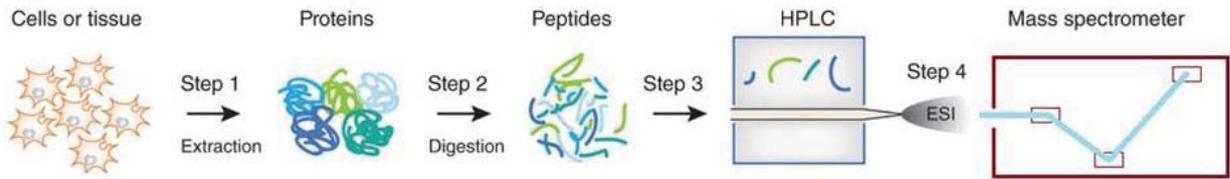
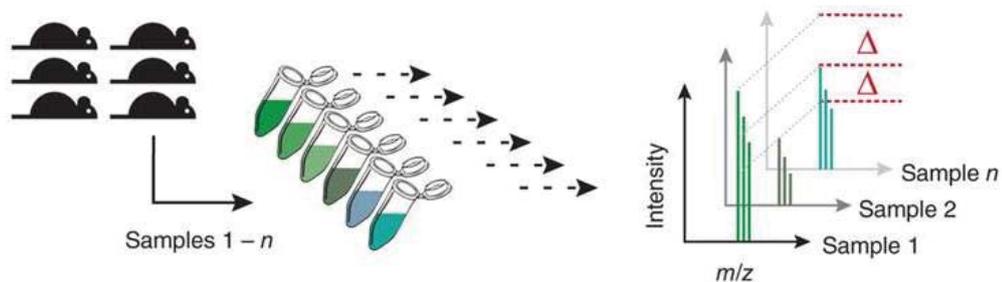
A**B**

Figure 4.1. Quantitative shotgun proteomics. (A) Outline of the process used for the MS analysis of mouse caecal tissue samples described in this chapter. A protocol for extraction of proteins from caecum was developed combining detergent and heat for protein solubilisation and denaturation. Purified extracted proteins were proteolytically fragmented using trypsin. HPLC was used to separate the complex mixture of peptides for MS analysis. Peptides were ionised on exit from the HPLC column, moving directly into the mass spectrometer for time of flight (TOF)-based detection of mass. (B) Diagram to illustrate the label-free intensity-based relative quantification method used in MS analysis described in this chapter. Individual biological samples were prepared separately and analysed sequentially by MS. Quantitation was based on the differential intensities of peptides of identical amino acid sequence and charge between separate MS runs. Δ indicates quantitative peptide differences. (A) and (B) are adapted from [245].

4.1.2 Post-transcriptional control of gene activity

Transcription is thought to be the foremost point of control in the conversion of genetic information into biologically active proteins. Regulation at the level of transcription gives rise to more efficient usage of nucleic acids, and can be achieved relatively quickly through the action of transcription factors. However the need for tight control of gene activity

demands regulatory mechanisms acting at multiple points after transcription, and these post-transcriptional mechanisms play substantial roles.

microRNAs (miRNA) are RNA molecules ~ 21 nucleotides in length which bind mRNA of complementary sequence along with miRNA-associated proteins. miRNA inhibit protein production by two mechanisms; preventing the formation of actively translating polysomes, and triggering degradation of mRNA. The relative importance of these processes varies for different miRNA-mRNA pairs, the reason for which is unknown. At least 1,000 miRNAs operate in humans and bioinformatic predictions suggest miRNAs might regulate 30% of genes in mammals [246].

Cellular localisation of mRNA is important in controlling the rate at which translation is initiated. Following translocation from the nucleus to the cytoplasm some mRNA are sequestered in large ribonucleoprotein (RNP) granules, preventing their translation. It is thought mRNA granules are a mechanism developed to store mRNA for translation under specific environmental conditions [247]. Stress granules are a specific type of RNP granule formed in response to triggers such as heat, oxidative conditions and hypoxia, and have been shown to contain mRNA encoding housekeeping functions. Formation of stress granules diverts translational machinery away from the production of proteins for general cellular upkeep to those important for protection and repair [248].

Ranging from global control mechanisms to targeted regulation of individual genes there are multiple translational regulatory mechanisms. Control of translation can occur at initiation and elongation through availability of specific protein factors involved in these processes, though unlike miRNA and RNP granules this type of control typically affects all transcripts relatively equally and is used as a general control of cell activity. Conversely eukaryotic mRNAs have *cis*-acting elements such as the 5' and 3' untranslated regions (UTR) with which sequence-specific *trans*-acting RNA-binding proteins associate for translational control at the individual gene level. Metabolism in particular has been identified as a group of pathways subject to a high level of translational control on account of the need for rapid changes in response to metabolites, nutrients and endocrine signals [249].

Following translation, protein activities may be controlled by post-translational modifications such as phosphorylation, acetylation and glycosylation. Similar to the control of translation by targeting of mRNAs to specific locations, protein activity is also influenced by

cellular location, and movement of proteins to specific locations is directed by cellular transport machinery. As well as modifications to amino acid side chains, the peptide backbone can itself be altered by proteolytic cleavage, for example many digestive enzymes, clotting factors and proteins involved in apoptosis are activated in this way. Finally, cellular levels of proteins are determined by the rates of both translation and protein degradation. Post-translational ubiquitination leads to targeting of proteins to proteosomal degradation pathways and is an important point of control.

4.1.3 Concordance between RNA transcript and protein abundances

The relative ease of transcriptional profiling compared with quantitative proteomics has resulted in the extensive usage of transcript levels as a proxy for gene activity. However as a result of the post-transcriptional mechanisms of control described, protein and RNA levels are often poorly correlated; a large number of studies spanning organisms from archaea to mammals report mRNA levels are not to be relied upon for the prediction of protein abundance. Examples of such studies include [250-253]. Indeed correlation in abundances of protein and mRNA has been reported to range from $r = 0.6$ in a study on yeast to $r = -0.025$ in a lung adenocarcinoma study [254-256].

The observed differences between protein and transcript abundance are thought to be the result of a complex combination of technical limitations and the biological effects described above. Some past studies have not examined RNA and protein from identical samples, an obvious flaw. The proteins examined are biased towards those which are most abundant due to the threshold of detection by MS, and the regulation of high abundance proteins may not be typical of the entire proteome. In addition there is a need for improved bioinformatic tools to facilitate such comparisons. Protein and RNA abundances are non-normally distributed and some previous attempts to assess correlation have not performed appropriate transformations for the assessment of correlation. Similarly measurement of protein or transcript abundance is influenced by protein or transcript length and these effects are not always accounted for [257].

Differences in protein-RNA correlation between studies may in part be due to biological variation between the samples involved. The correlation between RNA and protein abundance is not thought to be linear at the whole genome scale, indeed past studies have found functionally related groups of genes respond differently to treatment, with some groups

displaying a positive correlation in transcript and protein fold changes, while for others these are negatively correlated [258]. Similarly it is likely that the differential activities of post-translational regulatory mechanisms in different organisms and tissues result in true differences in RNA-protein correlation. A recent study of neutrophils stimulated with LPS demonstrated that different functional groups of genes undergo regulation by different mechanisms in response to certain triggers. A reduction in the level of housekeeping proteins was predominantly achieved by increased rates of protein degradation, while increases in proteins involved in the induced immune response were predominantly the result of increased rates of transcription [259].

In order to overcome the poor predictive power of mRNA abundance over protein levels, experimental techniques have been developed for selective sequencing of actively translated mRNA. Early efforts used sucrose density gradients to selectively purify mRNA associated with ribosomes for sequencing [260]. Ribosome profiling is a more recent approach in which short nucleotide fragments enveloped by the ribosome during translation are selected for sequencing through their protection from nuclease activity [261]. Although these approaches are closer indicators of true gene activity than measurement of global transcript levels, techniques which measure active translation of protein fail to account for the vast differences in protein half-lives, and consequently direct quantitation of protein remains a superior indicator of biological activity.

4.2 Aims of the work described in this chapter

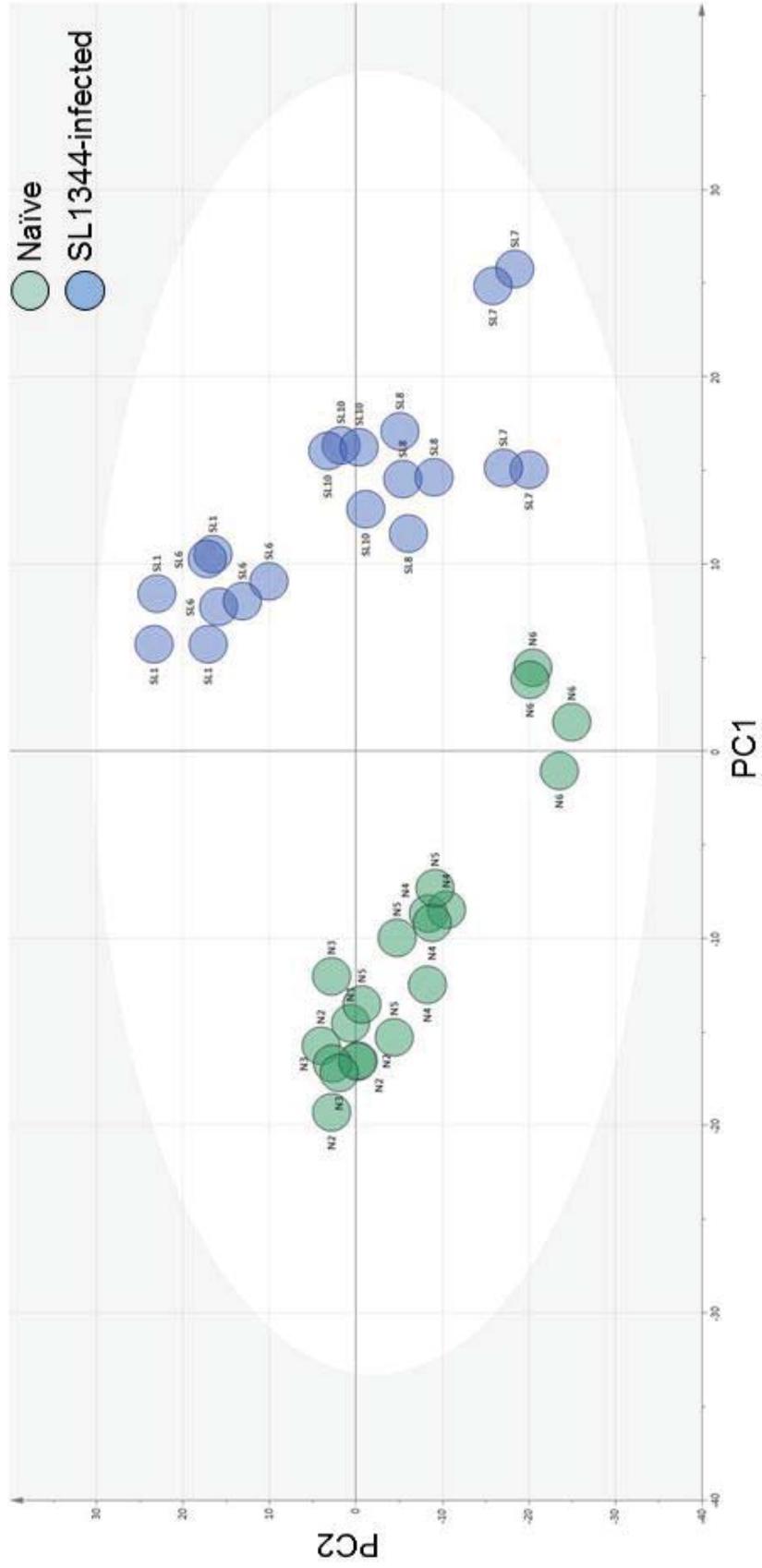
Mass spectrometry was used to quantitatively describe changes in protein abundance in the mouse caecum in response to *S. Typhimurium* infection. The infection-induced changes in the caecal proteome were compared with the transcriptomic dataset introduced in Chapter 3 and the relationship between these complementary datasets described. Genes which underwent coordinated regulation at the levels of both RNA and protein are strongly supported as subjects of regulatory activity during infection and were selected for pathway analysis.

4.3 Results

4.3.1 Proteomic analysis of *S. Typhimurium*-infected caecal tissue

Protein was extracted from caecal tissue pieces from the same *S. Typhimurium* SL1344-infected and naïve control mice as used for RNAseq analysis in section 3.3.6, (n = 5). Each sample was analysed in four mass spectrometry runs using a DIA MS^E approach, and through comparison with the Uniprot sequence database mouse proteins present in each run were quantified and identified. In total 7,499 proteins were identified (multiple peptides were detected, at least one of which was unique to the protein), and 3,590 proteins were quantified in multiple MS runs. In Figure 4.2 a PCA plot and dendrogram show clustering of individual MS runs based on their proteome profiles. Runs are seen to cluster according to biological group with the exception of a single naïve sample which clusters with SL1344-infected sample runs (N6 in Figure 4.2). After consideration it was decided this sample should be included in the naïve sample group in subsequent analysis despite the difference in clustering; the different grouping potentially reflecting true biological differences, though the alternative of unintentional sample cross-contamination cannot be excluded. MS runs of protein extracted from individual samples show some degree of clustering although runs of samples from individual mice are not clearly separated. In principal component analysis of the MS data naïve and infected samples are separated by the first principal component.

A



B

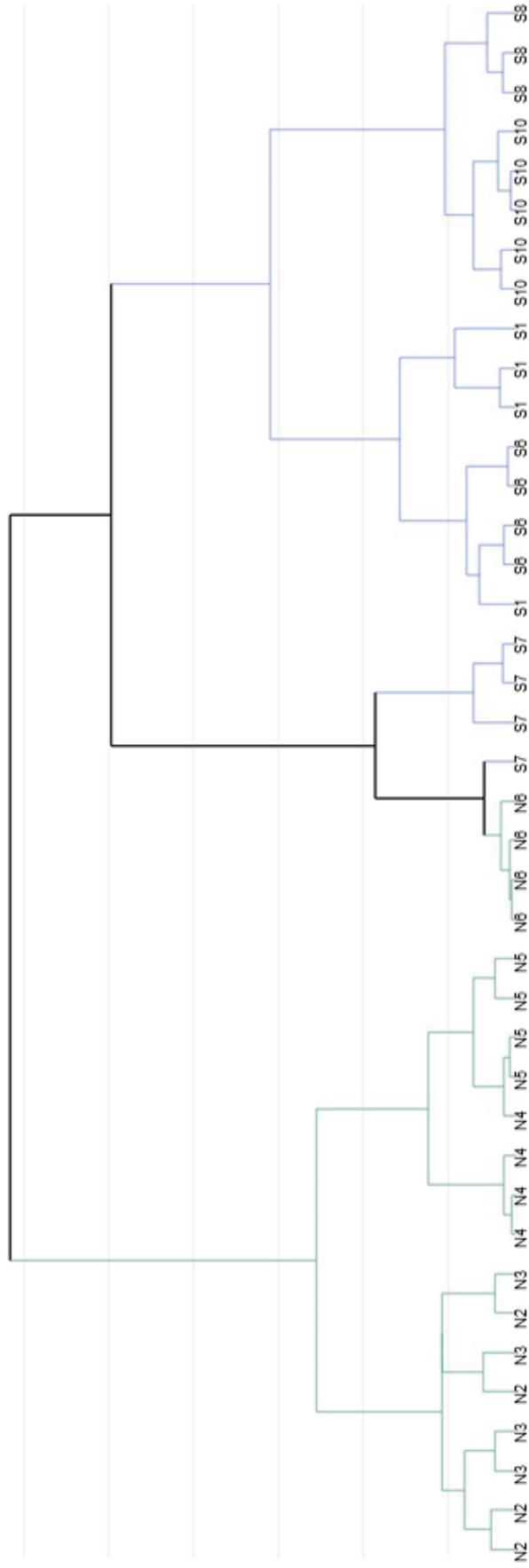


Figure 4.2. MS analysis of the murine caecal proteome at day 4 PI with *S. Typhimurium* SL1344. (A) Principal component analysis of caecal tissue proteome profiles. Proteins extracted from *S. Typhimurium* SL1344-infected and naïve control mice ($n = 5$) were analysed by MS with four runs per sample. Each circle represents one MS run; labels denote the biological samples from which profiles were produced. (B) Cluster dendrogram of MS profiles as in (A).

Proteins quantified in both naïve and *S. Typhimurium* SL1344-infected samples were tested for changed abundance upon infection. 59 proteins were found to be significantly increased in infection and 202 proteins significantly decreased (\log_2 fold change > 1 and < -1 respectively, $p < 0.05$). Due to the high detection threshold of MS many proteins were detected exclusively in samples from either the naïve control group or *S. Typhimurium* SL1344-infected group. These ‘single-condition proteins’ represent the amalgamation of proteins regulated in infection such that their abundance exceeds the MS detection threshold in one condition only, and proteins with a stable abundance in the region of the MS detection threshold (therefore detected exclusively in one group by chance). In order to exclude proteins of this latter type from the category of ‘regulated proteins’ in subsequent analyses a threshold was set; ‘single-condition proteins’ were required to be detected in a minimum of five MS runs (therefore detected in ≥ 2 biological samples) to be considered regulated in infection. 123 proteins detected exclusively in infected samples and 87 proteins detected exclusively in naïve samples satisfied the condition of detection in ≥ 5 runs. Throughout this chapter ‘proteins significantly upregulated in infection’ refers to the 59 significantly upregulated proteins detected in both naïve and infected sample groups and 123 proteins detected in infected samples only (≥ 5 MS runs) combined. Similarly ‘proteins significantly downregulated in infection’ refers to the 202 significantly downregulated proteins detected in both sample groups and 87 proteins detected in naïve samples only (≥ 5 MS runs) combined.

Peptides detected by MS were also compared with a *S. Typhimurium* protein reference dataset. In total 299 *S. Typhimurium* proteins were quantified. Of these 177 proteins were detected in extracts from *S. Typhimurium* SL1344-infected samples, and 153 were detected in protein extracts from naïve control samples. The average number of *S. Typhimurium* proteins quantified in naïve and infected samples were remarkably similar at 9.7 and 11.1 proteins respectively. As many of the detected *S. Typhimurium* proteins are components of highly conserved bacterial processes and widely present throughout the bacterial kingdom; for example DnaK in DNA replication and TalA in the pentose phosphate pathway; it is likely that many of these proteins were in actual fact *S. Typhimurium* protein homologues derived from species of the microbiota. Whilst 147 *S. Typhimurium* proteins were quantified exclusively in infected samples, of these only 11 were quantified in multiple runs (9 proteins were detected in two runs and 2 were detected in three MS runs), indicating bacterial proteins were relatively rare amongst the host proteins in tissue extracts.

4.3.2 Comparative analysis of transcriptomic and proteomic datasets

The mouse caecal protein profiles described in section 4.3.1 and corresponding mRNA profiles described in section 3.3.6. were compared. Fold changes in protein and gene abundance observed at day 4 PI with *S. Typhimurium* were tested for correlation, and genes which were observed to undergo similar regulation at both the protein and RNA level were identified.

4.3.2.1 The caecal transcriptome is more dramatically changed in *S. Typhimurium* infection than the proteome

Figure 4.3 displays numbers of proteins for which changes in abundance were detected in infection with the total number of proteins identified by MS, similarly numbers of regulated mRNA transcripts are shown with the total number of transcripts identified. Over three times more transcripts were identified compared with proteins. Compared with proteins 19.8 times more transcripts were significantly upregulated and 9.6 times more transcripts significantly downregulated in infection, and therefore the proportion of detected transcripts which were significantly regulated was higher than the equivalent proportion of proteins. Figure 4.4 displays the distribution of fold changes in protein abundance upon infection. Relative to fold changes in transcript abundance changes at the protein level are smaller overall. However a fold change cannot be calculated for proteins detected exclusively in one condition and these likely represent some of the most highly regulated proteins.

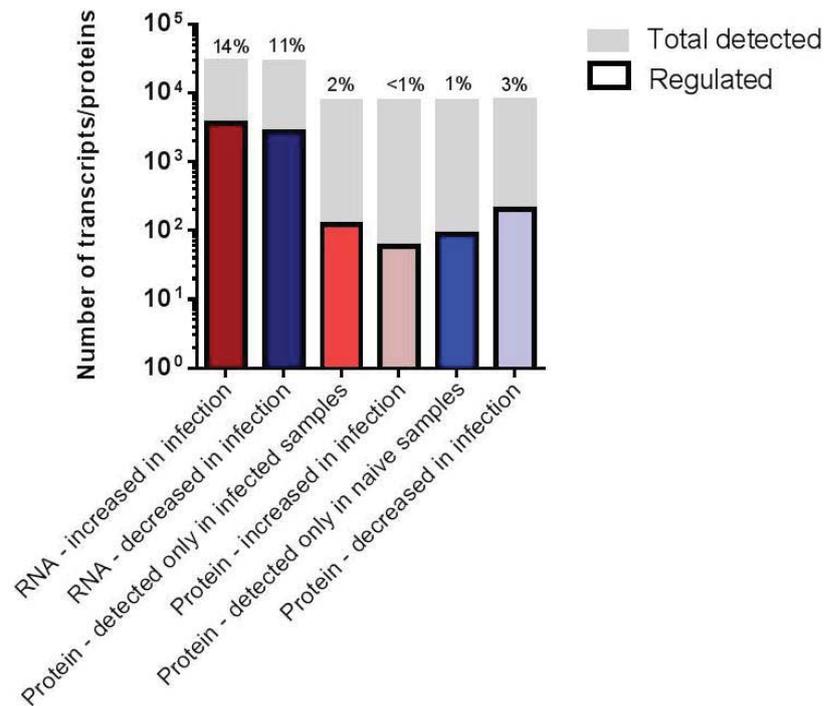


Figure 4.3. Comparison of transcripts and proteins differentially regulated during infection with *S. Typhimurium* in murine caecum. Bar graph of numbers of differentially regulated transcripts and proteins (coloured bars) as a proportion of total entities detected (pale grey bars). Percentages above bars indicate the proportion of the total transcripts/proteins with changed abundance in infection according to the specified condition. A total of 25,342 transcripts were identified by RNAseq and 7,499 proteins identified by MS (multiple peptides detected, at least one unique to the protein). Both transcripts and proteins changed in infection are defined as those with log₂ fold change < -1 or > 1, p-value < 0.05 after correction for multiple testing, with the exception of proteins detected exclusively in samples of a single condition for which these values cannot be calculated. For proteins detected exclusively in one condition only those detected in ≥ 5 machine runs are included.

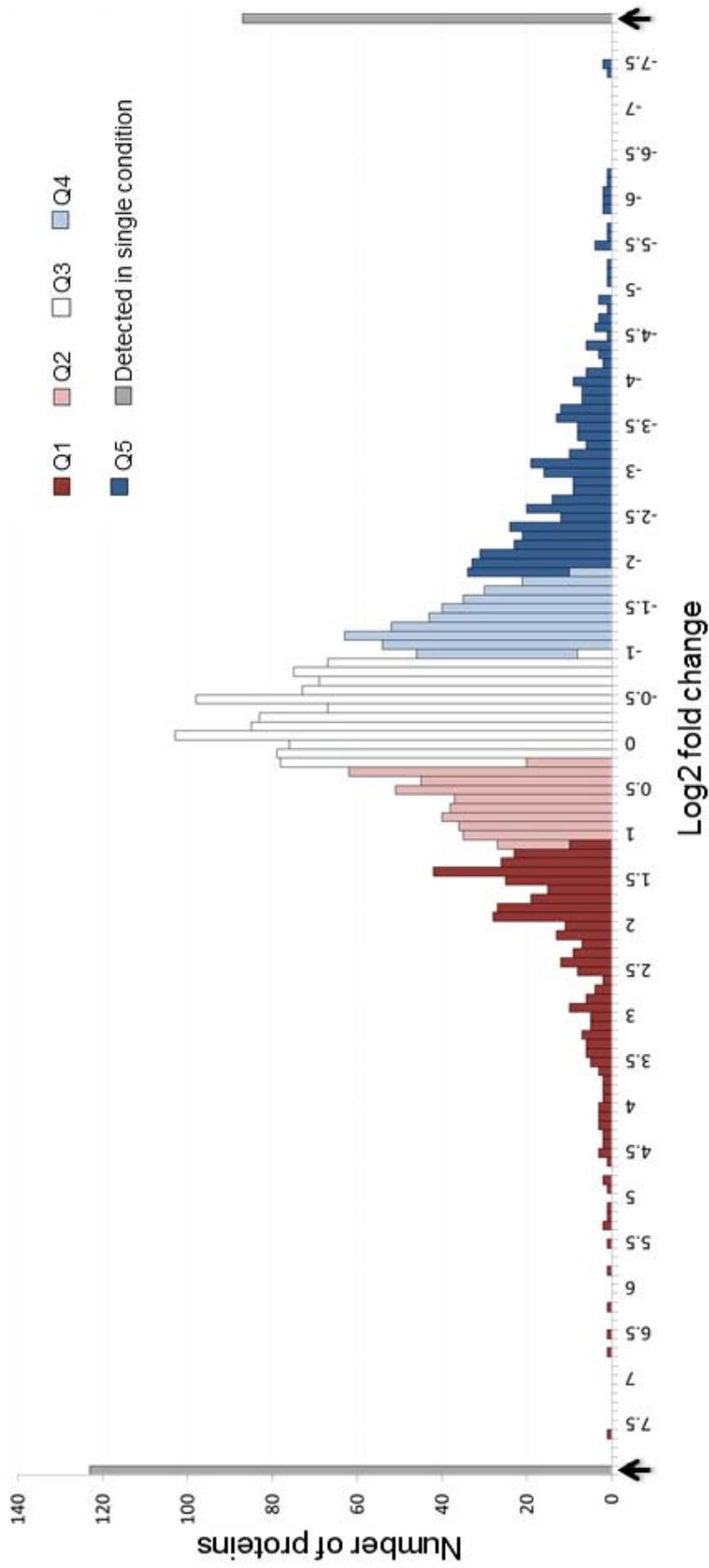


Figure 4.4. Quantile plot of log₂ fold changes in protein abundance during *S. Typhimurium* infection. Distribution of log₂ fold changes in protein abundance in caecal tissue at day 4 PI with *S. Typhimurium* SL1344. Quantiles are as follows: Q1: 0% - 15%, Q2: 15% - 30%, Q3: 30% - 70%, Q4: 70% - 85%, Q5: 85% - 100%. Plotted at the extremes of the x axis indicated by arrows are significant sample specific proteins (detected in ≥ 5 runs).

4.3.2.2 Changes in abundance of proteins encoded by transcripts most highly regulated in *S. Typhimurium* infection

Proteins encoded by the transcripts most highly regulated in infection were examined. Of the 50 genes most highly upregulated in infection at the level of RNA only five of the encoded proteins were detected to be significantly upregulated, all of which were detected exclusively in samples from *Salmonella*-infected mice. Of the 50 genes most downregulated at the level of RNA seven of the corresponding proteins were detected to be significantly downregulated by MS. Table 4.1 contains normalised abundance values, log₂ fold changes and adjusted p-values from both RNAseq and MS for these genes. The observed poor correspondence of RNA and protein regulation for this selection of 100 genes highly regulated at the RNA level is to a large extent the result of quantification of fewer proteins; in total just 29 of the proteins encoded by the 100 genes were quantified. Of the quantified proteins which were non-significantly regulated > 75% displayed an abundance change in infection in the direction observed for the corresponding transcript.

Gene	RNAseq			MS		
	Transcript abundance	Log2 fold change	Adjusted p-value	Protein abundance	Log2 fold change	Moderated T-test p-value
S100A9	7254	7.54	8.09E-54	0.22	Infected only (12)	-
S100A8	3722	7.28	6.89E-40	0.24	Infected only (5)	-
LCN2	12954	7.23	2.74E-84	0.06	Infected only (9)	-
NGP	508	7.09	4.49E-28	0.18	Infected only (19)	-
HP	4552	6.72	3.81E-47	1.6	Infected only (20)	-
CYP2C55	17747	-9.99	1.09E-215	0.34	Naive only (20)	-
HAO2	10026	-9.68	3.52E-105	0.22	Naive only (16)	-
GSDMC3	10349	-7.61	1.11E-62	0.01	Naive only (10)	-
GSDMC2	35129	-7.43	1.51E-63	0.13	Naive only (14)	-
UGT2B36	79	-6.42	4.40E-46	0.01	Naive only (6)	-
HSD3B3	4574	-6.11	5.23E-74	0.06	Naive only (5)	-
GSTM3	666	-6.06	2.78E-78	0.37	-3.14	1.51E-05

Table 4.1. Genes with the largest changes in transcript abundance upon *S. Typhimurium* infection for which significant regulation of the encoded protein was also detected. Genes included in the table are those in the 50 most highly upregulated and 50 most highly down-regulated genes at the transcriptional level, with significant regulation in the proteome also. Transcript abundance is the baseMean output from DESeq2 (the average of the normalised count values over all samples). Protein abundance is the normalised summed top three peptide intensities for each protein averaged for all samples. For proteins detected exclusively in samples in either the *S. Typhimurium* SL1344-infected or naïve control groups the number in brackets in the column ‘log2 fold change’ is the number of MS runs in which the protein was detected.

4.3.2.3 Correlation between changes in transcript and protein abundance during *S. Typhimurium* infection

Fold changes in protein and RNA abundance for all genes quantified in both RNAseq and MS were plotted (Figure 4.5) and the correlation between these two variables assessed. Proteins detected exclusively in infected samples were assigned an arbitrary log2 fold change of 10 and proteins detected exclusively in naïve samples an arbitrary log2 fold change of -10 for the purpose of this analysis. Pearson’s r for the correlation of all genes was equal to 0.46, indicating a fairly small positive correlation. The correlation for genes with protein quantified in both naïve and *S. Typhimurium*-infected sample groups was extremely modest; $r = 0.16$, while for ‘single condition’ genes correlation between RNA and protein fold changes was good; $r = 0.73$.

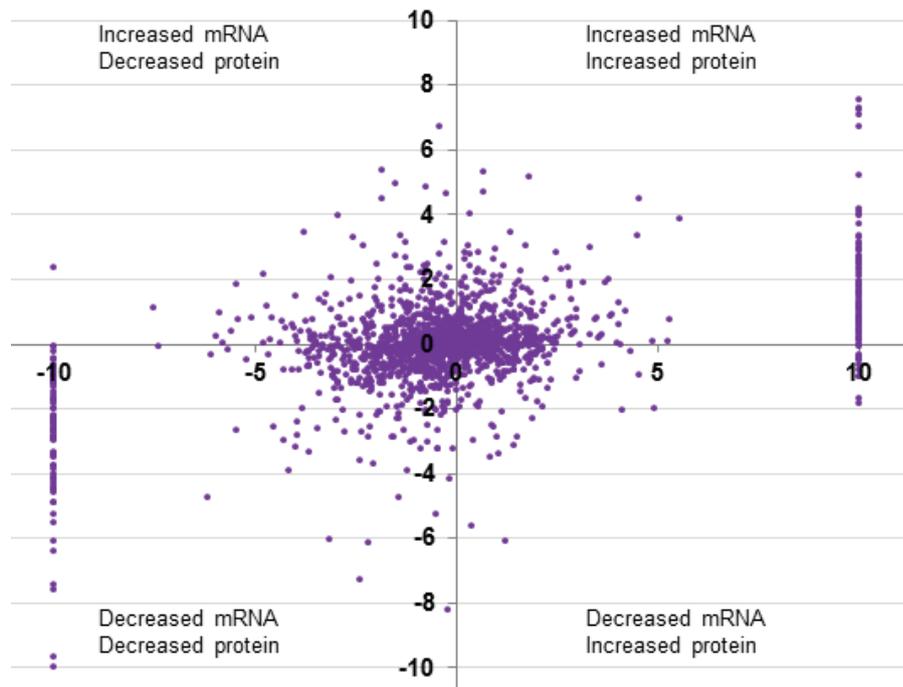


Figure 4.5. Correlation between fold changes in transcript and protein abundance at day 4 PI with *S. Typhimurium* in mouse caecum. x-axis displays log₂ fold change in protein abundance and y-axis the log₂ fold change in transcript abundance during *S. Typhimurium* infection, as determined by MS and RNAseq respectively. Proteins detected exclusively in a single condition in ≥ 5 runs were assigned an arbitrary log₂ fold change: 10 for proteins detected exclusively in infected samples and -10 for proteins detected exclusively in naïve control samples. Proteins detected in < 5 runs, and genes for which multiple RNA or protein isoforms were detected, were excluded.

Transcripts and proteins significantly differentially regulated in *Salmonella* infection in the caecum were compared in order to identify genes with evidence of regulation at both the RNA and protein level. The Venn diagram in Figure 4.6 displays numbers of significantly regulated transcripts and proteins and the extent to which these overlap. Where multiple isoforms of a transcript or protein were similarly regulated these were counted as a single entity. Isoforms significantly regulated in opposing directions upon infection were detected for a single gene only: protein isoforms of the actin-binding protein Filamin A were detected in each of the up- and down-regulated protein groups. 74 genes were upregulated at day 4 PI at the level of RNA and protein, while 115 genes were downregulated in both datasets. In agreement with the modest correlation between changes in RNA and protein described above, less than half of up- or down-regulated proteins were encoded by genes also significantly regulated at the level of RNA. Small numbers of genes were regulated in opposing directions

at the transcript and protein level upon infection; six genes displayed increased protein abundance upon infection while RNA was simultaneously decreased, and seven genes were decreased at the protein level with increased RNA.

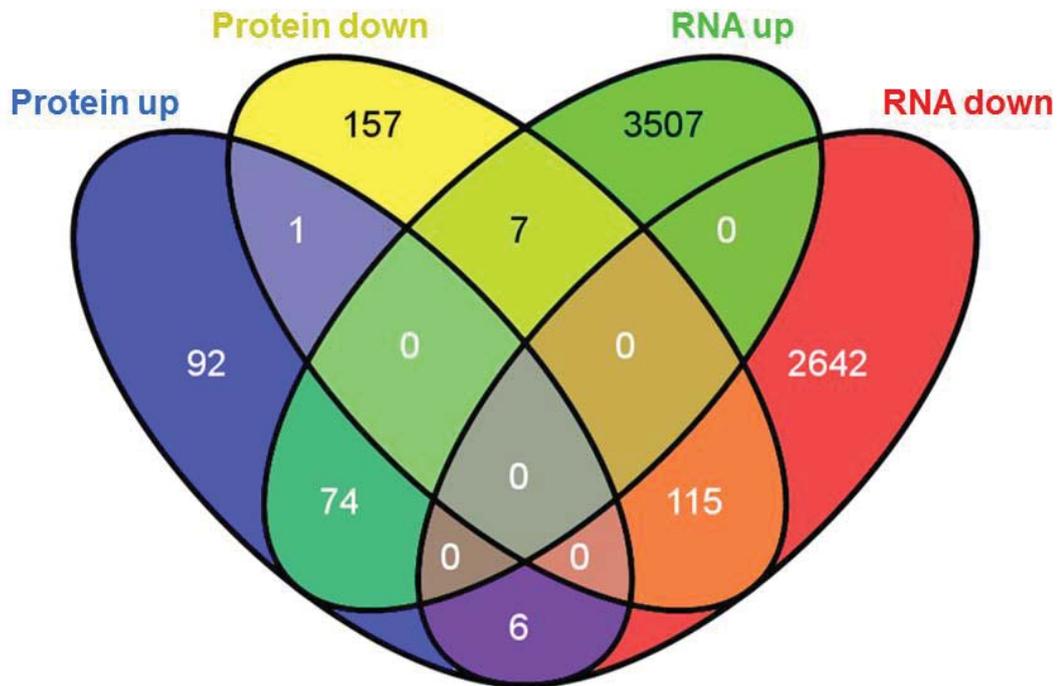


Figure 4.6. Venn diagram to show overlap between significantly regulated transcripts and proteins. Numbers of transcripts and proteins regulated in caecal tissue at day 4 PI with *S. Typhimurium* (transcripts: log₂ fold change < -1 or > 1, adjusted p-value < 0.05, proteins: log₂ fold change < -1 or > 1, adjusted p-value < 0.05 and proteins detected exclusively in *S. Typhimurium* SL1344-infected or naïve samples in ≥ 5 runs). Protein or RNA isoforms of the same gene were considered as one where the direction of regulation upon infection was the same.

4.3.3 Pathway analysis of consensus regulated genes in transcriptome and proteome datasets

The 74 genes upregulated in both the transcriptome and proteome were analysed using the InnateDB pathway analysis tool to identify pathways in which these genes are significantly overrepresented. The ten pathways most significantly associated with these ‘consensus’ upregulated genes are listed in Table 4.2 and the full list of significantly associated pathways detailed in Appendix 4. In total 34 pathways were significantly

associated with the consensus upregulated genes (corrected p-value < 0.05). Pathway analysis was similarly performed for consensus downregulated genes and 40 significantly associated pathways were identified, listed in Appendix 5.

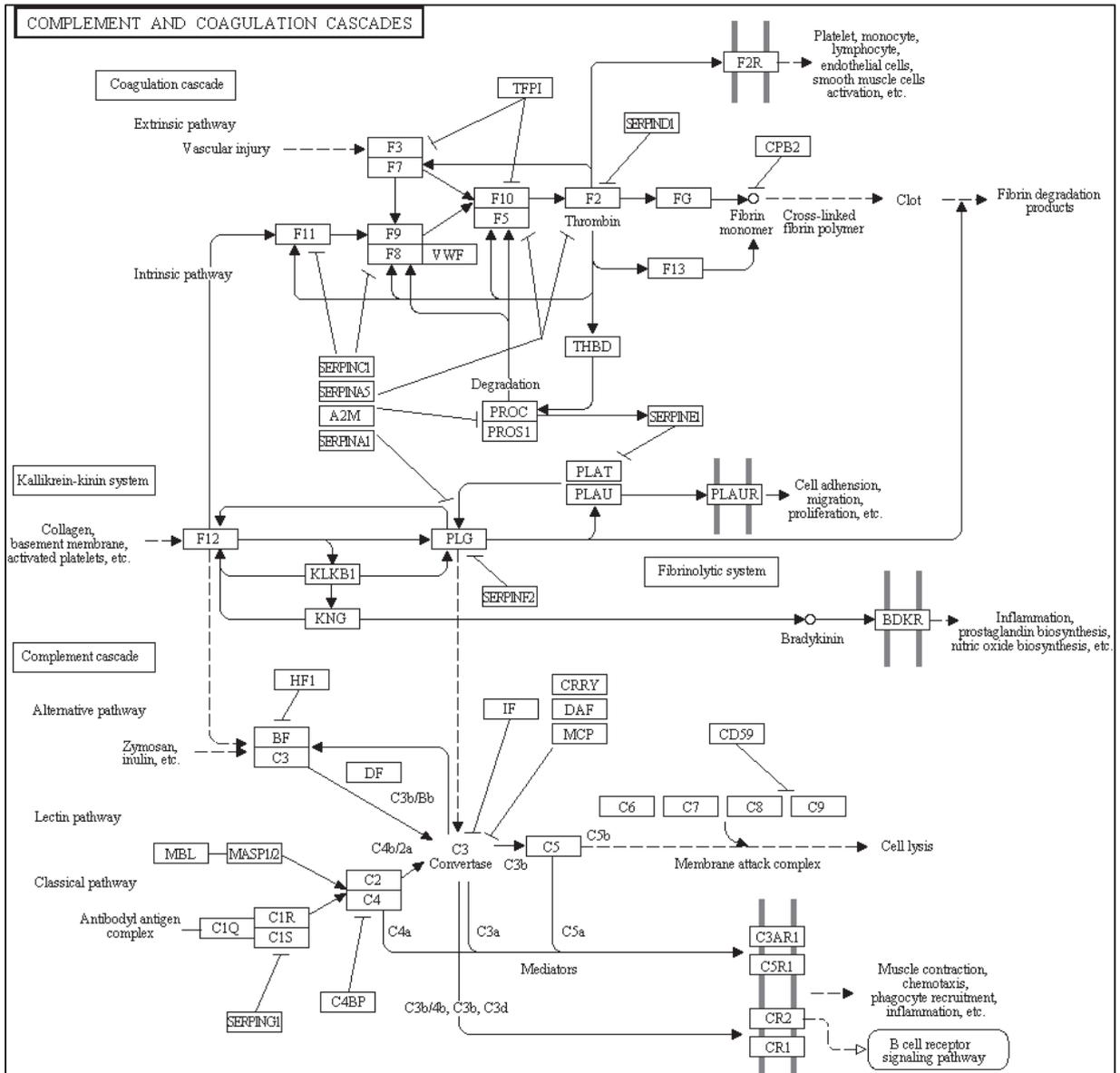
Pathway name	Up-regulated gene count	% of pathway genes up-regulated	Pathway p-value (corrected)	Gene symbols
Regulation of complement cascade	4	22	3.77E-05	C3, C4b, Cfb, Cfh
Activation of C3 and C5	3	50	4.12E-05	C3, C4b, Cfb
Endosomal/vacuolar pathway	3	38	9.17E-05	B2m, Ctss, H2-K1
Complement cascade	4	13	1.62E-04	C3, C4b, Cfb, Cfh
Interferon signalling	5	7	2.42E-04	B2m, Gbp2, H2-K1, Isg15, Ptpn6
Antigen processing-cross presentation	5	7	2.52E-04	B2m, Ctss, H2-K1, Psmb8, Tapbp
Interferon gamma signalling	4	9	3.79E-04	B2m, Gbp2, H2-K1, Ptpn6
Initial triggering of complement	3	18	4.49E-04	C3, C4b, Cfb
Antigen presentation: folding, assembly and peptide loading of class I MHC	3	16	5.87E-04	B2m, H2-K1, Tapbp
Alternative complement activation	2	50	7.42E-04	C3, Cfb

Table 4.2. The 10 Reactome pathways most significantly associated with genes upregulated in both the transcriptome and proteome during *S. Typhimurium* infection. ‘Upregulated gene count’ indicates the number of genes upregulated at the protein and RNA level annotated to each pathway; the names of which are listed in the column ‘Gene symbols’.

More than half of pathways associated with consensus upregulated genes were identical to pathways associated with transcripts upregulated during infection, and of the remainder pathway terms appeared closely related to those associated with upregulated transcripts. For consensus downregulated genes ‘metabolism’, ‘biological oxidations’ and ‘phase II conjugation’ remained the three most highly associated pathways, and pathways relating to the metabolism of amino acids, sugars and lipids were again highly associated. A striking difference in the results of pathway analysis of consensus upregulated transcripts and proteins was the rise in the relative significance of pathways relating to the complement cascade. The complement pathway term most significantly associated with upregulated transcripts; ‘initial triggering of complement’; was placed 37th in the list of pathways with a corrected p-value of 3.6×10^{-4} . In the pathways associated with consensus upregulated genes complement-related pathways feature more highly; the most significantly associated pathway is ‘regulation of complement cascade’, and complement-related pathways comprise 5 of the 10 most highly associated pathways.

Figure 4.7 shows the KEGG pathway ‘complement and coagulation cascades’ with highlighting to denote proteins and transcripts significantly upregulated in the caecum at day 4 PI with *S. Typhimurium*. The complement pathway is clearly subject to extensive upregulation during infection, with many components regulated both at the level of the transcriptome and proteome.

A



B

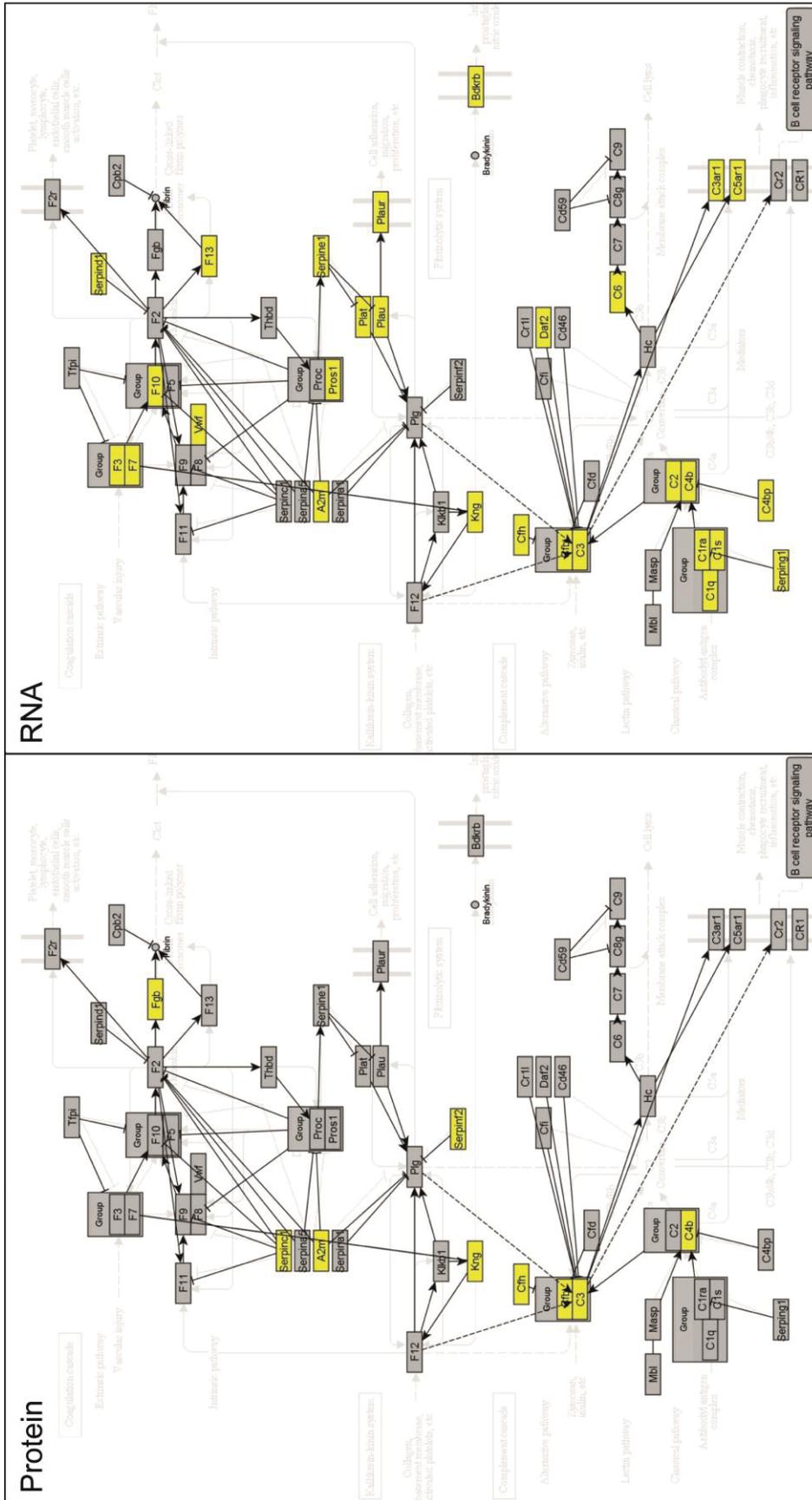


Figure 4.7. Regulation of the KEGG pathway ‘Complement and coagulation cascades’ in *S. Typhimurium* infection. (A) Pathway diagram for the KEGG pathway ‘Complement and coagulation cascades’ in mouse. (B) ‘Complement and coagulation cascades’ pathway with components significantly upregulated in the caecal proteome (left) and transcriptome (right) at day 4 PI with *S. Typhimurium* coloured in yellow. The analysis tool InCroMAP was used to generate the pathway diagrams in (B) [262].

4.4 Discussion

Integrative analysis of ‘omics’ data is an emerging area of interest, and predicted to become a major strategy for gaining insight into interactions taking place in complex biological systems [263-265]. In the work described in this chapter transcriptomic and proteomic data were integrated to identify processes subject to regulation at multiple levels during *S. Typhimurium*-induced inflammation in the caecum.

Both transcriptomic and proteomic data can describe changes in the abundance of biological molecules; however differences between them introduce challenges in their integration. The number of entities quantified by RNAseq and MS with current technology is a major difference; in this work the number of proteins quantified in mouse caecal tissue was less than a third of the number of transcripts, despite the greater mass of caecal material used for protein extraction. As described in [257] this is a significant challenge in the integrative analysis of proteomic and transcriptomic data, and improved analysis methods are required to appropriately account for the absent proteins. For some genes under certain conditions regulatory mechanisms may control RNA and protein levels such that transcripts are present in the absence of encoded protein. Indeed it has been reported that for many genes with a relatively low level of transcription no protein product is translated [266]. However, true biological differences make a relatively minor contribution to the existence of ‘missing proteins’.

The difference in the sensitivities of RNAseq and MS compared with the dynamic ranges in transcript and protein abundance is the foremost reason for the failure to detect some proteins encoded by genes shown to be actively transcribed. The massive dynamic range in the proteome; approaching seven orders of magnitude; creates a major hurdle in detection of proteins of lower abundance [267]. In contrast the dynamic range in the transcriptome is much smaller, between three and four orders of magnitude [268]. The number of proteins quantified in this work is comparable to (and in the majority of cases exceeds) numbers reported in

proteomic studies published in only the past one or two years [235, 269]. The major improvements in MS technology making whole tissue profiling possible occurred only recently. As protein abundance more accurately reflects gene activity compared to transcript abundance improvements in MS sensitivity are welcome. However profiling of the proteome with coverage comparable with RNAseq profiling of the transcriptome remains a possibility of the distant future.

In addition to proteins which are entirely absent from the proteomic dataset the lower sensitivity of MS compared with RNAseq and the larger dynamic range in protein abundance results in the detection of proteins exclusively in samples from a single treatment group. As fold changes cannot be calculated for these proteins, the manner in which to appropriately include these in analysis must be decided upon - there is no standard method to deal with these cases. Here detection in five MS runs was chosen arbitrarily as the minimum requirement for calling proteins detected exclusively in one condition 'regulated'. Unfortunately there is no simple way to overcome the fact that failure to detect a protein in both conditions means information describing the extent of regulation is missing.

Detection of bacterial proteins within infected tissue presents the possibility of identifying *S. Typhimurium* virulence factors, in particular effector proteins injected into host cells through *Salmonella* T3SS. For this reason *S. Typhimurium* proteins in caecal tissue samples were of interest. Our findings demonstrate detection of possible *S. Typhimurium* proteins in infected caecal tissue; however there exists difficulties in distinguishing *Salmonella* proteins from widely conserved bacterial proteins. Peptides which aligned to *Salmonella* proteins were detected in both naïve control and *S. Typhimurium*-infected samples. As peptides detected exclusively in infected samples were of low abundance their apparent condition-specific distribution may have occurred simply by chance. Basic local alignment search tool (BLAST) searches could be performed to distinguish those proteins which are unique to *Salmonella* and those which are widely conserved across bacterial species. Greater sensitivity in protein detection is required to detect virulence factors. This could be achieved by proteome fractionation approaches or faster sequencing [270]. Alternatively dissociation of the caecal mucosa from the remainder of the organ and protein extraction from this region might increase the proportion of infected cells and *S. Typhimurium* proteins. In addition introduction of washing steps to clean faecal material

from the caecum and remove associated microbiota might reduce proteins from contaminating bacteria.

When equivalent fold change and significance thresholds were applied a smaller proportion of identified proteins were found to be regulated during *Salmonella* infection compared with transcripts. In Figure 4.3 proportions of proteins identified by MS which were regulated are shown with equivalent proportions for transcripts. Whilst considering the regulated fraction of all identified entities is logical in the case of transcripts detected by RNAseq, it is arguably more appropriate to consider regulated proteins as the proportion of proteins quantified in multiple MS runs, since proteins must be detected in multiple runs for differences in abundance between groups to be detected. Even so, proteins up- and down-regulated in infected samples form 5% and 8% of proteins quantified in multiple runs respectively, compared with 14% and 11% for transcripts. Several factors may give rise to the seemingly wider regulation of the transcriptome. MS detects only the most highly abundant proteins, the regulation of which may not be typical of proteins generally. A complex relationship between RNA and protein abundance might exist such that transcripts are truly more dramatically changed upon infection than proteins. The abundance of stable proteins may take longer to respond to changes at the RNA level than the time elapsed between the occurrence of these changes and sample collection. Further differences between the proteome and transcriptome may arise from imported and secreted proteins. Whilst changes in transcription of genes encoding secreted proteins should be detected as for any other gene, secreted proteins may be under-represented in the proteome. The reverse is true of potential imported proteins transcribed and translated in other tissues and travelling to the caecum during infection.

The difference in sensitivity between MS and RNAseq is once again likely to be a major factor. Many proteins which are truly regulated are not consistently quantified in individual MS runs and therefore fall short of the significance threshold for regulation. Proteins detected exclusively in samples from one condition are a related problem; this group is likely to be 'hiding' some of the largest fold changes in protein. In section 4.3.2.2 greater than 75% of non-significantly regulated proteins encoded by the 50 most highly up- and down-regulated transcripts showed evidence of regulation in the same direction as the transcript. With deeper proteomics data it is likely a large proportion of these would be found to be significantly regulated.

A previous study which investigated changes in the proteome of RAW 264.7 macrophages in *S. Typhimurium* infection found 24% of identified macrophage proteins were significantly changed in abundance in infected samples at one or all of the 2, 4, and 24 h time points sampled [242]. In contrast to our study where a log₂ fold change of > 1 or < -1 (absolute fold change of > 2) was considered to indicate regulation, a higher threshold of a five-fold difference was applied in the macrophage study. Our finding of 6.3% of detected proteins (or 13% of proteins detected in multiple runs) regulated in infection despite the lower threshold used to define regulated proteins is substantially different. However the proportion of cells infected in whole caecal tissue during an *in vivo* infection is small compared with cultured macrophages following 24 h incubation with *S. Typhimurium* at a multiplicity of infection of 100. In addition, macrophages are highly adapted to respond to bacteria. Many of the cells which become infected in caecal tissue are non-immune cells which may not possess such elaborate mechanisms to direct a response. Perhaps surprising though is the minimal overlap between the proteins regulated in the macrophage study and in caecal tissue; 9 of the 244 proteins found to be regulated in RAW264.7 macrophages were also regulated in caecal tissue, and of these the direction of regulation was in agreement for just 6 (Itih2, Cs, Met, Pgm2, Adh5, Idh1). Hadhb, involved in mitochondrial β -oxidation of long chain fatty acids, was upregulated in infected macrophages and downregulated in infected tissue. Psap, a precursor of proteins involved in catabolism of glycosphingolipids, and Pgm2 involved in carbohydrate metabolism, were both downregulated in infected macrophages and quite considerably upregulated in infected tissue (approximately 20-fold and 10-fold respectively).

The correlation between the fold changes in transcripts and proteins in *S. Typhimurium*-infected caecum was found to be positive, although poor. Assignment of an arbitrary fold change to proteins detected exclusively in either naïve or *S. Typhimurium*-infected caecum samples and inclusion of these increased the correlation substantially as considered alone the fold change in protein and RNA for the ‘single-condition proteins’ showed good positive correlation ($r = 0.73$). The correlation observed here is in line with previous studies, though only a few report correlation in fold changes upon changing conditions rather than correlation between RNA and protein abundances under a single condition [271, 272]. This finding of a limited correlation between transcript and protein fold changes upon infection adds further support to the idea that post-translational regulatory mechanisms are extensive. Given the opportunity it would be interesting to compare the correlation of transcript and protein fold changes over several time points to gain insight into

the importance of different regulatory mechanisms as *Salmonella* infection progresses. Post-transcriptional regulatory mechanisms have greater prominence under certain conditions and in response to particular stimuli; it will be interesting to discover how infection fits into this picture.

Several options were available for the integrative analysis of RNA and proteomics data, and further a strong argument can be made for analysing proteomics data in isolation given that protein abundances are more accurate indicators of gene activity than transcript abundances. Analysis tools for the integrative analysis of ‘omics’ data, though in their infancy, have emerged in recent years. For example the web tool IMPalA is designed for integration of metabolomics data and transcriptomics data, and another freely available analysis tool, InCroMAP, for the integration of a multitude of data types including DNA methylation, protein modifications, metabolomics and gene-based abundance data [262, 273]. A third and conceptually simple approach was to select genes regulated in the same manner at the RNA and protein level for pathway analysis. These genes where regulation is independently validated by separate techniques can be considered strongly supported as subjects of regulation. Each of the options described has its merits and its disadvantages and all three were investigated during the course of this work. Unsurprisingly the pathways determined to be significantly associated with the relevant datasets were largely the same, although with different supporting genes and degrees of association in each case. More detailed investigation is required to appreciate the similarities and differences between the three outputs in finer detail and to determine if there exist pathways significantly associated with the data by a single approach. Interestingly the significance of the complement pathway is relatively strong in every case. Using InCroMAP for integrative analysis of all transcripts and proteins, including information on the magnitude of the fold change detected in each case but irrespective of p-value, the KEGG pathway ‘complement and coagulation cascades’ was the third most highly associated pathway with a p-value of 6.6×10^{-9} . Further, pathway analysis of all upregulated proteins in InnateDB identified this pathway to be the most highly associated pathway in the KEGG database.

While 5 of the 10 pathways most highly associated with consensus upregulated genes were related to the complement cascade these pathways are overlapping in their annotated genes and in fact only four genes; the activation pathway components C3, C4b and Cfb, and the additional regulatory factor Cfh, are identified from the consensus gene dataset in all of

these pathways. Therefore further work is needed to identify the importance of different sub-pathways within the broader umbrella of the complement cascade.

The results of additional proteomics analysis not described in this chapter lend further support to a particular involvement of complement in *S. Typhimurium* infection in the caecum. During extraction and purification of proteins from caecal tissue for MS analysis proteins were separated according to molecular weight, with the analysis of proteins in the > 30 kDa molecular weight fraction described throughout this chapter. Fractions containing proteins smaller than 30 kDa were pooled according to the initial sample group (naïve control or *S. Typhimurium*-infected tissue) and prepared and analysed by MS separately. Following filtering of data to exclude weakly supported proteins just 4 of the 83 protein groups identified displayed an increased abundance in infection and the remainder appeared downregulated. As analysis was performed with just two pools of samples it would be inappropriate to surmise that infection is associated with a dramatic downregulation of small proteins and peptides, though these findings warrant further investigation with independent samples. Interestingly however the four proteins increased in the pool of proteins from infected tissue included Myeloperoxidase, a major protein in neutrophil granules and the major complement protein C3.

5 Signatures of the streptomycin mouse model: the complement system in mucosal *S. Typhimurium* infection

5.1 Introduction

5.1.1 The complement system in the immune response to infection

The complement system is a collection of > 30 proteins, including both soluble proteins activated through a cascade of proteolytic activity, and receptors for detection of activated complement protein fragments. Three pathways of activation converge to generate effector functions essential in immunity to many common pathogens. As a powerful defence mechanism tight controls upon the complement system prevent excessive activation and collateral damage to the host. The system is evolutionarily ancient and many functions in innate immunity have been gradually uncovered [274]. Complement has also been shown to participate in the stimulation of adaptive immunity [275]. Complement proteins are present at high concentrations in blood plasma and are produced rapidly in the acute phase response (APR) to inflammation; consequently the liver was originally thought to be the sole producer of biologically significant levels of complement proteins. However recent findings suggest a role for locally produced complement in inflammation and infection.

5.1.1.1 Three complement activation pathways converge upon a C3 convertase

The complement system undergoes activation via three pathways and activation by different proteins facilitates responses to a wide range of pathogen triggers. The first pathway to be described, known as the ‘classical pathway’, is founded around a complex of three complement proteins, C1q, C1r and C1s, that are together called the C1 complex. Upon binding of C1q to the Fc region of antigen-bound ‘complement-fixing’ antibodies (primarily IgG and IgM subclasses), the C1 complex serine proteases C1r and C1s are autocatalytically activated. Protease activation leads to cleavage of complement proteins C2 and C4. The larger fragments of protein resulting from these cleavage reactions associate to form the complex C4bC2a. Described as a C3 convertase, C4bC2a cleaves C3 into 9 kDa and 185 kDa fragments; C3a and C3b respectively.

The lectin pathway is similar to the classical pathway although dependent upon PRRs rather than antibodies for the initial triggering of serine protease activity. Upon recognition of

PAMPs, mannose binding lectin (MBL) and other lectins associate with serine proteases analogous to those in the C1 complex of the classical pathway. Here also serine protease activation leads to C2 and C4 cleavage, and formation of the C3 convertase C4bC2a.

The alternative pathway is mechanistically different to the classical and lectin pathways, and leads to formation of an alternative C3 convertase. Spontaneous hydrolysis of C3 produces an activated thioester bond to which the protein Factor B binds, resulting in cleavage of Factor B into fragments Bb and Ba catalysed by Factor D. The Bb fragment bound to hydrolysed C3 possesses C3 convertase activity, cleaving C3 into C3a and C3b. Subsequently where C3b associates with pathogen surfaces it binds to Bb forming C3bBb, the predominant convertase of the alternative pathway. C3bBb requires stabilisation by an additional factor called properdin or Factor P. The alternative pathway is considered a dual system with a role as a recognition and activation pathway similar to the classical and lectin pathways, and a role as an amplification system for these [276].

The three activation pathways described above are represented in the upper part of Figure 5.1. Whilst the initial triggers and complement proteins involved are different, all three pathways result in formation of a C3 convertase complex for cleavage of C3 to C3a and C3b fragments. Cleavage of C3 initiates the three major effector arms of the complement response; mediated by the membrane attack complex (MAC), opsonins, and anaphylotoxins.

5.1.1.2 The three major effector functions of the complement system

The MAC is an organised assembly of activated complement proteins, capable of forming a stable pore of up to 10 nm in diameter in bacterial outer membranes. Pore formation destroys the osmotic stability of the target, inducing cell lysis. Upon cleavage of C3, the C3b fragment is deposited on bacterial cell membranes where it interacts with C4bC2a and C3bBbC3b, forming C5 convertases. C5 cleavage exposes a binding site for C6, leading to a C5bC6 complex which serves as the foundation for pore assembly. Integration of the complement protein complex into the phospholipid bilayer occurs upon joining of C7, and fully stable pore formation is completed upon association of further complement proteins C8 α , C8 β and C9.

Host cell complement receptors enable phagocytes to bind and ingest material bound by activated complement fragment opsonins. Complement receptors on erythrocytes are

important for transport of immune complexes to the macrophage system for clearance, while receptors on neutrophils and monocytes mediate phagocytosis on binding of complement protein fragments C4b, C3b, iC3b, C3dg and C1q. In addition to directly activating phagocytes through complement receptors, complement on pathogen surfaces provides sites for antibody binding to further encourage clearance.

The complement protein fragments C3a, C4a, and C5a are collectively called anaphylotoxins and act as potent pro-inflammatory signalling molecules. Highly homologous in amino acid sequence, their many overlapping functions include promoting release of other inflammatory mediators to amplify the response, increasing vascular permeability, inducing smooth muscle contraction, and recruitment of leukocytes.

In addition to the innate immune functions described, more recent evidence demonstrates the importance of complement for inducing adaptive immunity, both in B and T cell-mediated processes [275]. Studies in which mice were depleted of systemic C3 demonstrated that complement is involved in shaping the B cell repertoire [277]. Multiple mechanisms linking complement proteins and B-lymphocytes have now been uncovered. The formation of a complex between the B cell receptor (BCR), complement receptor 2 and C3 activation fragments has been shown to influence several aspects of B cell biology, in particular by substantially lowering the threshold required for B-lymphocyte stimulation through the BCR [278]. Complement interaction with T cells has also been demonstrated however the mechanisms are less well understood than for B cells [279].

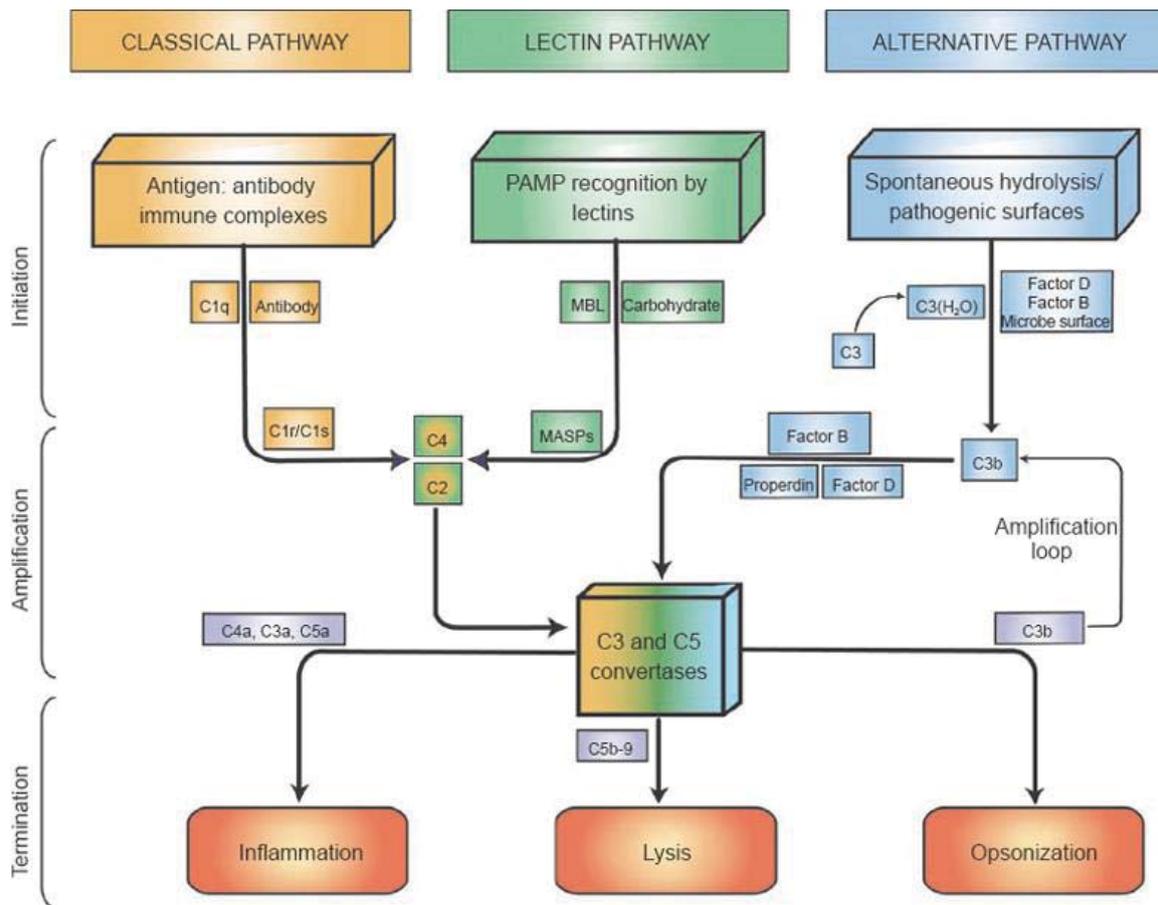


Figure 5.1. Complement activation pathways and major effector functions of activated complement. Three pathways of complement activation, the classical, lectin and alternative pathways converge in formation of a C3 convertase, leading to three major effector functions. In the classical pathway C1q in the C1 complex binds the Fc region of antigen-bound antibody, activating the serine proteases C1r and C1s to cleave C2 and C4. Activation of the lectin pathway is initiated by binding of MBL to pathogen surface carbohydrate motifs, activating the MBL-associated serine proteases (MASPs), again leading to cleavage of C2 and C4. Finally the alternative pathway is initiated by spontaneous hydrolysis of C3, an early C3 convertase forming as a result of cleavage of Factor B by the constitutively active serum protease Factor D, stabilised by properdin. C3b formed by this convertase binds available Factor B creating new C3bBb in a positive feedback loop. The three effector functions triggered by C3 convertase production are performed by different collections of complement proteins. Small protein fragments produced in complement cleavage; C3a, C4a and C5a; are potent inflammatory signalling molecules. C5b-9 associate to form a membrane-spanning pore for target cell lysis and opsonins such as C3b bind surfaces of foreign targets to facilitate phagocytosis and antibody binding. Figure taken from [280].

5.1.1.3 Regulation of the complement cascade

Tight control of the complement pathway is required to ensure the restriction of activation to surfaces of foreign cells, and generation of a proportionate response rapidly terminated on resolution of infection to prevent damage to nearby host cells. Regulatory

mechanisms act at multiple stages to inhibit the complement cascade, examples of which are outlined in Figure 5.2.

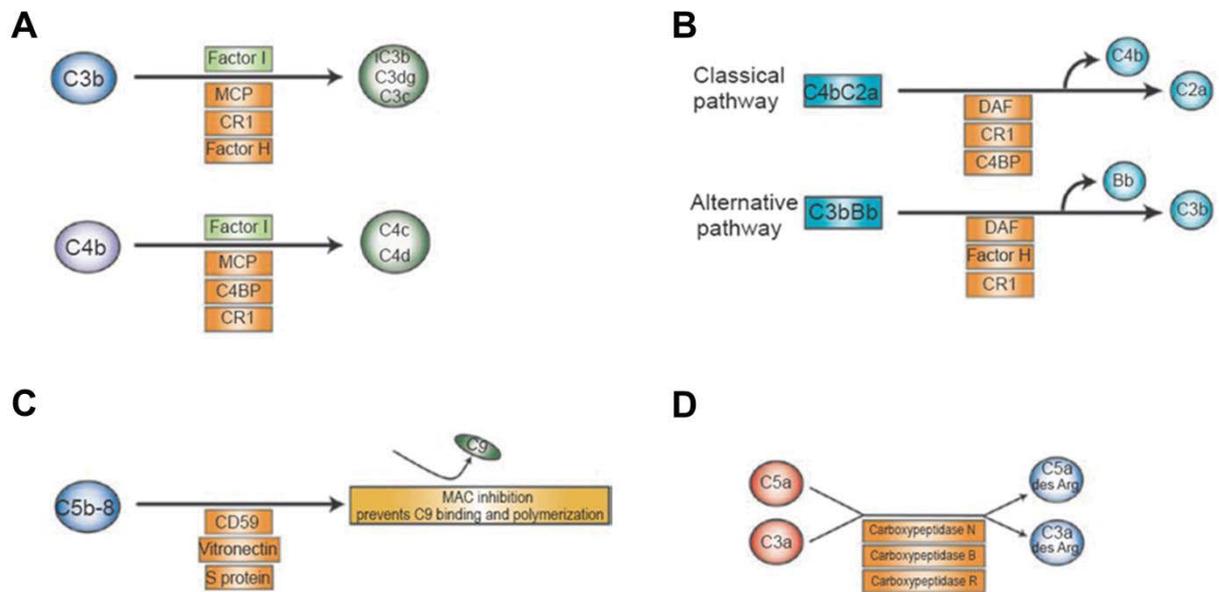


Figure 5.2. Complement activity is controlled by multiple regulatory mechanisms. (A) Serine protease Factor I cleaves complement activation products C3b and C4b into inactive fragments in the presence of proteins which identify host cells; including host cell membrane intrinsic proteins such as MCP, and soluble proteins which bind components specific to host membranes such as Factor H. Cleavage prevents the formation of the C3 convertase on host cells. (B) Convertase inhibitors and decay accelerating factors inhibit assembly and shorten the half-life of existing convertases. (C) During extensive complement activation MAC inhibitors limit complex assembly and pore formation. (D) Serum carboxypeptidases remove N-terminal arginine residues from anaphylotoxins when these are no longer required to limit their interaction with cognate receptors.

5.1.1.4 Proteolytic cleavage of the major complement protein C3

C3 is central to the three complement activation pathways and therefore critical to the effector functions of the complement system, making it an important target of complement regulatory mechanisms as outlined above. C3 is also the most highly abundant complement protein, present in blood plasma from healthy individuals at around 2 mg/ml. The proteolytic processes which act upon C3 during complement activation have been studied in detail. Awareness of the functional roles of C3 fragments, generated during processing both to activate and terminate C3 activity, is essential in order to understand the impact of C3 in a particular tissue or condition. Figure 5.3 outlines the steps involved in the processing of C3 and the activities of the intermediate forms produced.

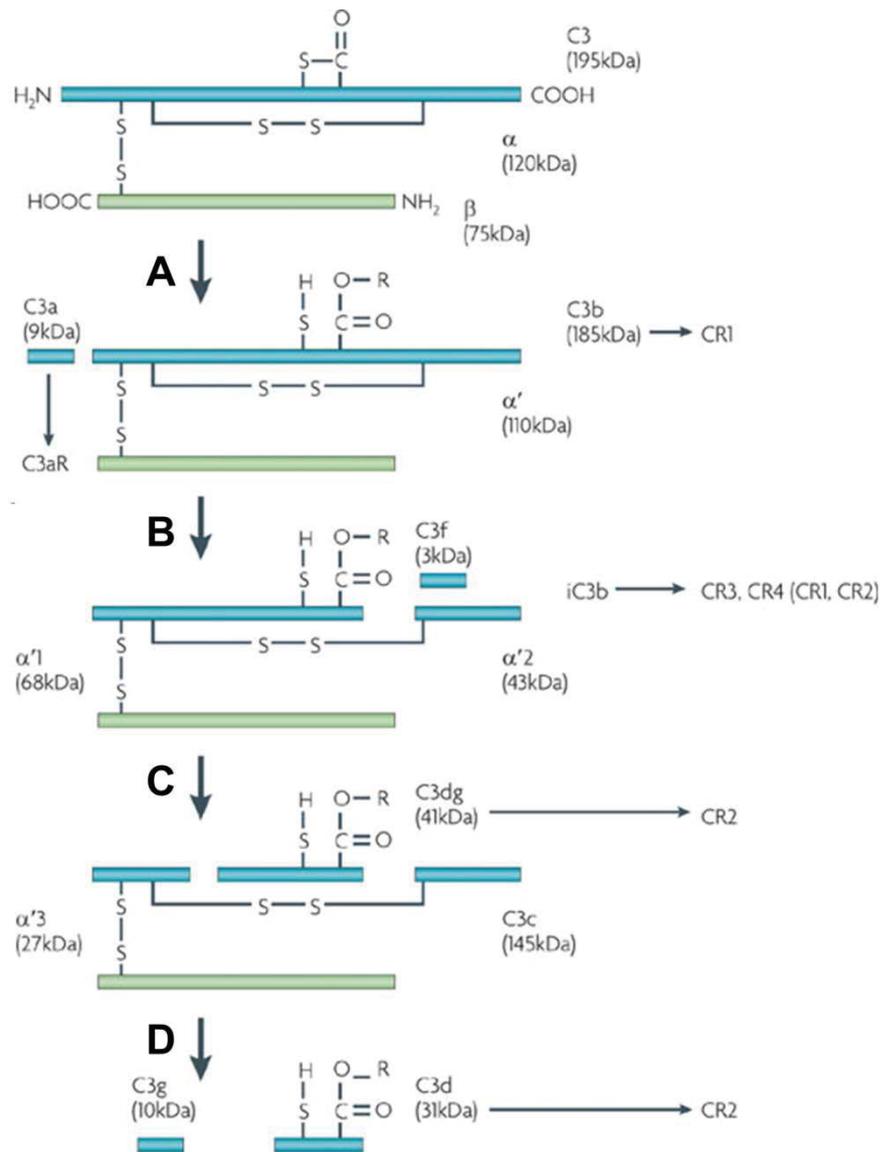


Figure 5.3. Proteolytic processing of C3 is important for both activation and termination of effector functions. Circulating C3 is constructed of disulphide bonded α and β chains. C3 convertase activity in (A) produces the 9 kDa anaphylotoxin C3a, and the larger fragment C3b with an exposed thioester bond susceptible to hydrolysis or bonding with surfaces including those of microbes. C3b interacts with Factor B to stimulate the alternative pathway, with complement receptor 1 (CR1) for activity as an opsonin, or with C3 convertases to form C5 convertases. When inhibition of complement activity is required C3b undergoes cleavage by Factor I to release the small fragment C3f and produce iC3b. Although unable to form convertases, the cleavage product iC3b retains the ability to be bound by complement receptors to mediate phagocytosis of complement-coated pathogens. In (C) Factor I mediates further degradation of iC3b to C3c which is released and C3dg, which remains bound to antigen and interacts with B cells and dendritic cells through CR2. Finally, inflammatory proteases mediate degradation of C3dg to C3g, which is released while C3d remains bound to antigen and, like C3dg, also mediates activation through CR2. Figure adapted from [281].

5.1.1.5 Complement control of *Salmonella*

Both cell-free complement-mediated killing and opsonization of bacteria by complement fragments have been shown to be important in the control of *Salmonella* infection in certain contexts. A study into the mechanisms controlling *S. Typhimurium* in human blood investigated the kinetics of complement activation, terminal complex formation and phagocytosis. Following rapid opsonization of *Salmonella* by C3 and specific antibodies phagocytosis was shown to be initiated immediately whilst the onset of terminal pathway activity is more delayed. This work suggests that slow formation of the membrane attack complex results in significant numbers of bacteria avoiding complement-mediated lysis, instead being taken up by phagocytes including macrophages and transported to an intracellular environment more favourable for their survival [282]. A further discovery was the requirement of *Salmonella*-specific antibody for complement mediated killing mechanisms. This finding is in line with work showing specific antibody is essential for protection of African children against iNTS infection, in part explaining the observed age distribution of childhood iNTS [283].

Studies have shown that the structure of the *Salmonella* O surface antigen is important in determining the rate at which activation of the alternative complement pathway, and subsequent opsonization of *Salmonella* by C3 occurs. In mice infected with *S. Typhimurium* by IP injection more rapid opsonization of bacteria in an O-antigen dependent manner leads to faster uptake of *Salmonella* by phagocytic cells, resulting in enhanced control of infection. This effect is absent when *Salmonella* are delivered intravenously due to the large excess of C3 in the blood [284, 285].

Important differences between humans and mice have been demonstrated in activity of the complement cascade. Over half a century ago it was shown that mouse serum lacks bactericidal activity against many microorganisms effectively killed by complement-dependent lysis following treatment with serum from humans or other mammals [286]. Recent work to investigate the absence of *Salmonella* killing by mouse serum demonstrated restoration of bactericidal activity by the addition of human complement factors, showing that the lack of killing arises from differences in the complement factors themselves rather than the *Salmonella*-specific antibody response [287]. Despite the apparent absence of cell-free complement-mediated killing in mice, C1qa^{-/-} mice are less able to control *S. Typhimurium* following IP and intravenous infection, indicating that other effector mechanisms mediated by

the classical complement activation pathway are important in the murine response to *Salmonella* infection [288]. Of note, differences in the degree of opsonization of *Salmonella* by complement may contribute to observed mouse strain-specific differences in control of *Salmonella* [289].

5.1.1.6 Local complement protein synthesis

The high concentrations of complement proteins in plasma and rapid stimulation of production in the APR led to the view of complement as a systemic system of proteins synthesised in the liver [290]. However, studies demonstrating the presence of complement proteins at sites where penetration of plasma is poor have suggested that local synthesis of complement may also be important, opening a new area of research. A huge range of cell types have been shown to be capable of synthesising complement components *in vitro*, for example production of MAC has been demonstrated in cultured monocytes [291]. Yet the *in vivo* contribution of individual cell types and the functional importance of complement produced outside the liver is poorly understood. In strong support of local complement synthesis is a study finding intense staining for complement in a type of glial cell in and around areas of pathology in inflammatory brain disease. A hepatic origin for complement here is unlikely due to exclusion of plasma by the blood brain barrier [292]. Multiple studies also support the kidney as a site of local complement production with kidney cell types shown to produce complement during inflammation, in addition to infiltrating monocytes and macrophages [293]. Local complement production has also been implicated in the autoimmune condition rheumatoid arthritis [294].

5.1.1.7 The role of complement in the gastrointestinal mucosa

Relatively little work has addressed the possibility of local complement production in the intestinal mucosa, or has investigated the function of intestinal complement in health and disease. Evidence for local production of complement in the intestine is largely provided by IBD studies, with the effect of bacterial infection upon complement production receiving minimal attention [295]. Two studies of complement in the host response to *C. rodentium* infection described later in this section find an important protective role. However more work is needed to understand how these findings relate to enteric infections more generally.

Support for the local production of complement proteins in the intestinal mucosa is reviewed in [295]. Briefly, several studies have provided evidence for expression of complement components at the transcriptional level using immortalised intestinal cell lines, primary cells extracted from intestinal tissue and human tissue biopsies [296, 297]. Further studies confirm the presence of complement pathway proteins in intestinal cells or tissue, including [298] (MBP), [299] (C3b and MAC proteins) and [300] (C3 and C3d). The presence of MAC components in intestinal tissue is a subject of conflicting reports. However, the majority of studies which address this topic report MAC proteins are not detected, and as such lytic functions of complement do not play a major role in the mucosa. Local production of MAC proteins has been detected at extraintestinal locations [301]. In contrast to terminal pathway components, the presence of both C3 and C3 activation fragments is relatively well supported, as is increased concentrations of these proteins in IBD.

Despite the studies described above the precise cellular sources of intestinal complement proteins *in vivo* remain unresolved. During infection and inflammation vascular permeability is increased and the normal barrier function of the epithelium impaired. Passive leakage of serum proteins, including those of the complement cascade, into surrounding tissue and the gut lumen is likely. Further work is required to determine the contributions of serum complement leakage and active intestinal protein production toward the total complement protein present. Further, complement proteins are produced by infiltrating leukocytes, for example properdin is stored in neutrophil granules for release on cytokine or anaphylotoxin stimulation, and the relative contribution of complement produced by immune cells is also an important question [302].

Whilst studies of complement in IBD are in agreement regarding increased production of complement proteins in inflammation, both protective roles and detrimental effects of complement have been proposed. For example $C3^{-/-}$ mice were protected against DSS-induced colitis, and deficiencies in the anaphylotoxin receptors C3aR and C5aR were also protective [303, 304]. Conversely C5aR deficiency was shown to result in more severe disease in a DSS colitis chronic inflammation model. Similarly C5-deficient mice also displayed more extensive inflammation compared with wild type controls in DSS-induced colitis [305, 306]. Deficiencies in early complement pathway proteins (C1 - C4) are associated with inflammation and autoimmune disease as a result of defective immune complex clearance

[307]. Additional work is required to understand the basis of these seemingly contradictory findings.

The relative importance of the three major complement activation pathways in the intestine is also unknown. Factor B is one of the most transcriptionally-induced proteins in epithelial cells in IBD, suggesting the alternative pathway may be important under inflammatory conditions [308]. The interaction of IgG antibodies with C3-coated bacteria in [309] indicates the classical activation pathway may also be important.

Two published works have investigated the role of complement in infection with an intestinal pathogen, *C. rodentium*. Both studies suggest the presence of complement-mediated protective effects. The earlier of these studies indicates a potential role for luminal complement in conjunction with complement fixing-isotypes of IgG. Binding of C3 and serum IgG to *C. rodentium* in the intestinal lumen was demonstrated by flow cytometry, and colocalisation of C3b and IgG at the epithelial surface of the mucosa by immunofluorescence staining. Finding C3 and IgG inhibited the growth of *C. rodentium in vitro* led to a model in which classical pathway activation leads to opsonization of *C. rodentium* by complement, in turn directing Th1-dependent antibody responses against *C. rodentium* antigens, in particular production of IgG2c subtypes. Using *C3^{-/-}* mice and depletion of systemic C3 by cobra venom factor the presence of complement was shown to be an important determinant of infection severity.

Published earlier this year, a study of *C. rodentium* infection in properdin-deficient mice demonstrated a protective role for this C3/C5 convertase-stabilising factor. Properdin was shown to be important for generation of the anaphylotoxin C5a and subsequent stimulation of epithelial cell IL6 production. In the absence of properdin mice displayed exacerbated colitis and increased bacterial colonisation [310].

In summary, evidence from studies of autoimmune disease and infection support important biological functions for the complement system in the intestinal mucosa. However the source of intestinal complement, pathways involved in activation, and effects upon susceptibility to intestinal pathology are areas requiring attention.

5.2 Aims of the work described in this chapter

Regulation of complement during *S. Typhimurium* infection in the streptomycin mouse model was examined in RNAseq and proteomic datasets described in Chapters 3 and 4. Upregulation of C3 in caecal tissue was verified by Western blotting. Further, Western blotting was also employed to detect activation of C3 in caecal tissue and plasma from *S. Typhimurium*-infected mice. The localisation of C3 protein (including C3 fragments) in caecal tissue was investigated using immunofluorescence staining approaches.

5.3 Results

5.3.1 Complement upregulation during *S. Typhimurium* infection in the gastrointestinal mucosa

A highly significant association of complement-related pathways with transcripts and proteins upregulated in *S. Typhimurium* SL1344 infection in caecal tissue was described in Chapter 4. Abundance values, and fold changes compared with naïve tissue with associated p-values, for the four complement pathway genes significantly upregulated at both the level of RNA and protein are displayed in Table 5.1.

	RNAseq			MS		
	Transcript abundance	Log2 fold change	Adjusted p-value	Protein abundance	Log2 fold change	Moderated T-test p-value
C3	78079	4.50	1.29E-48	0.77	4.52	1.66E-06
C4b	16675	3.12	1.73E-39	0.09	Infected only (6)	-
Cfb	31644	3.05	2.73E-52	0.02	Infected only (5)	-
Cfh	3309	2.21	1.25E-27	0.08	Infected only (13)	-

Table 5.1. Complement pathway genes upregulated both in RNA and protein in *S. Typhimurium*-infected caecal tissue. Transcript abundance is the baseMean output from DESeq2 (the average of the normalised count values over all samples). Protein abundance is the normalised summed top three peptide intensities for each protein averaged over all samples. For proteins detected only in *S. Typhimurium* SL1344-infected samples the number in brackets in the column ‘log2 fold change’ denotes the number of MS runs in which the protein was detected.

C3 is the most highly abundant complement protein in blood. We find C3 to be the most abundant complement protein and most highly expressed transcript in caecal tissue. Furthermore of all complement and complement pathway factor genes detected by RNAseq, C3 is the most highly upregulated transcript. C3 is also highly upregulated at the level of protein; of proteins quantified in both naïve control and *S. Typhimurium* SL1344-infected samples, C3 was the fifth most highly upregulated. The dramatic increase in levels of C3 upon infection led us to investigate C3 further by Western blotting and immunofluorescence staining.

Significantly increased both in protein and RNA in addition to C3 are the C3 convertase-forming components C4b, Complement Factor B (Cfb), and the convertase inhibitor and decay accelerating factor, Complement Factor H (Cfh). Highly upregulated at the level of RNA, the proteins encoded by these three genes were present in ≥ 5 of the total 20 MS runs for infected caecal tissue extracts, and were below the threshold of detection in naïve samples.

Complement proteins C1qb, C2, and C8b, and negative regulators of complement activation, Decay Accelerating Factor (CD55) and Vitronectin were also detected by MS. These proteins were detected in ≤ 4 of the 40 MS runs in total for naïve and infected samples, however detection of C1qb and Vitronectin exclusively in infected samples (≥ 3 runs), and the non-significant upregulation of C2 in infection are in line with the upregulation of complement activity in infection. C8b and CD55 were detected in a single MS run only.

5.3.2 Detection of C3 and C3d by Western blotting

Western blotting for C3 was performed on intestinal content and faeces from *S. Typhimurium*-infected and naïve control mice to investigate possible translocation into the intestinal lumen as reported previously in *C. rodentium* infection [309]. An anti-human C3c antibody with cross-reactivity for mouse protein was used, reported to bind C3c (~ 145 kDa) and larger forms of C3 which contain this fragment (C3, C3b and iC3b). Caecal tissue protein extracts from *S. Typhimurium* SL1344-infected and naïve control mice at day 4 PI were investigated by Western blotting with the C3c antibody also. Extracted liver proteins and serum from untreated mice were used as positive controls. Tissue extracts and plasma from *S. Typhimurium* SL1344-infected and naïve control mice at day 4 PI were tested for complement activation by blotting with an anti-C3d antibody. The C3d antibody is reported to

bind the stable activation product C3d (~ 31 kDa) and larger C3 fragments containing the polypeptide chain liberated as C3d by proteolytic processing.

5.3.2.1 Failure to detect C3 in luminal content or faeces by Western blot

Multiple attempts were made to detect C3 in luminal content and faeces obtained from *S. Typhimurium*-infected and naïve control mice by Western blotting with anti-C3c antibody. An example blot is shown in Figure 5.4. While serum controls consistently produced a polypeptide between 100 and 150 kDa, no polypeptide of this size or smaller was detected in intestinal content or faecal protein extracts. As SDS-PAGE was performed under reducing conditions, the disulphide bond linking the 120 kDa α -chain and 75 kDa β -chain of C3 should be broken. Accordingly the polypeptide detected in serum is the correct size for the C3 α -chain. In extracts from naïve control mice a single polypeptide in the region of 250 kDa was detected, significantly larger than full length C3 (~ 195 kDa). The detection of this product exclusively in naïve control mice, which have a comparatively rich microbiota, suggests the product might potentially arise from cross-reactivity of the polyclonal antibody with a microbial protein. Immunoprecipitation and MS identification could be performed for confirmation of the band identity. The polyclonal C3c antibody used in Figure 5.4 consistently produced high background and dark marks on blots making interpretation difficult. Based on scrutiny of several blots performed with this antibody the diagonal lines in lanes 4 and 8 appear to be regions of non-specific binding.

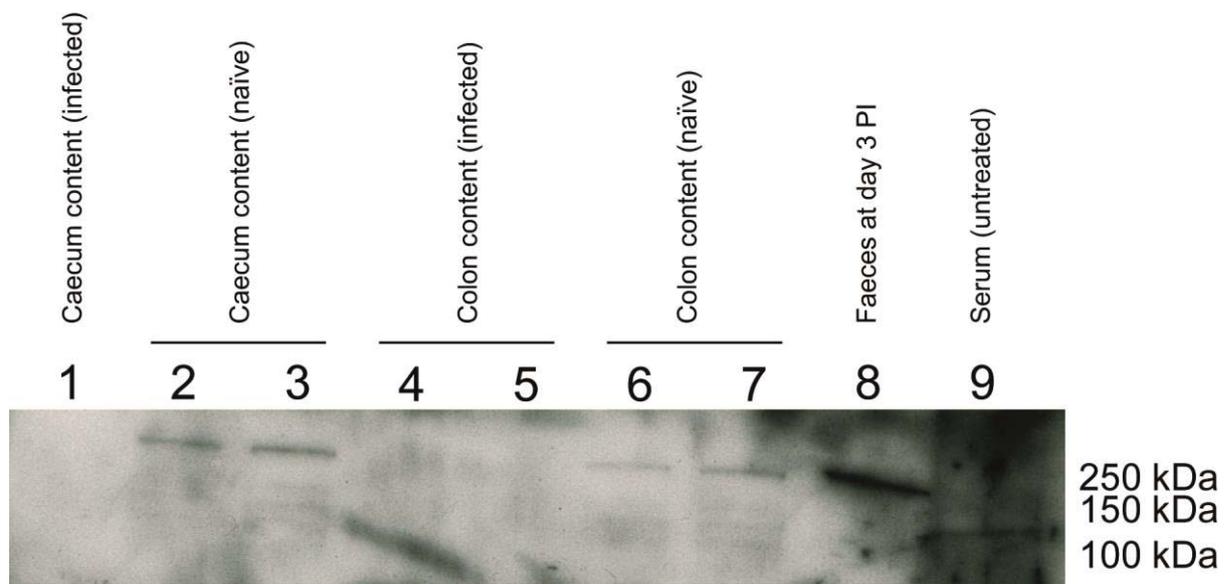


Figure 5.4. Western blotting for C3 in intestinal content and faeces. Western blot produced using Dako human C3c antibody. Lanes are numbered as follows: 1 = caecum content extract from an *S. Typhimurium* SL1344-infected mouse, 2 & 3 = caecum content extracts from individual naïve mice, 4 & 5 = colon content extracts from individual *S. Typhimurium* SL1344-infected mice, 6 & 7 = colon content extracts from individual naïve mice, 8 = protein extract from faeces collected at day 3 PI from an *S. Typhimurium* SL1344-infected mouse, 9 = serum from an untreated mouse. All lanes were loaded with 40 µg protein/lane with the exception of lane 9 in which 20 µg serum protein was loaded. Blot image was produced with a 2 min exposure.

5.3.2.2 C3 and C3 activation fragments in caecal tissue and plasma

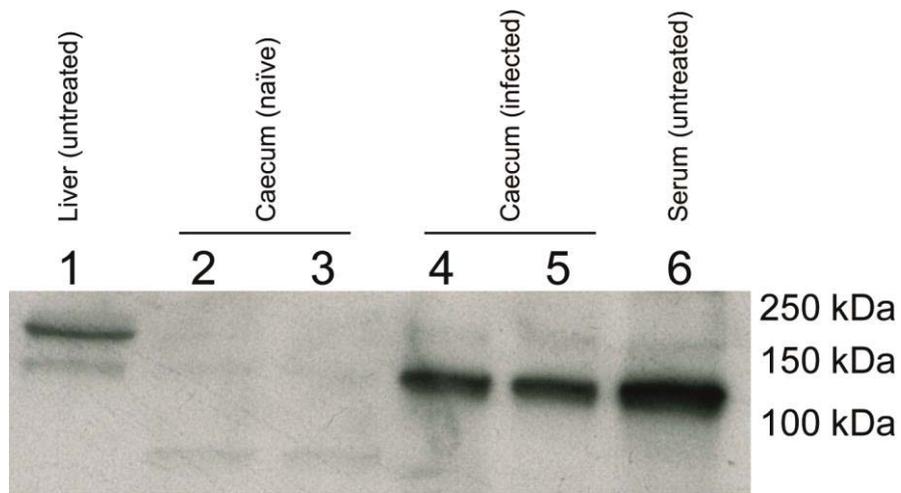
Western blotting was performed on caecal tissue extracts from *S. Typhimurium*-infected and naïve control mice using anti-C3c antibody (Figure 5.5A). A dark product at ~120 kDa was detected in *S. Typhimurium*-infected tissue extracts whereas in naïve control extracts this was barely visible. These findings are in agreement with the MS data.

Having demonstrated the presence of significant quantities of C3 in infected caecal tissue we next investigated the activation state of caecal C3 using an anti-C3d antibody (Figure 5.5B). At around 100 kDa a product corresponding to full-length C3 was detected in all samples. As with the C3c antibody the C3 polypeptide was clearly and consistently more concentrated in caecum extracts from mice at day 4 PI with *S. Typhimurium* compared with naïve controls. At around 30 kDa a polypeptide corresponding to the C3 activation product C3d was detected exclusively in *S. Typhimurium*-infected caecum extracts. The C3d product

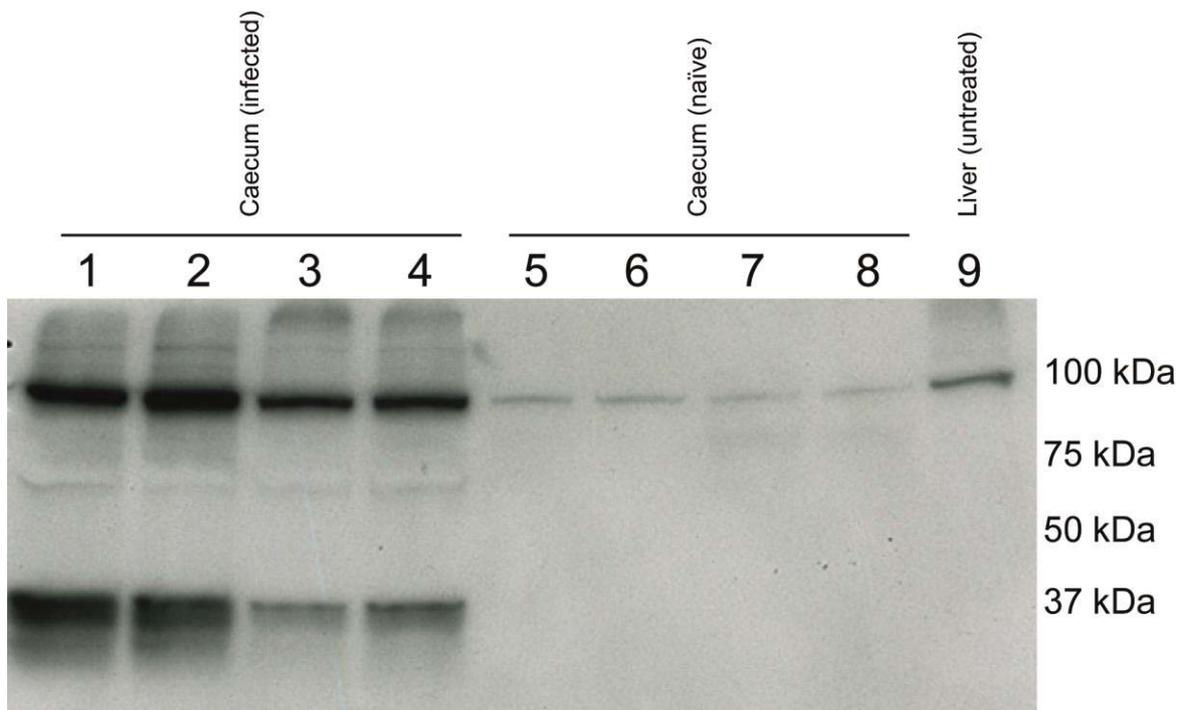
was observed to be consistently fainter than the signal corresponding to full-length C3, indicating that much of the C3 present in infected tissue remains as the full-length form.

We were interested to determine whether C3 activation could be detected in blood upon *S. Typhimurium* infection. Figure 5.5C shows a blot produced with anti-C3d antibody and plasma samples from *S. Typhimurium*-infected and naïve control mice at day 4 PI. The signal intensity for full length C3 appeared highly similar between *S. Typhimurium*-infected and naïve control plasma samples. Densitometry analysis indicated slightly lower intensity for naïve control mice however this was not statistically significant (student T-test). C3d signals were dramatically darker in plasma samples from infected mice; densitometry analysis indicated levels of C3d approximately 100-fold higher in infected plasma versus naïve plasma for which a C3d signal was visible. Both full-length C3 and the C3d fragment were detected at higher levels in plasma compared with tissue.

A



B



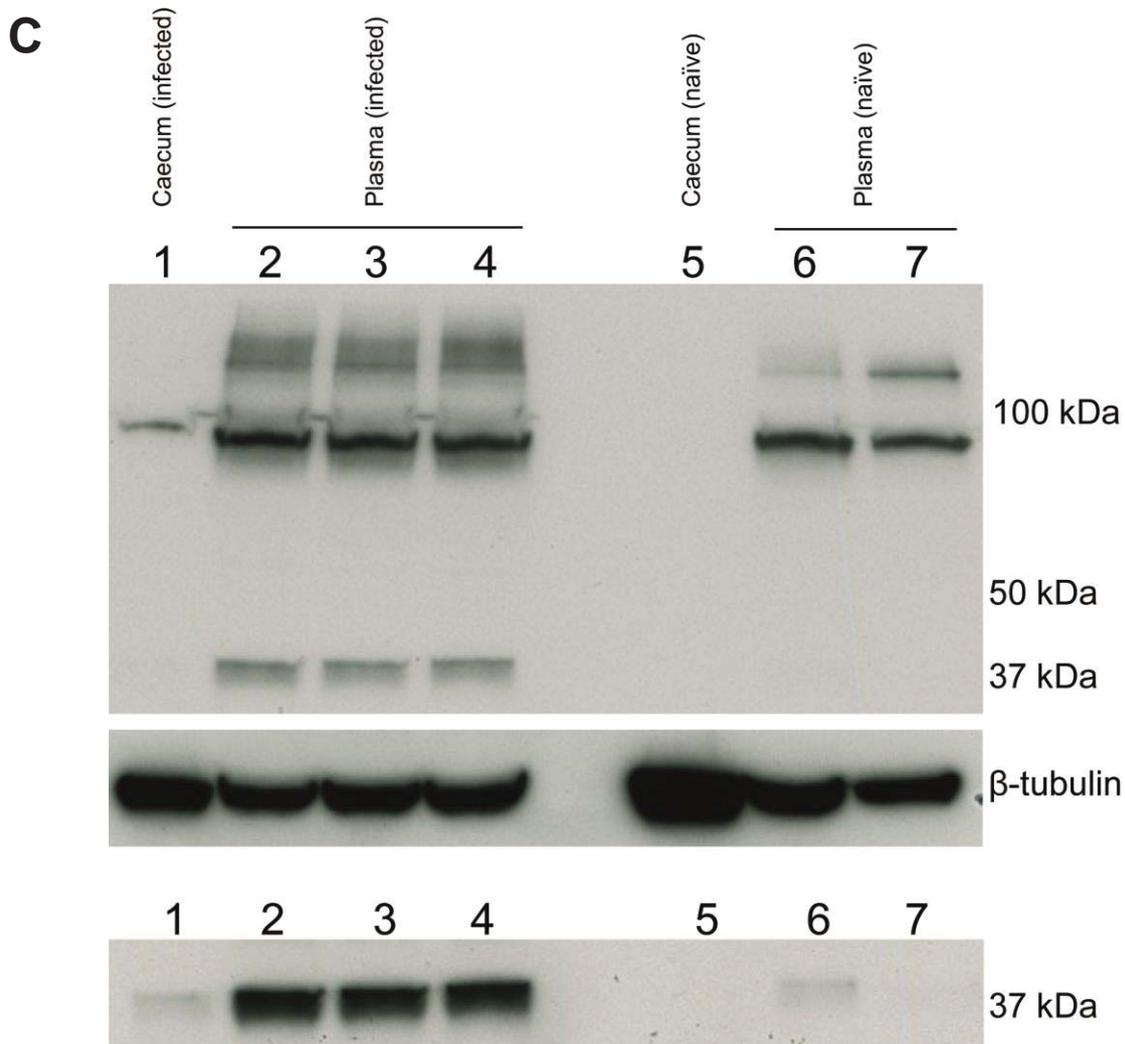


Figure 5.5. Western blotting for C3 and the activation fragment C3d. (A) Western blot produced using Dako human C3c antibody to compare effect of *S. Typhimurium* infection on C3 in caecal tissue. Lanes are numbered as follows: 1 = liver protein extract from an untreated mouse, 2 & 3 = caecal tissue extracts from individual naïve mice, 4 & 5 = caecal tissue extracts from individual *S. Typhimurium* SL1344-infected mice, 6 = serum from an untreated mouse. All lanes were loaded with 30 μ g protein/lane with the exception of lane 6 in which 20 μ g serum protein was loaded. Blot image was produced with a 1 min exposure. (B) Western blot produced using R & D systems mouse C3d antibody to compare effect of *S. Typhimurium* infection on C3 and C3d in caecal tissue. Lanes are numbered as follows: 1 - 4 = caecal tissue extracts from individual *S. Typhimurium* SL1344-infected mice, 5 - 8 = caecal tissue extracts from individual naïve mice, 9 = liver extract from an untreated mouse. All lanes were loaded with 30 μ g protein/lane. Blot image was produced with a 2 min exposure. (C) Western blot produced using R & D systems mouse C3d antibody to compare effect of *S. Typhimurium* infection on C3 and C3d in plasma, with β -tubulin loading control. Upper blot image was produced with a 1 min exposure. Lower blot image shows \sim 30 kDa region of the same blot following a 2 min exposure. Lanes are numbered as follows: 1 = caecal tissue extract from an *S. Typhimurium* SL1344-infected mouse, 2 - 4 = plasma from individual *S. Typhimurium* SL1344-infected mice, 5 = caecal tissue extract from a naïve mouse, 6 & 7 = plasma from individual naïve control mice. All lanes were loaded with 20 μ g protein/lane.

5.3.3 Localisation of C3 in caecal tissue by immunofluorescence staining

Immunofluorescence staining was performed to investigate the localisation of C3 in caecal tissue from *S. Typhimurium* SL1344-infected and naïve control mice. 5 µm sections of frozen tissue were stained with C3 and C3d antibodies for the detection of C3 and activation fragments. Staining for the endothelial cell marker CD34 was used to indicate the position of caecal vasculature, and staining for *Salmonella* common structural antigens (CSA) for visualisation of *Salmonella*.

Figure 5.6 shows typical images of co-staining for C3 and CD34. C3 staining was more intense and extensively distributed in *S. Typhimurium*-infected mice compared with naïve controls, in line with the detection of increased C3 in infected caecal tissue by MS and Western blotting. Strong colocalisation of C3 and CD34 was clearly observed in infected tissue. In places C3 staining appeared to extend a little beyond the CD34-stained regions, although these inconsistencies could be the result of differences in the intrinsic brightness of the fluorophores used. Staining performed with C3d antibody showed a distribution indistinguishable from the C3 antibody staining.

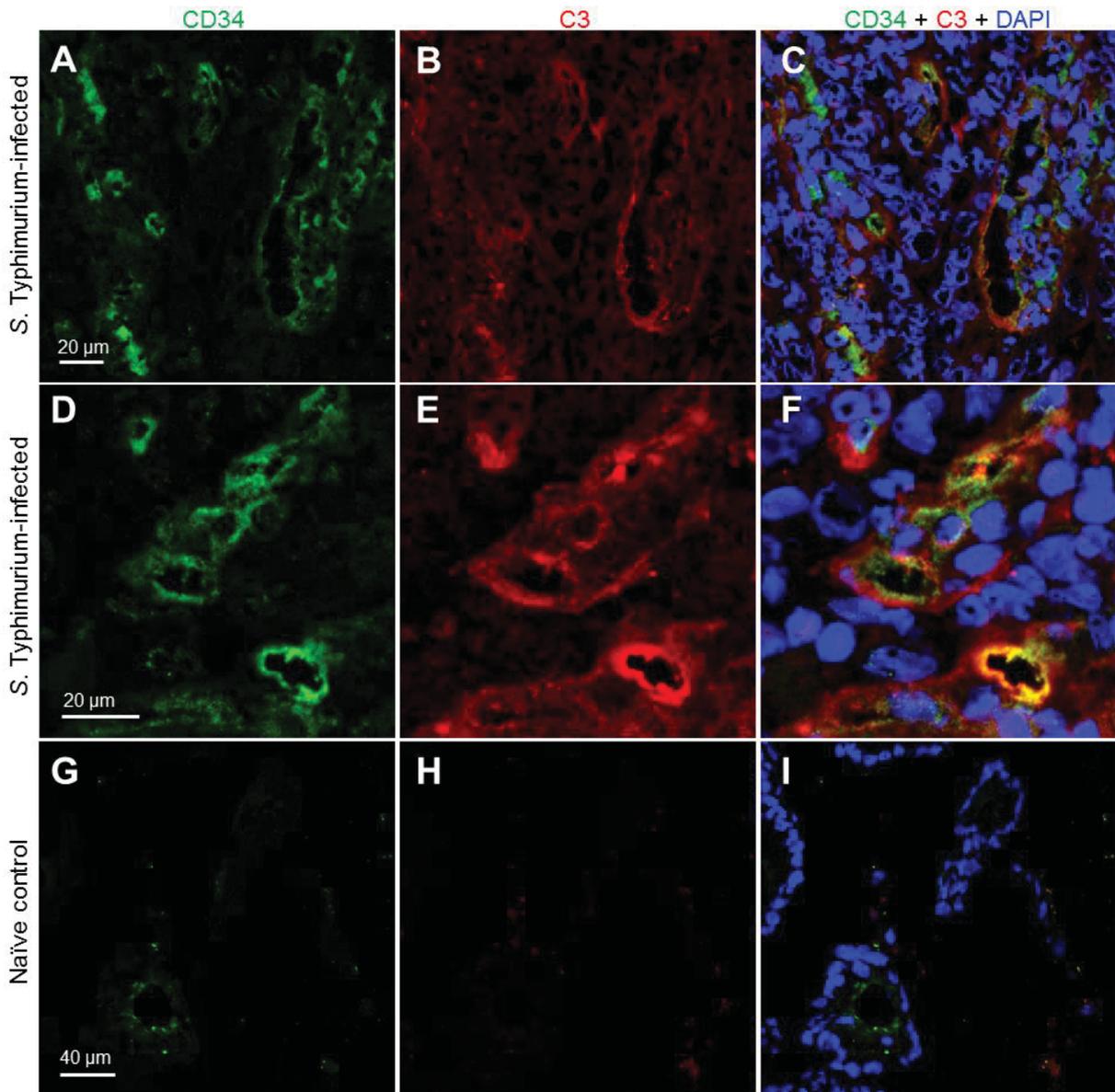


Figure 5.6. Immunofluorescence staining of C3 and CD34 in caecal tissue. 5 µm sections of frozen caecal tissue were stained with C3 and CD34 primary antibodies with appropriate secondary antibodies for C3 staining in red and CD34 staining in green. DAPI was used to stain DNA. Confocal images to the left and in the centre display fluorescence in the individual green and red channels respectively. Images to the right display the overlay of the red, green and blue channels. A - C = *S. Typhimurium*-infected caecum, low magnification, D - F = *S. Typhimurium*-infected caecum, high magnification, G - I = naïve control caecum.

Salmonella CSA staining of 5 µm sections was performed to determine the distribution of complement relative to *Salmonella* within infected caecal tissue. Regions of complement staining were often observed in the proximity of *Salmonella* bacteria but CSA and C3 staining did not appear to be directly overlapping. Individual host cells, possibly

macrophages, brightly stained with CSA antibody were observed in infected caecal tissue though their presence was relatively rare. These cells were located in regions of tissue with bright staining for C3 and were stained for C3 themselves.

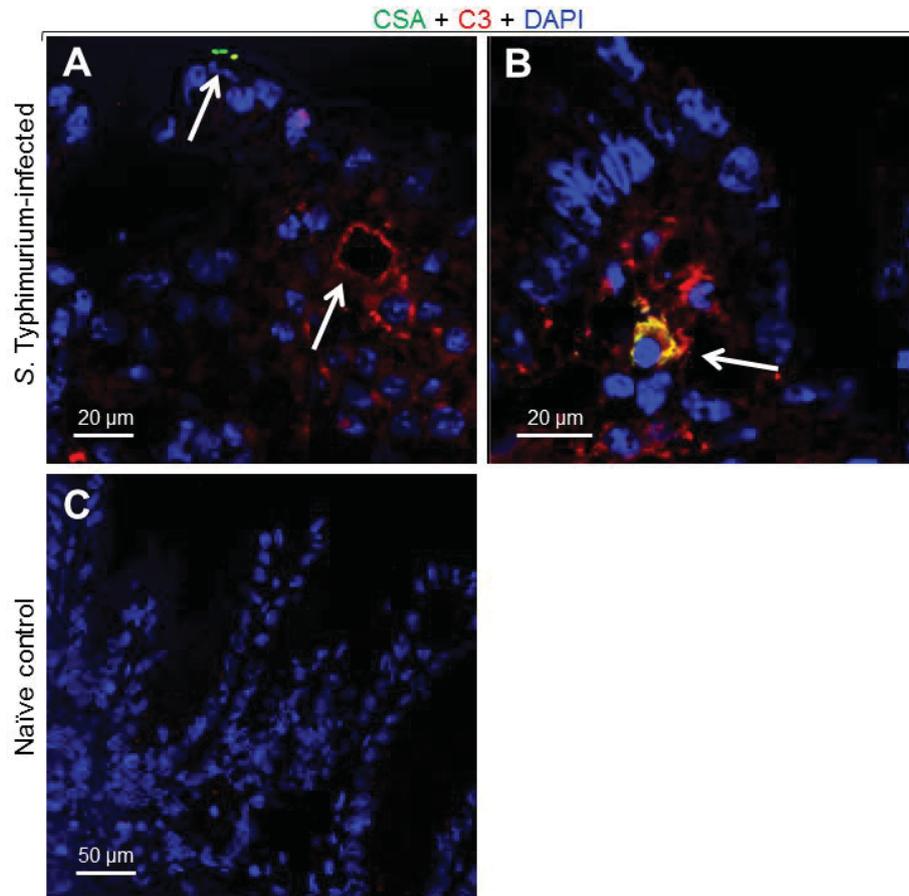


Figure 5.7. Immunofluorescence staining of C3d and CSA in caecal tissue. 5 µm sections of frozen caecal tissue were stained with C3d and CSA-FITC primary antibodies with appropriate secondary antibody for C3d staining in red. DAPI was used to stain DNA. All images are the overlay of red, green and blue channels detected by confocal imaging. A = *S. Typhimurium*-infected caecum, upper arrow points to *Salmonella*, lower arrow indicates a region of bright C3d staining. B = *S. Typhimurium*-infected caecum, arrow points to a cell brightly stained by both C3d and CSA antibodies. C = naïve control caecum.

Whole-mount three-dimensional imaging of caecal tissue demonstrated a greater degree of overlap between C3 and *Salmonella* compared with the staining of 5 µm tissue sections (Figure 5.8). *S. Typhimurium* bacterium associated with epithelial cells at the luminal surface, and in the lamina propria, were brightly stained for C3. Regions of more diffuse C3 staining were also present throughout caecal tissue.

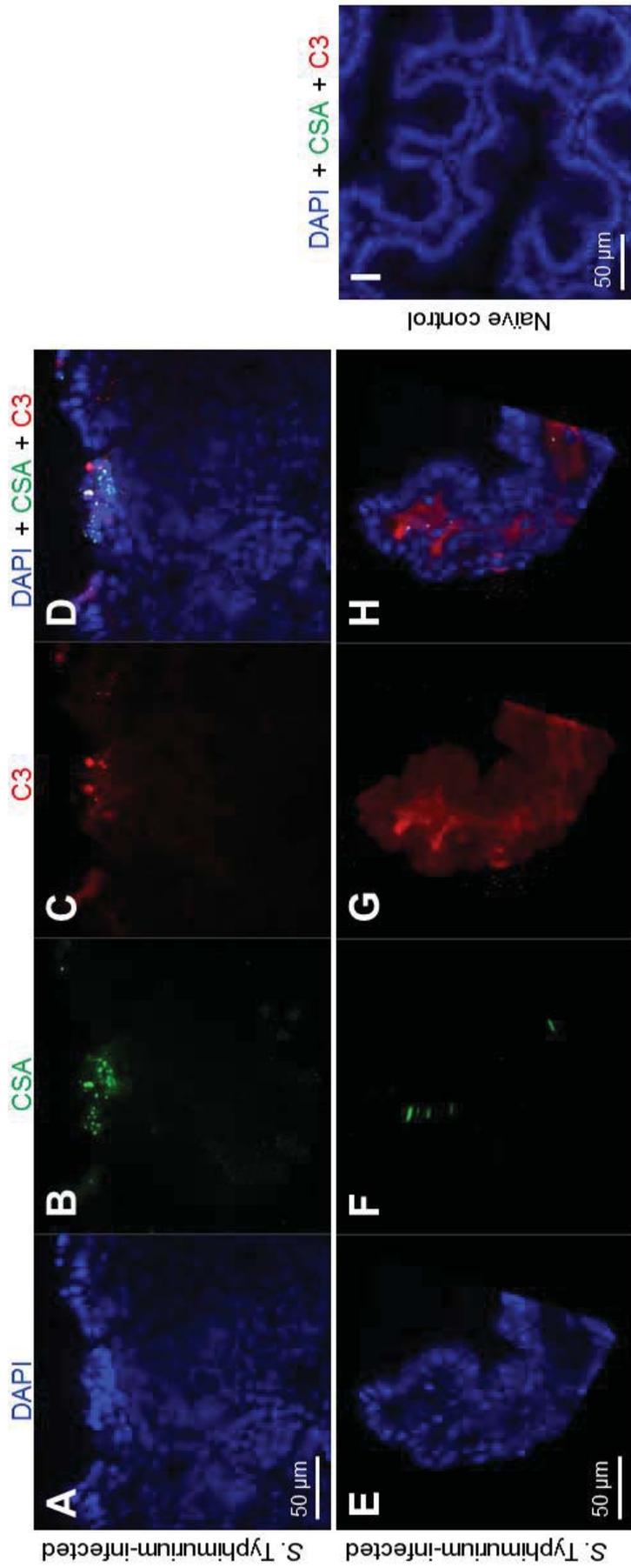


Figure 5.8. Three-dimensional confocal imaging of C3 and CSA in caecal tissue. PFA-fixed caecal tissue was stained with C3d and CSA-FITC primary antibodies with appropriate secondary antibody for C3d staining in red. DAPI was used to stain DNA. From left to right images display fluorescence captured in the individual blue green and red channels. Images D, H and I, display the overlay of the three channels. A - D = *S. Typhimurium*-infected caecum with *Salmonella* associated with the epithelial surface. E - H = villus-like projection in *S. Typhimurium*-infected caecum with *Salmonella* inside the central core. I = naïve control caecum.

5.4 Discussion

The increased abundance of several complement pathway proteins in caecal tissue during *S. Typhimurium* infection suggests complement may play an important role in enteric infection. The source of these proteins may be the intestinal tissue itself, as a result of increased local production. Alternatively, translocation of circulating liver-derived complement proteins into caecal tissue may take place during infection. Moreover the increase in complement pathway proteins likely arises from a combination of these two effects. Past studies support an increase in local complement protein production during inflammation of the intestinal mucosa [311]. Our finding that mRNAs encoding complement pathway proteins are dramatically increased in *Salmonella*-infected caecal tissue suggests an increase in complement production within the caecal tissue may be taking place. However a contribution from systemic complement to total tissue complement protein is not excluded. Given the increase in vascular permeability which occurs upon inflammation, leakage of blood complement proteins into tissue seems likely. Additionally, as MS and Western blotting were performed upon protein extracted from whole caecal tissue, protein from inside blood vessels within the tissue makes an unknown contribution to total complement pathway protein. The impact of increased vascular permeability might be assessed by measuring levels of a stably-produced blood protein similar in size to C3 in naïve and *Salmonella*-infected caecal tissue. An example of such a molecule is creatinine (113 kDa), commonly used as an indicator of renal function in medical practice.

The source of intestinal complement in infection was partially addressed by staining caecal tissue for C3 and C3 stable activation products. The colocalisation of C3 and the endothelial cell marker CD34 indicates C3 is deposited on the walls of blood vessels within the intestinal mucosa, and diffuses into surrounding tissue. However C3 appeared also to be present at sites remote from vessels, perhaps indicating either diffusion or transport of C3 through the tissue, or alternatively the production of C3 *in situ*. Individual cells interspersed throughout the caecum which were brightly stained for CSA also stained for C3. C3 was also observed diffusely localised in the surrounding region. In the CSA-positive cells rod-shaped structures or obvious bacteria were not visible, therefore these cells may represent phagocytes which have ingested and lysed *Salmonella* causing release of CSA into the cytoplasm or its display on the cell surface. Activated immune cells including polymorphonuclear cells and

dendritic cells have been shown to be producers of complement proteins [312, 313], therefore we suggest these cells may represent activated phagocytes, possibly macrophages. Co-staining of intestinal tissue for immune cell surface markers and C3 would confirm this and aid understanding of the contribution of immune cells to C3 production. The work presented in this chapter has provided insight into potential sources of the complement protein C3 in intestinal infection. Further work is required to determine if similar patterns are observed with other components of the complement pathway.

Cell isolation experiments may be a means by which to further address the source of increased C3 in the caecum during infection [314, 315]. Isolation of individual intestinal cell types followed by protein extraction and MS or blotting would allow the contribution of different cell types to be quantified. Early trials of separating intestinal cell types were performed, however contamination of epithelial cell fractions with CD45⁺ cells was a recurring problem, and prohibited the analysis of pure cell populations. Ensuring cells are washed sufficiently to remove complement associated with the cell surface or in surrounding fluid would be essential but difficult to determine.

Cobra venom factor has been used extensively to deplete systemic C3 [309, 316, 317]. Systemic complement depletion could provide insight into the contribution of locally-produced complement protein in *S. Typhimurium* infection. The efficacy of cobra venom factor is not precisely defined and therefore systemic C3 may contaminate tissue extracts. However, examining the differential impact of systemic C3 depletion on naïve and infected tissue may help to further determine the sources of C3.

A previous study demonstrated the presence of C3 on the surface of a bacterial pathogen in faeces [309]. In the study, C3 was detected on the surface of luminal bacteria in C3^{-/-} mice treated with wild type mouse serum, indicating C3 was able to translocate into the lumen from the circulatory system. C3 was also detected on the surface of luminal bacteria from mice treated with cobra venom factor, suggesting locally produced complement may contribute in addition to complement from a systemic source. Despite multiple attempts we were unable to detect C3 in intestinal content or faeces by Western blotting, irrespective of infection. In light of the greater epithelial damage incurred in infection with *S. Typhimurium* compared with *C. rodentium* it seems highly unlikely that systemic leakage of C3 occurs exclusively in *C. rodentium* infection or indeed intestinal production of complement is stimulated specifically by *C. rodentium*. C3 is susceptible to heat degradation and is degraded

by certain species of bacteria, most likely as an immune evasion strategy [318, 319]. The intestinal lumen is rich in enzymatic activity, mediated by both bacterial and host-derived digestive enzymes. Therefore following translocation into the intestinal lumen C3 may be quickly broken down. Future work to investigate the presence of luminal C3 might include flow cytometry analysis of faeces from mice infected with labelled *Salmonella* using C3 antibodies, as previously described for *C. rodentium*.

By Western blotting we detected the stable product of C3 activation C3d in *S. Typhimurium*-infected caecal tissue. C3d contains a highly reactive thioester bond through which it covalently attaches to pathogen surfaces. The importance of an interaction between C3d on pathogen surfaces and the B-cell CR2 receptor in directing the adaptive immune response is well described [275, 320]. The differential intensity of signatures corresponding to C3 and C3d in Western blots indicates the majority of C3 protein present in *Salmonella*-infected caecal tissue remains in the full length form. This is likely the result of a situation in which a massive influx and/or upregulation of C3 floods the tissue with the complement protein, while activation is restricted to the locations of bacteria within the caecum. Indeed we detected colocalisation of C3 and/or activation fragments and *Salmonella* using three-dimensional confocal imaging, but also complement in regions of tissue without bacteria. Regions of colocalisation likely represent sites of C3 activation, though the pathway by which activation may occur is unknown. The apparent lack of commercial antibodies which specifically recognise activated forms of C3 is regrettable; ideally the location of activation products within tissue would be examined by immunofluorescence and compared with the distribution of full length C3. Further work could include co-staining for C3 and complement-fixing Ig subtypes in the intestinal mucosa to investigate the possibility of complement activation by the classical pathway.

Western blotting demonstrates a vast increase in C3d in the plasma of *S. Typhimurium*-infected mice compared with naïve controls at day 4 PI in addition to the increase in caecal tissue. Indeed C3d detected in plasma was shown to form a significantly larger proportion of total plasma protein compared with the proportion of C3d in protein extracted from caecal tissue. On the one hand the observed increase in systemic C3d could occur through the transport away from the tissue of C3d fragments produced by activation of C3 in the intestine. Alternatively C3d detected in plasma may have been generated by C3 activation in the liver and spleen, organs which by day 4 PI in the streptomycin mouse model

carry a considerable burden of *Salmonella* (section 3.3.1). Given that C3 is thought to translocate from vessels into caecal tissue upon infection C3d may travel similarly into the caecal tissue. Therefore the detection of activated C3 fragments in the caecum cannot be attributed conclusively to C3 activation in tissue. Although the presence of a high bacterial burden in the caecum suggests complement activation likely is taking place, similar to C3 we cannot determine the proportion generated *in situ* and outside the tissue. As C3d is reported to be covalently bound to pathogen surfaces, C3d moving from the circulation to the intestinal tissue would likely be accompanied by bacteria. Based on the much greater signal intensities observed for plasma compared with caecal tissue from *S. Typhimurium*-infected mice this might suggest that more C3 activation is occurring in the blood than within the intestinal mucosa.

A statistically significant increase in blood plasma C3 at day 4 PI with *S. Typhimurium* compared with uninfected mice was not detected. For confirmation blotting should be repeated with loading of reduced amounts of protein. Relative to the high concentration of C3 in plasma in the naïve condition a small change in the total amount of C3 protein will be difficult to detect.

The systemic spread of *S. Typhimurium* in the later stages of the streptomycin mouse model is a significant deviation from the most commonly observed outcome of *S. Typhimurium* infection in humans; self-limiting gastroenteritis. The streptomycin mouse model, and in particular later time points in the infection, is poorly suited for studying the presence and role of complement activation in such cases of gastroenteritis. It might be argued that the course of infection in the mouse model more closely approximates bacteremic and extraintestinal focal *Salmonella* infections caused by non-typhoidal *Salmonella* species. Invasive infections occur in approximately 5% of people infected with non-typhoidal *Salmonella* serotypes and are more prevalent in certain groups such as the elderly and those with immunosuppressive conditions. *C. rodentium* studies can provide some insight into the role of complement in a non-invasive pathogen but studying an invasive enteric pathogen is more difficult. A variation of the streptomycin mouse model in which an *sseD* mutant derivative of *S. Typhimurium* gives rise to a self-resolving infection might be used to study complement in intestinal *Salmonella* infection. Alternatively infection of *Salmonella* resistant mouse strains (*Nramp1*^{+/+}) in which intestinal inflammation occurs without development of systemic infection might provide a more appropriate system for these studies. [50].

In summary the work presented in this chapter raises the possibility of important roles for complement in bacterial infection of the intestinal mucosa. Further experiments to explore these roles will include *S. Typhimurium* infection of $C3^{-/-}$ mice. C3-deficient mice were previously reported to be highly susceptible to *C. rodentium* infection [309]. Infection of $C3^{-/-}$ mice in the streptomycin mouse model will determine if the increase in C3 in the caecum upon *S. Typhimurium* infection is required for protection from inflammation. However due to the eventual systemic spread of *S. Typhimurium* in the streptomycin mouse model, study of multiple time points or passive transfer of C3 in serum to $C3^{-/-}$ mice, may be required to separate the effect of C3 deficiency on intestinal inflammation, from effects upon systemic control of *Salmonella*.

6 Signatures of the streptomycin mouse model: investigating host susceptibility using mutant mice

6.1 Introduction

6.1.1 Systematic screening of knockout mice for the detection of novel phenotypes in infection susceptibility

The International Knockout Mouse Consortium (IKMC) was established in 2006 with the goal of mutating all protein-coding genes in the mouse, and making a library of mutant embryonic stem (ES) cells available to the scientific community [321]. The generation of mutant ES cell lines is based on a targeted approach using homologous recombination to introduce the desired mutation into the genome. The main strategy adopted in the IKMC is introduction of a DNA fragment containing a *lacZ* marker gene directly upstream of the critical exons of the targeted gene, thereby impeding gene function without deleting the targeted gene (Figure 6.1). The allele disrupted as described is referred to as the ‘tm1a’ allele. Breeding of mouse tm1a lines with lines expressing Flp and Cre recombinases can be used to either knock out the gene entirely or inactivate the gene in a tissue-specific manner through generation of a conditional line [322].

Mutant mouse lines are systematically screened for phenotypes at centres in the International Mouse Phenotype Consortium (IMPC). The Wellcome Trust Sanger Institute (WTSI) is a founding member of the IMPC and has made a substantial contribution to the colossal task of screening the ~ 20,000 mutant lines which represent all known and predicted mouse genes. Indeed in 2013 production of more than 900 mutant lines by the WTSI mouse genetics project (MGP) was reported, with close to 500 lines screened for viability and fertility, and 250 tested for adult phenotypes [323].

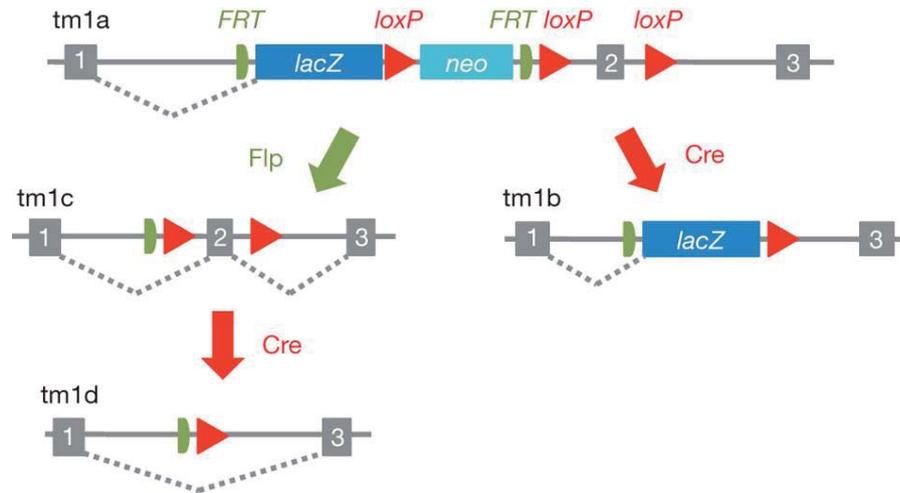


Figure 6.1. Mutant allele design used in the IKMC programme to mutate all protein coding genes in the mouse strain C57BL/6N. Generated by homologous recombination the tm1a allele contains a *lacZ* cassette and floxed cassette for neomycin resistance inserted into the intron preceding the critical exon (here exon 2). Conversion of a tm1a allele to a conditional allele with restored gene activity (tm1c) is performed with Flp recombinase. Crossing tm1c lines with lines in which expression of Cre recombinase is driven by a tissue-specific promoter results in frameshift mutations, forming the allele tm1d in these specific tissues. Alternatively Cre recombinase can be used to delete the neo cassette from the tm1a allele to generate a *lacZ*-tagged knockout allele (tm1b). Image taken from [322].

MGP phenotyping involves a comprehensive array of tests to examine many major organ systems of the adult mouse. Testing of the immune system is an important component of the screening process and is coordinated by the Infection, Immunity and Immunophenotyping (3i) consortium (<http://www.immunophenotyping.org/>). Both ‘observational’ and ‘challenge’ elements contribute to the overall assessment of immune function. In observational tests the immune compartment of organs in healthy animals is examined in flow cytometric analysis of spleen, lymph nodes, bone marrow and peripheral blood. The challenge component investigates the response of mutant mice to a range of infectious organisms and chemical stresses. At the time of writing 115 mutant lines have been screened for susceptibility to the parasitic worm *Trichuris muris*, 437 screened with DSS, 385 with influenza virus and 1,153 lines with *S. Typhimurium* M525 (TetC).

The immune challenges of the 3i program were selected to evaluate the function of broad aspects of innate and adaptive immunity. Multiple phenotypes can be interrogated using different challenges and the inclusion of pathogen challenges can identify distinct aspects of

the immune system that are involved in the responses to these infections. Also, a consideration involved in the design of the individual challenges was minimising variability in the measured outcomes between wild type control mice. Limited variability permits the use of small numbers of mice in screening; reducing the time and cost given the large number of lines, and in keeping with the principles of the ‘three Rs’ framework for the humane use of animals in research (replacement, reduction, refinement).

Observational screen			Challenge screen		
Organ / test	no of mice	readout	Challenge	no of mice	readouts
Spleen, MLN, bone marrow	3F3M	flow cytometry	Salmonella	4F4M	bacterial counts
Blood (6 weeks) Blood (16 weeks)	4F4M 7F7M		Influenza	3F3M	BAL washes
Antinuclear antibodies	3F3M	microscopy	Trichuris	6F	worm burden
Ear epidermis	2F2M		DSS	4M	histology
Cytotoxic T lymphocytes	2F2M	kill assay	Tuberculosis	selected	bacterial counts

Figure 6.2. Summary of current screening for phenotypes relating to infection and immunity directed by the 3i consortium. Image taken from <http://www.immunophenotyping.org/methods>.

The MGP infection challenges have successfully recognised phenotypes in genes with previously published roles in infection such as *MyD88* and *NF-κB*, [130, 324, 325]. Many genes with no prior involvement in immunity have also been implicated, and for some mechanistic details have been uncovered. For example a phenotype in the *Salmonella* challenge was identified for *Mysm1* mutant mice with further investigation finding *Mysm1* plays an important role in maintenance of bone marrow stem cell function and haematopoiesis [326]. In addition to defective bone formation mutants in gastric intrinsic factor (*Gif*) were found to exhibit an altered susceptibility to both *C. rodentium* and *S. Typhimurium*

challenges, with further secondary phenotyping demonstrating metabolic and immunological defects associated with vitamin B12 deficiency [327].

The MGP *Salmonella* challenge tests the involvement of host genes in protection in an *S. Typhimurium* typhoid model. Four female and four male mice around eight weeks of age are infected by intraperitoneal (IP) injection of 10^5 *S. Typhimurium* strain M525 expressing the tetanus toxin fragment C (TetC). At days 14 and 28 PI half of the infected mice are culled and the liver is removed for bacterial enumeration and histopathological examination. A blood sample is collected and plasma prepared for analysis of titres of antibody (total Ig, IgG1 and IgG2a) against TetC by ELISA.

We proposed that a *S. Typhimurium* infection challenge, which employs the streptomycin model of *S. Typhimurium* gastroenteritis, might provide complementary information on the role of genes in host protection against *Salmonella* infection. Host genes have been reported previously which affect susceptibility to infection via the oral route exclusively [328]. Such genes may not be identified in the existing *Salmonella* IP challenge. Streptomycin pre-treatment and *S. Typhimurium* infection could be performed as part of secondary phenotyping of lines alongside phenotyping involving other tests of immunity. Additionally the streptomycin-*Salmonella* challenge might be performed on mice with mutations in genes which were prioritised for gene targeting following identification in, for example, GWAS studies of IBD.

In this work *tm1a* mouse lines with mutations in *IL10rb*, *IL22ra1* and *BC017643* were selected for oral challenge with *S. Typhimurium* on the basis of phenotypes observed in infection with *C. rodentium* and *Salmonella* at the WTSI, and in the case of *IL10rb* and *IL22ra1* due to the known importance of these genes in immunity. In all tests of the primary phenotyping pipeline *IL10rb*^{tm1a/tm1a} mice were comparable to wild type mice with the exception of infection with *C. rodentium* where the *IL10rb* mutants displayed delayed bacterial shedding and increased colon weights at day 28 PI. *IL22ra1*^{tm1a/tm1a} mice were phenotypically normal in all tests with exception of the skin epidermal structure, as assessed by whole-mount staining of the tail, and the *C. rodentium* challenge [329]. *BC017643*^{tm1a/tm1a} displayed a number of phenotypes in the primary screening pipeline, most notable of these was strongly increased susceptibility to infection in the *Salmonella* challenge. At day 3 PI *BC017643* mutant mice showed severe clinical signs of illness, with *Salmonella* CFU in the liver and spleen around three orders of magnitude greater compared with wild type mice.

These findings suggest a critical role for this unpublished gene in defence against bacterial infection. Using the streptomycin model we sought to investigate a potential role for BC017643 in mucosal bacterial defence. The functions of the cytokine receptor subunits IL10rb and IL22ra1 in host protection are introduced here briefly.

6.1.2 The role of IL10rb in infection and immunity

The IL10 family of cytokines consists of interleukins including 10, 20, 22, 24, 26, 28 (α and β) and 29, and is defined based on the structure and location of the genes which encode them, the structures of these proteins, and the receptor complexes they activate. While these features are relatively conserved across all family members the functional roles they perform are diverse. All IL10 family cytokines induce activation of target cells through heterotetrameric receptor complexes and all may activate cells through complexes which contain a common IL10rb subunit combined with cytokine-specific partners. Accordingly *IL10rb* is constitutively expressed in most cell types, while individual partner subunits display more restricted expression [330].

Produced by a wide range of immune cell types, IL10 is a major anti-inflammatory cytokine, guarding against excessive immune responses and limiting the extent of tissue damage during infection. Its many functions include suppressing the secretion of pro-inflammatory cytokines such as IL1, IL6, IL12 and TNF α , and Th1 cytokines such as IL2 and IFN γ , and controlling differentiation and proliferation of macrophages and B and T lymphocytes [331]. Figure 6.3 summarises the major points of action of this cytokine.

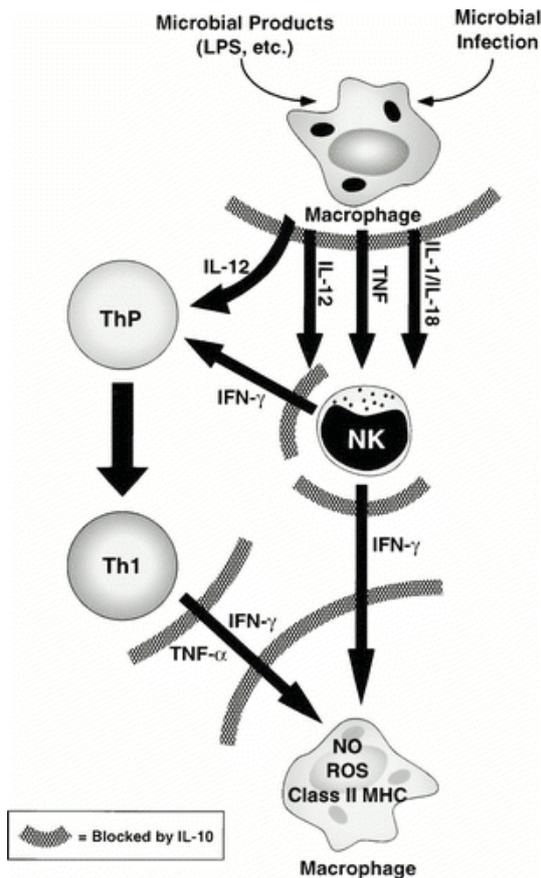


Figure 6.3. IL10 inhibits the major steps in activation of an inflammatory immune response. The detection of microbial molecules by macrophages and monocytes induces a combination of cytokines, which stimulate IFN γ production by NK cells, and polarisation of T cells to the Th1 subtype. Th1 cell cytokines as well as IFN γ from NK cells induce microbial killing mechanisms in macrophages including production of reactive oxygen and nitrogen species, and phagocytosis. IL10 inhibits this response system at all the sites indicated by thick grey curved lines. The combined effect of inhibition at all these individual sites is a large suppression of the overall inflammatory response. Image taken from [332].

The importance of IL10 and receptors in controlling inflammation, particularly in the intestinal mucosa, is illustrated by inactivating mutations which cause monogenic autosomal recessive IBD. Intestinal inflammation resulting from *IL10* mutation is early in onset and only successfully treated by haematopoietic stem cell transplantation to provide immune cells with a functional IL10 signalling pathway [331, 333]. The consequences of human *IL10* mutations are mirrored in mice; a spontaneous colitis phenotype is observed in lines carrying inactivating mutations in *IL10rb* and *IL10* [334, 335].

Whilst the involvement of IL10 pathway mutations in autoimmunity is unequivocal their impact upon control of infection is variable and context dependent; both protection and increased susceptibility have been reported. Suppression of IL10 activity is beneficial during systemic infection of mice both with *Listeria monocytogenes* and *S. Choleraesuis* [336, 337]. IL10-deficient mice display less acute infection and associated inflammation in infection with *C. rodentium*, surprising given the importance of IL10 in restricting inflammatory Th1 and Th17 responses. A second interesting outcome of *C. rodentium* infection of IL10-deficient mice is protection against the onset of spontaneous colitis [338]. Conversely in other cases of

pathogen-induced intestinal inflammation continued Th1 responses in the absence of IL10 can generate severe intestinal pathology; IL10-deficient mice show increased intestinal inflammation and mortality in *Trichuris muris* infection, and loss of IL10 also increases susceptibility to *Helicobacter hepaticus* [339, 340]. During co-infection of mice with malaria and *S. Typhimurium* the parasite infection induces IL10, resulting in reduced inflammatory pathology yet increased systemic *Salmonella* dissemination. This demonstrates how malaria can increase the risk of iNTS infection in Africa through a mechanism reliant upon IL10 [341]. Together these findings demonstrate that the impact of IL10 is highly dependent on factors such as the severity and location of infection, and the specific host factors involved in the response.

6.1.3 The role of IL22ra1 in the intestinal mucosa

IL22 signalling is important for induction of tissue repair and host defence at barriers including the skin, and the mucosae of the airways and intestine. By promoting epithelial proliferation IL22 restores barrier integrity to damaged epithelia. Interaction of antigen presenting cells with microorganisms induces production of IL23, which stimulates a variety of immune cells including IELs, ILCs and Th17 cells to produce IL17 and IL22. IL22ra1 accompanies IL10rb to form the receptor complex for IL22, and is present almost exclusively in non-haematopoietic cells; major targets of IL22 signalling are epithelial cells and keratinocytes. An important function of IL22 in the intestinal mucosa is restriction of commensal bacteria to the intestinal lumen by inducing the production of antimicrobial peptides including Reg3 β and γ , lipocalin-2, and S100A8 and S100A9 to form calprotectin, also by the stimulation of production of mucins within goblet cells [342, 343].

Like IL10, IL22 signalling has been reported to be both detrimental and protective in different infections. Intestinal pathogens such as *Salmonella* which induce an upregulation of antimicrobial peptides via IL22 signalling, exploit this pathway to assist them in overcoming colonisation resistance and therefore IL22-deficient mice are protected in *Salmonella* infection. As described in Chapter 1, *S. Typhimurium* possesses a number of resistance mechanisms for avoidance of the harmful effects of antimicrobial peptides. Indeed *IL22* was one of the most highly induced genes in a previous study of mucosal *S. Typhimurium* infection and was found to be highly induced in infection in the work described here in section 3.3.6 also (log₂ fold change = 4.26) [152].

Conversely, and again due to effects of IL22 signalling on the microbiota, IL22ra1-deficient mice produced at the WTSI were shown to be highly susceptible to infection with *C. rodentium*. Here IL22-induced production of the fucosyltransferase Fut2 is required for generation of oligosaccharides to feed the microbiota during the period of poor nutrient availability, which occurs during infection. In the absence of fucosylation the opportunistic bacterium *Enterococcus faecalis* is able to outgrow the microbiota, damage the epithelium and cause systemic infection [329, 344]. Similarly IL22-deficient mice also display systemic bacterial burden, increased epithelial damage and mortality [345].

Though classically IL22 is thought to protect against infection at body barriers, a recent paper has described a role for IL22 in protection during systemic infection through induction of hepatic C3, and potentially C3 induction at non-hepatic sites also. It is likely that further roles of IL22 remain undiscovered [346, 347].

6.2 Aims of the work described in this chapter

The potential of oral *S. Typhimurium* infection following streptomycin pre-treatment as a secondary phenotyping test in the MGP mutant mouse screening pipeline was examined. Three mutant lines were selected for investigation on the basis of phenotypes observed in primary phenotyping challenges. Mutant mice were investigated for differences in *Salmonella* burden in caecum, colon, liver and spleen, with further phenotyping performed for the *IL22ra1*^{tm1a/tm1a} mutant line. To gain insight into processes controlled by the targeted genes during infection RNAseq was used to investigate the caecal tissue transcriptome of mutant lines.

6.3 Results

6.3.1 Expression of *IL22ra1* and *IL10rb* in mice homozygous for the tm1a mutant allele

In contrast to *IL10rb*^{-/-} mice reported previously, spontaneous development of colitis was not observed in the *IL10rb*^{tm1a/tm1a} mutant line produced at the WTSI [334]. As explained in the introduction section of this chapter, in tm1a mutant alleles the gene of interest is interrupted rather than deleted. Where the splice acceptor in the gene trap cassette is not 100% efficient alternative splicing can result in ‘skipping’ of the gene trap cassette, giving rise to ‘leaky’ expression of the wild type transcript [323]. We hypothesised that the *IL10rb*^{tm1a/tm1a}

allele may be hypomorphic, with sufficient functional *IL10rb* transcripts produced by unwanted splicing events to prevent colitis in mutant mice. Such an effect may be tissue-specific in distribution. Initially we attempted to investigate the extent of IL10rb-deficiency in the *IL10rb*^{tm1a/tm1a} mouse line by Western blotting for IL10rb in cultured macrophages from *IL10rb*^{tm1a/tm1a} and wild type mice. Unfortunately despite efforts with two antibodies directed against mouse IL10rb we were unsuccessful in detection of a polypeptide of the appropriate size for this protein even in wild type mice.

Expression of *IL10rb* and *IL22ra1* in their respective tm1a mutant mouse lines was therefore examined by qPCR and RNAseq. In *IL10rb*^{tm1a/tm1a} mice the *IL10rb* gene is interrupted by insertion of the targeting cassette in the intron 5' to the critical third exon. Commercially available qPCR primers were selected to investigate levels of RNA for regions upstream (exons 1 & 2) and downstream (exons 3 & 4) of the inserted cassette in a range of tissues (Figure 6.4A). Signal from both regions of *IL10rb* was reduced in RNA samples from *IL10rb*^{tm1a/tm1a} mice compared with wild type controls, with a markedly greater reduction in the region of exons 3 & 4. However consistent detection of signal from the region of exons 3 & 4 in RNA samples from *IL10rb*^{tm1a/tm1a} mice indicates potentially incomplete abolition of full-length protein-coding *IL10rb* transcripts in the mutant line.

For further investigation RNA was extracted from bone marrow and peritoneal macrophages obtained from *IL10rb*^{tm1a/tm1a} and wild type mice (n = 3) and qPCR was performed to assess the reduction in RNA in the region of exons 3 & 4 of *IL10rb*. *IL10rb*^{tm1a/tm1a} BMDM displayed a 274-fold reduction in RNA from exons 3 & 4 (normalised against expression of the housekeeping gene *Gapdh*), while a 381-fold reduction was observed in peritoneal macrophages.

Further, differential expression analysis was performed on caecal transcriptome profiles from *S. Typhimurium*-infected wild type and *IL10rb*^{tm1a/tm1a} mice (day 4 PI) generated by RNAseq (n = 5) (further detail in section 6.3.4). An average normalised read count across all samples of 2,873 for *IL10rb*, relative to the median across all genes of 357, indicates *IL10rb* is highly expressed in the infected caecum. *IL10rb* transcripts were decreased in *IL10rb*^{tm1a/tm1a} *S. Typhimurium*-infected caecal samples compared with wild type controls with a log2 fold change of -4.32, equivalent to a ~ 20-fold reduction, with a p-value of 10⁻²⁴³ following correction for multiple testing.

In summary these findings indicate ‘leaky’ expression of wild type *IL10rb* occurs in *IL10rb*^{tm1a/tm1a} mice, with levels of full length transcripts controlled by cell- or tissue-dependent factors. Further work is needed to determine how conditions such as infection impact upon levels of full length *IL10rb* transcripts in these mice.

Possible production of full length *IL22ra1* transcripts in *IL22ra1*^{tm1a/tm1a} mice was investigated with qPCR with primers for the amplification of a region spanning the fourth and fifth exons of the mouse *IL22ra1* gene. A range of tissues were selected for qPCR to assess level of *IL22ra1* expression in healthy wild type and *IL22ra1*^{tm1a/tm1a} mice (Figure 6.4B). No obvious relationship was found between organ levels of *IL22ra1* expression in wild type mice and the reduction in expression in these organs in *IL22ra1*^{tm1a/tm1a} mutants. *IL22ra1* transcripts were virtually undetectable in RNA from the liver and pancreas of *IL22ra1*^{tm1a/tm1a} mice. However in several tissues expression of *IL22ra1* in the mutant line was considerable; for example in the colon expression in samples from *IL22ra1*^{tm1a/tm1a} mice was 15.7% of the wild type level.

RNA extracted from wild type and *IL22ra1*^{tm1a/tm1a} naïve and *S. Typhimurium*-infected (day 4 PI) caecal samples were sequenced. For both the naïve and infected samples, which were analysed in separate sequencing runs, the normalised read count for *IL22ra1* across all samples was greater than the median for all genes, indicating *IL22ra1* expression is substantial. Differential expression analysis of *IL22ra1*^{tm1a/tm1a} and wild type naïve samples produced a log₂ fold change in *IL22ra1* of -1.6 (approximately a 3-fold reduction in the mutant) with an adjusted p-value of 1.6×10^{-8} . In contrast analysis of the transcriptional profiles from *S. Typhimurium*-infected tissue showed *IL22ra1* was not significantly reduced in the mutant (log₂ fold change = -0.44, adjusted p-value = 0.248). The read counts from the individual samples used in this analysis are displayed in Figure 6.4C.

These findings indicate that expression of *IL22ra1* is incompletely abolished in *IL22ra1*^{tm1a/tm1a} mice, and that levels of residual expression are both tissue-dependent, and involve other factors including infection. For detailed investigation of the role of *IL22ra1* the tm1a allele should be converted to a tm1b allele in which the critical exon is deleted. Although we observed substantial levels of residual *IL22ra1* expression in *IL22ra1*^{tm1a/tm1a} mice this has not precluded detection of phenotypes previously, as demonstrated by the increased susceptibility of this line in *C. rodentium* infection [329]. In addition, the majority of mutant animals which are currently phenotyped in the MGP pipeline contain a tm1a allele

in the targeted gene. For these reasons and due to the absence of the time required for production of full knockout mice the tm1a line was studied throughout this work.

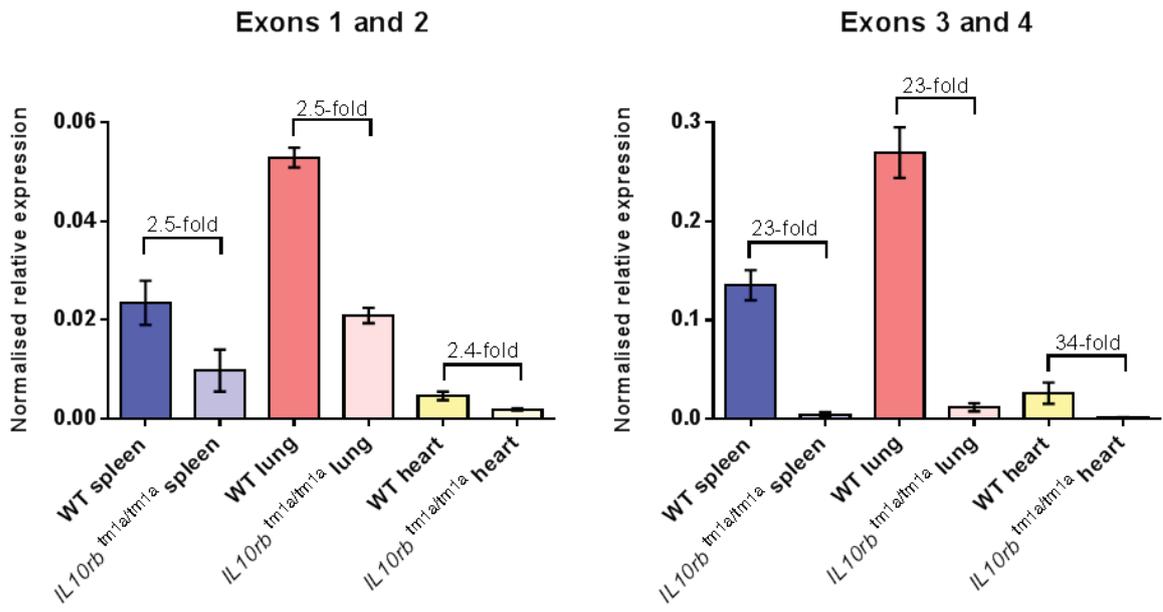
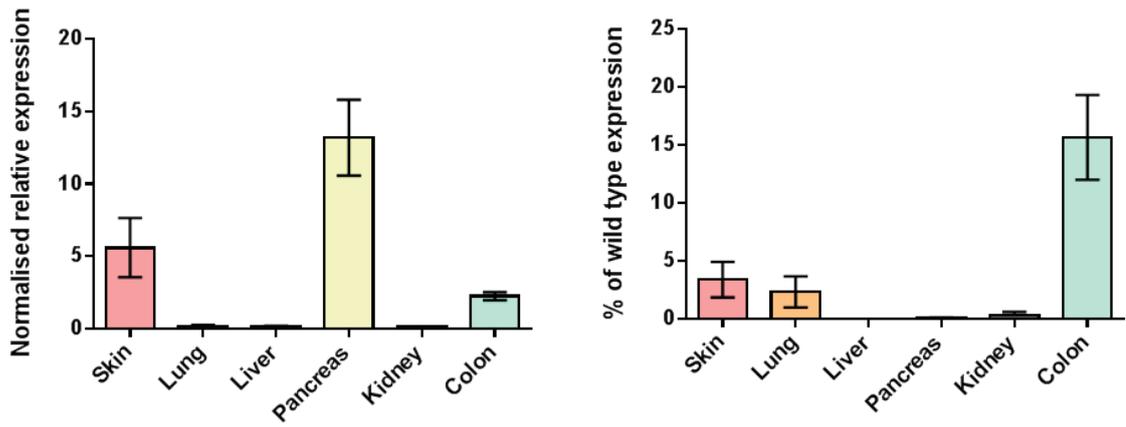
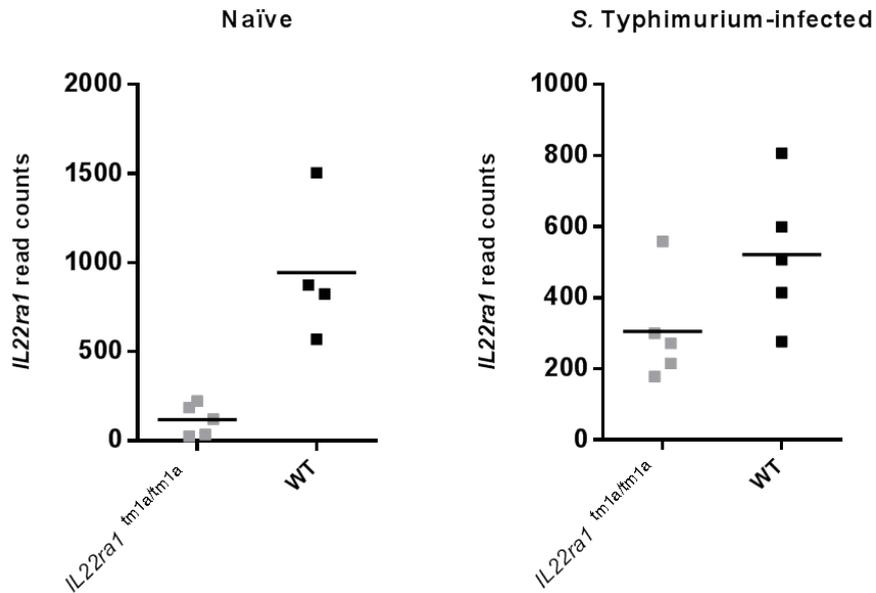
A**B****C**

Figure 6.4. Analysis of *IL10rb* and *IL22ra1* expression levels in *tm1a* mutant mouse lines. (A) Relative expression of regions of the *IL10rb* gene spanning the first and second, and third (critical) and fourth exons, normalised against expression of the housekeeping gene *Gapdh*. Tissue pieces were taken from two wild type (WT) and two *IL10rb^{tm1a/tm1a}* mice. All amplification reactions were performed in duplicate. Error bars display standard error on the mean. Fold reduction between the average expression for wild type and *IL10rb^{tm1a/tm1a}* is displayed for each tissue. (B) Left - Relative expression of *IL22ra1* normalised against expression of the housekeeping gene *Gapdh* in a range of tissues from wild type mice (n = 3). Amplification reactions were performed in triplicate. Right - Percentage of wild type normalised expression levels detected in samples from *IL22ra1^{tm1a/tm1a}* mice (n = 3). Error bars indicate the standard error on the percentage. (C) *IL22ra1* read counts from RNAseq of naïve and *S. Typhimurium*-infected caecal tissue. Naïve and infected samples were sequenced in separate sequencing runs and therefore cannot be compared directly.

6.3.2 *S. Typhimurium* colonisation and dissemination in *IL10rb^{tm1a/tm1a}* mice following streptomycin pre-treatment

IL10rb^{tm1a/tm1a} mice were treated with streptomycin 24 h prior to oral infection with *S. Typhimurium* SL1344 in two separate experiments. Representative data for *Salmonella* CFU in the spleen, liver, colon and caecum are presented in Figure 6.5. The results of both experiments suggest control of *Salmonella* in *IL10rb^{tm1a/tm1a}* mice is comparable to wild type animals.

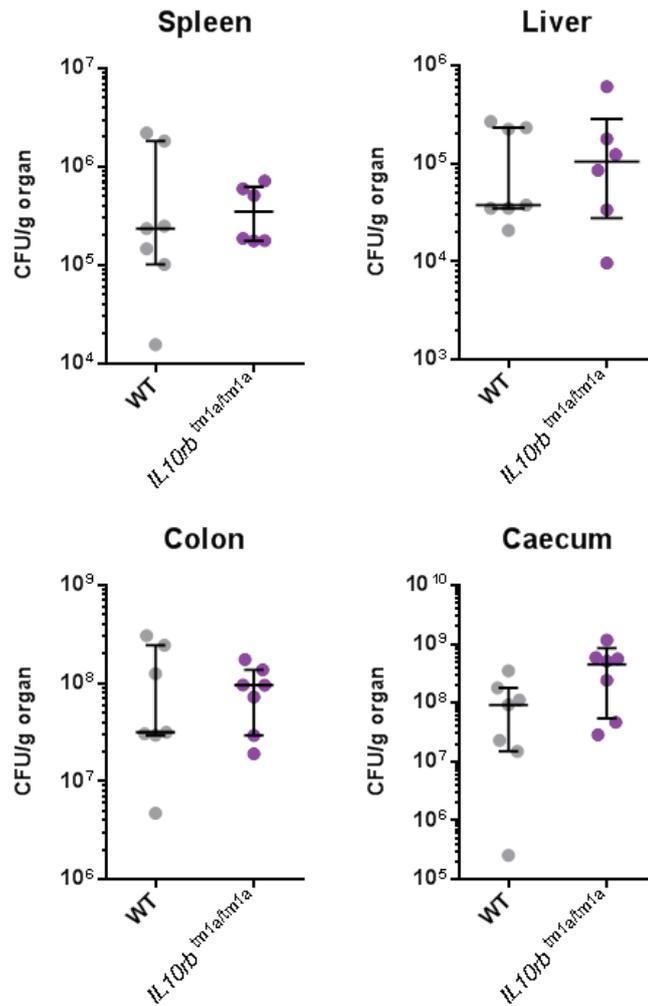
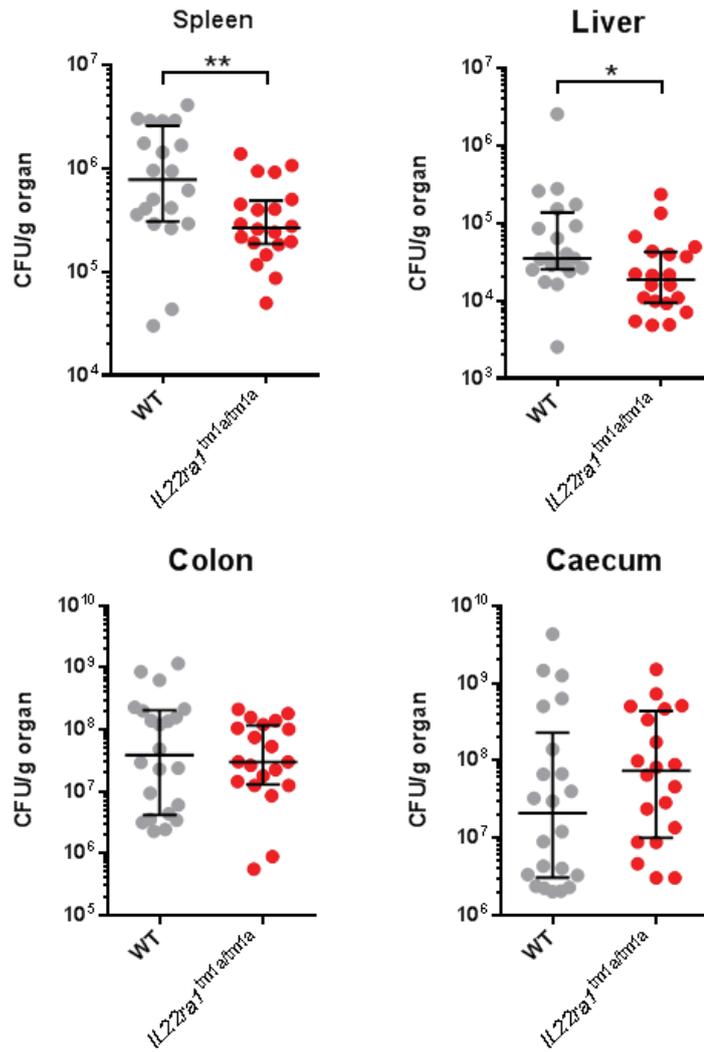
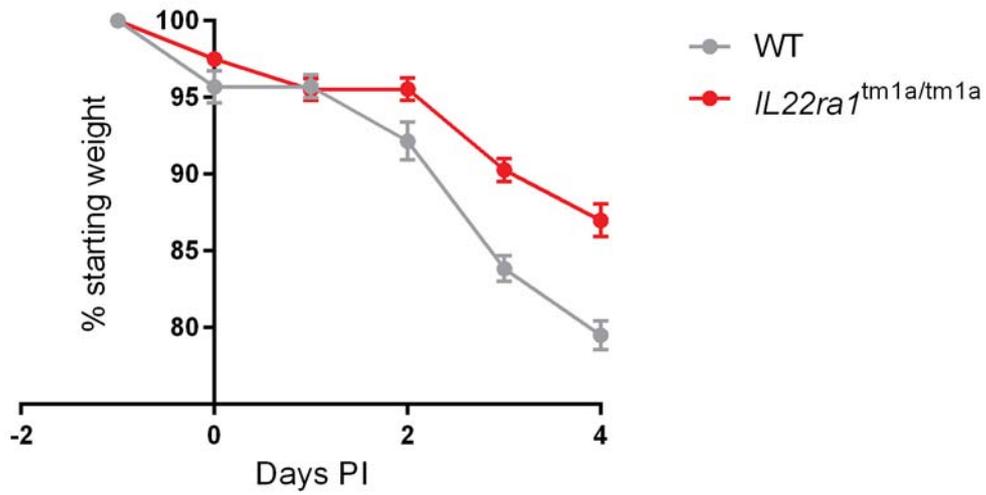


Figure 6.5. *Salmonella* burden in $IL10rb^{tm1a/tm1a}$ mice at day 4 PI with *S. Typhimurium* following streptomycin treatment. *Salmonella* organ CFU from oral infection of wild type (WT) and $IL10rb^{tm1a/tm1a}$ mice with *S. Typhimurium* SL1344 (1.3×10^4 CFU). Bars show median and interquartile range. Significance testing was performed with a Mann Whitney U test.

6.3.3 *S. Typhimurium* infection of $IL22ra1^{tm1a/tm1a}$ mice in the streptomycin model

$IL22ra1^{tm1a/tm1a}$ mice were treated with streptomycin 24 h prior to oral infection with *S. Typhimurium* SL1344 in three separate experiments. Reduced *Salmonella* colonisation of the liver and spleen in $IL22ra1^{tm1a/tm1a}$ mice was observed in multiple experiments, whilst CFU in the intestinal tissues did not appear to be affected by host genotype. Figure 6.6A shows the combined *Salmonella* CFU from the three experiments. Overall *Salmonella* counts in the spleen and liver were significantly reduced with p-values of 0.006 and 0.0142 respectively. $IL22ra1^{tm1a/tm1a}$ mice displayed reduced weight loss compared with wild type mice; a representative weight loss curve is presented in Figure 6.6B. Scoring to reflect the

severity of inflammation in H & E stained tissues from *IL22ra1*^{tm1a/tm1a} and wild type mice indicated no significant difference in tissue pathology between mice of different genotypes (n = 6) (Figure 6.6C).

A**B**

C

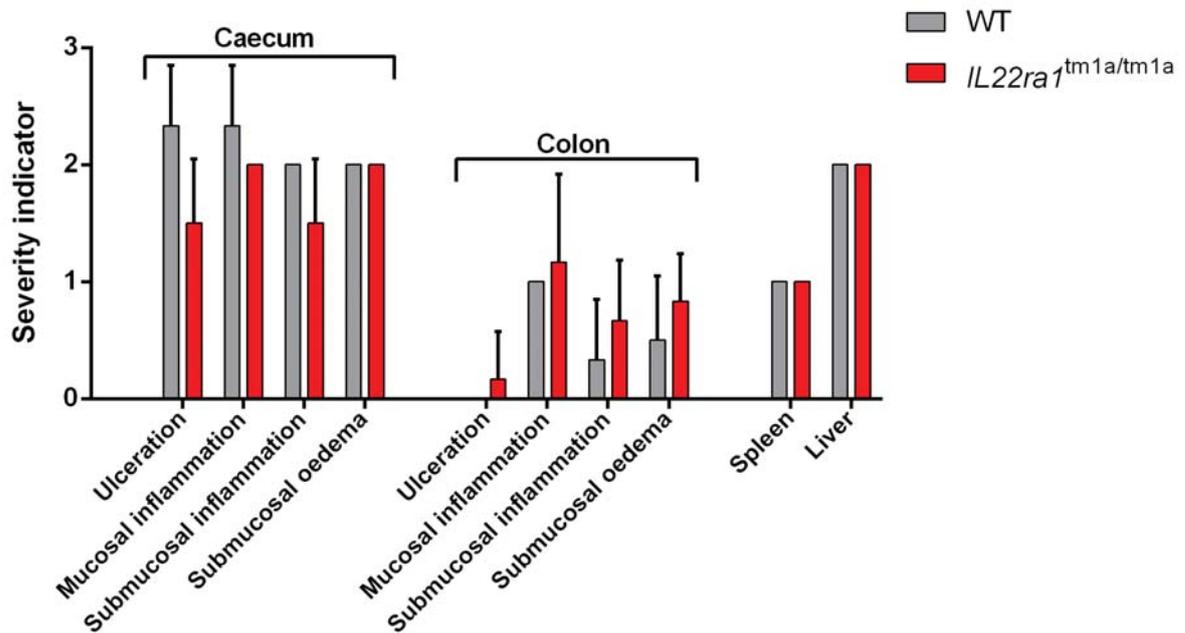


Figure 6.6. Burden of *Salmonella* in organs, weight loss, and intestinal inflammation in *IL22ra1^{tm1a/tm1a}* mice at day 4 PI with *S. Typhimurium* following streptomycin treatment. (A) *Salmonella* organ counts data from infection of wild type (WT) and *IL22ra1^{tm1a/tm1a}* mice with *S. Typhimurium* SL1344. Data is pooled from three individual experiments with oral delivery of a bacterial suspension of 1.2×10^4 , 1.4×10^4 and 1.2×10^4 CFU. Bars show median and interquartile range. Significance was assessed with a Mann Whitney U test. (B) Weight curve for wild type (WT) and *IL22ra1^{tm1a/tm1a}* mice infected with *S. Typhimurium*. Points indicate the mean of the mouse weights and error bars the standard error of the mean. Wild type: n = 7, *IL22ra1^{tm1a/tm1a}*: n = 8. (C) H & E-stained caecum, colon, liver and spleen from wild type (WT) and *IL22ra1^{tm1a/tm1a}* mice infected with *S. Typhimurium* were examined for indicators of inflammation. Caecum and colon were assigned scores to reflect the severity of ulceration, mucosal and submucosal inflammation and submucosal oedema. Liver and spleen were assigned a single score to reflect the degree of inflammation (0, 1, 2, or 3 for absent, mild, moderate & severe), n = 6.

6.3.4 The transcriptome of *IL22ra1^{tm1a/tm1a}* and *IL10rb^{tm1a/tm1a}* mice in *S. Typhimurium* infection

Analyses described in the previous two sections addressed potential phenotypes of *S. Typhimurium* infection in the *IL22ra1^{tm1a/tm1a}* and *IL10rb^{tm1a/tm1a}* mutant mouse lines. To examine possible differences in molecular signatures of infection between mutant and wild type mice RNA extracted from *S. Typhimurium*-infected caecal tissue was analysed by

RNAseq. *IL22ra1*^{tm1a/tm1a}, *IL10rb*^{tm1a/tm1a} and wild type mice were pre-treated with streptomycin and infected with *S. Typhimurium* SL1344 (n = 5). Samples of caecal tissue were collected at day 4 PI for RNA extraction followed by sequencing. RNA extracted from naïve *IL22ra1*^{tm1a/tm1a} and wild type mice (oral treatment with streptomycin followed by PBS only) were also sequenced in a separate sequencing run (n = 4). Differential expression analysis was performed to detect abnormalities in the response to infection in the mutant lines at the level of the transcriptome, and differences in basal gene expression in naïve *IL22ra1*^{tm1a/tm1a} mice. Fold change and p-value thresholds used to define differentially expressed genes were as previously used in Chapter 3 (log2 fold change < -1 or > 1, adjusted p-value < 0.05).

In samples from *S. Typhimurium*-infected *IL10rb*^{tm1a/tm1a} mice 83 genes were significantly increased and 77 significantly decreased compared with samples from wild type mice. A larger number of differentially expressed genes were detected for *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a} mice; 219 genes were upregulated and 190 downregulated relative to wild type controls. In contrast a relatively small number of differentially expressed genes were detected in naïve *IL22ra1*^{tm1a/tm1a} caeca compared with wild type controls; 15 genes were upregulated and 20 downregulated. Surprisingly little overlap was observed between differentially expressed genes for naïve and infected *IL22ra1* mutant mice; as few as 9 genes significantly altered in naïve *IL22ra1* mutants were also altered in the *S. Typhimurium*-infected animals relative to wild type controls with equal treatment. The 50 most highly increased and decreased genes for each comparison (*S. Typhimurium*-infected *IL10rb*^{tm1a/tm1a} verses wild type, *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a} verses wild type, and naïve *IL22ra1*^{tm1a/tm1a} versus wild type) are listed with corresponding transcript abundances, fold changes and associated p-values in Appendix 6.

Genes differentially expressed in *IL22ra1*^{tm1a/tm1a} mice were compared with a list of 41 genes reported to be regulated by IL22 signalling in previously published studies [152, 329, 348-353]. Eight genes reported to be upregulated by IL22 were found to be significantly decreased in *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a} mouse caecum compared with wild type mice. These were genes encoding the antimicrobial peptides Reg3 β , Reg3 γ and lipocalin-2, the antimicrobial chemokine CXCL1, the cytokine IL10, the fucosyltransferase Fut2, a phospholipase of the antibacterial phospholipase A2 family (Pla2g5), and Socs3 - part of a negative feedback system to limit cytokine signalling through the JAK/STAT pathway

[354]. In naïve *IL22ra1*^{tm1a/tm1a} caecum only one of the 40 genes, *Reg3β* was differentially expressed. Further many members of protein families implicated as targets of IL22 signalling were downregulated in the *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a} samples; for example expression of three matrix metallopeptidases, three members of the serum amyloid A family and six serine or cysteine peptidase inhibitors (SERPINs) were decreased in *IL22ra1*^{tm1a/tm1a} caecum relative to wild type control samples. Interestingly, while mucins are known to be upregulated by IL22, three mucins (2, 4, and 20) were upregulated in *IL22ra1*^{tm1a/tm1a} infected caecum relative to wild type controls, perhaps suggesting compensatory mechanisms may support mucin gene expression in the absence of IL22, or the cytokine may play different roles in regulation of the many members of this protein family [350]. The presence of known targets of IL22 signalling amongst the transcripts downregulated in *IL22ra1*^{tm1a/tm1a} mice validates the ability of this approach to detect targets of IL22 in infection.

Genes differentially regulated in *IL10rb*^{tm1a/tm1a} mice were also compared with published signalling targets; initially a list of over 40 genes previously reported to be regulated by the IL10 signalling pathway, many of which were supported by evidence from several studies [355-358]. Unexpectedly there was no overlap between the differentially expressed genes in *IL10rb*^{tm1a/tm1a} mice and the published genes. As IL10rb is an essential component of the IL22 receptor complex, possible overlap with the 41 previously reported IL22-regulated genes was also examined. Four genes induced by IL22 were found to be downregulated in the caeca of *IL10rb*^{tm1a/tm1a} mice; *Reg3β*, *Reg3γ*, *Pla2g5* and *Fut2*.

6.3.4.1 Signatures of T cell activation in the caecum of *IL10rb* mutant mice during *S. Typhimurium* infection

Despite the lack of changes in known IL10-regulated genes in the caecum of *Salmonella*-infected *IL10rb* mutants we observed a large number of genes important in T cell activation to be upregulated. T lymphocytes are one of the many cell types regulated by IL10; in particular IL10 is an important inhibitor of cytokine production by the Th1 subset and is essential for regulatory T cell function [359, 360]. We observed many proteins of the TCR/CD3 complex; the surface receptor complex for T cell activation by antigen and coupling to cellular signal transduction; were upregulated. Specifically eight proteins of the TCR including constant, variable and joining regions, and the δ , ϵ , and γ chains of CD3 were all significantly upregulated in *S. Typhimurium*-infected *IL10rb*^{tm1a/tm1a} caecum relative to wild type controls. Many other genes involved in T cell signalling pathways were also

increased, including *Zap70*, Linker for activation of T cells (*Lat*), Inducible T cell co-stimulator (*Icos*), and *IL2ra*. *Zap70* is important for signal transduction from the TCR and *IL2ra* encodes a chain of the receptor for IL2, a cytokine which acts primarily on T cells and promotes differentiation into effector and memory cell types. Upregulation of *CD40LG*, a T cell surface protein which regulates B cell function, and the B lymphocyte kinase *Blk*, also indicates possible effects on B cells in the *IL10rb* mutants.

6.3.4.2 Common genes are differentially expressed in *S. Typhimurium*-infected caecum from *IL10rb* and *IL22ra1* mutant mice

Given the involvement of *IL10rb* in the signalling receptor complex for IL22 we anticipated an overlap between genes differentially expressed in *IL22ra1* and *IL10rb* mutants in comparison with wild type mice. Indeed approximately half of transcripts significantly downregulated in caecal tissue from *IL10rb*^{tm1a/tm1a} were also significantly downregulated in the caeca of *IL22ra1*^{tm1a/tm1a} mice; the common genes accounting for one fifth of genes downregulated in the *IL22ra1* mutants. 28% of transcripts upregulated in *IL10rb* mutant tissue were similarly regulated in *IL22ra1* mutants, the common genes making up 11% of transcripts upregulated in the *IL22ra1* mutants. The numbers of genes which display common regulation are reported in Figure 6.7.

The greater overlap in downregulated genes in the two mutants is accompanied by a tighter correlation between fold changes for downregulated compared with those for upregulated genes; Pearson's *r* for the fold changes of the common downregulated genes was 0.54 compared with 0.21 for the upregulated genes. Notably transcripts encoding the serine protease *Prss27*, G-protein *Gna14* and antimicrobial peptide *Reg3γ*, were in the top five most downregulated genes for both mutants, with the absolute reduction in transcript levels ranging from a factor of 7.8 to 5.1.

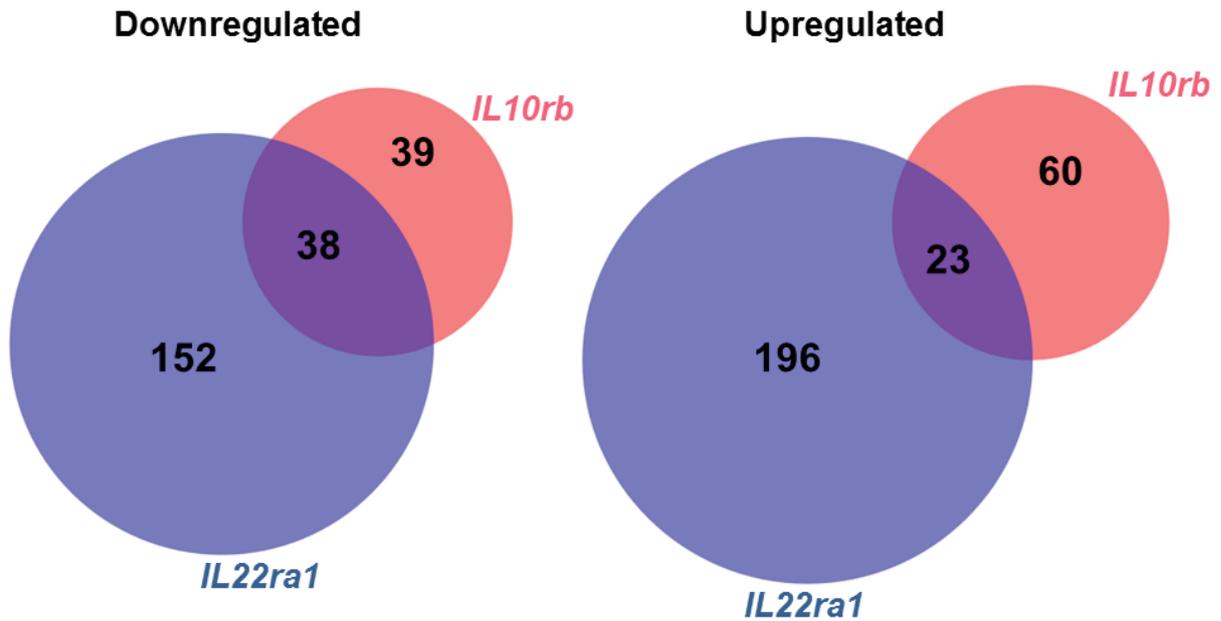


Figure 6.7. Overlap between genes differentially expressed in caecal tissue from *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a} and *IL10rb*^{tm1a/tm1a} mice compared with wild type controls. RNA was extracted from caecal tissue of *S. Typhimurium*-infected *IL22ra1*^{tm1a/tm1a}, *IL10rb*^{tm1a/tm1a} and wild type mice at day 4 PI and analysed by RNAseq. DESeq2 was used to perform differential expression analysis to identify genes expressed at different levels in mutant and wild type mice. Transcripts with a log₂ fold change < -1 or > 1 and adjusted p-value < 0.05 are considered differentially expressed. Collections of genes with different expression in the *IL22ra1* and *IL10rb* mutant lines are indicated here by circles, with the circle size proportional to the number of genes it represents.

6.3.5 *S. Typhimurium* infection of *BC017643*^{tm1a/tm1a} mice in the streptomycin model

As mentioned in section 6.1.1, *BC017643*^{tm1a/tm1a} mice display dramatically increased susceptibility to *S. Typhimurium* following intraperitoneal delivery of the *S. Typhimurium* strain M525. Investigation of the protein encoded by *BC017643* has indicated an important role in generation of reactive oxygen species through positive regulation of levels of endoplasmic reticulum NADPH oxidase complex components (D. Thomas and S. Clare, unpublished). In this work we sought to investigate how deficiency in *BC017643* affects mice orally infected with *S. Typhimurium* in the streptomycin model of gastroenteritis.

24 h post-treatment with streptomycin *BC017643*^{tm1a/tm1a} and wild type control mice were infected with *S. Typhimurium* SL1344 (n = 7). At day 3 PI *BC017643*^{tm1a/tm1a} mice displayed overt signs of illness and consequently all mice were culled. Organ plating was performed for the liver, spleen, colon and caecum for elucidation of *Salmonella* counts.

Relative to wild type controls *BC017643*^{tm1a/tm1a} mice displayed highly increased *S. Typhimurium* CFU in the liver and spleen (both $p = 0.0006$), and significantly increased counts in the caecum ($p = 0.011$) (Figure 6.8). These findings indicate that the defect in *BC017643* has an effect both in intestinal tissue, and in the liver and spleen following systemic dissemination of *S. Typhimurium*.

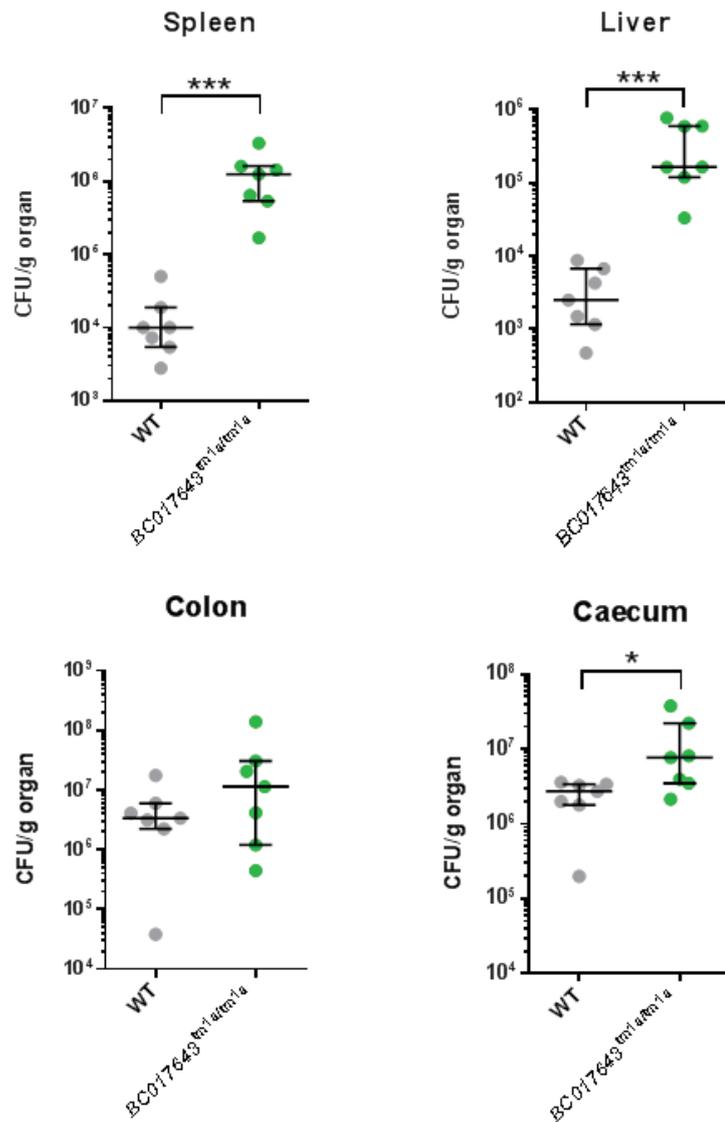


Figure 6.8. Burden of *Salmonella* in organs of *BC017643*^{tm1a/tm1a} and wild type mice at day 3 PI. *Salmonella* organ counts from mice infected with *S. Typhimurium* SL1344 (9×10^3 CFU) following streptomycin treatment. Bars show median and interquartile range. Statistical significance was assessed with a Mann Whitney U test.

6.4 Discussion

IL10rb^{tm1a/tm1a} mice failed to develop spontaneous colitis as observed in *IL10rb* and *IL10* mutant lines by others previously, despite the dramatic reduction in *IL10rb* transcripts in macrophages and *S. Typhimurium*-infected caecal tissue from *IL10rb*^{tm1a/tm1a} mice relative to wild type controls [334, 335]. Several non-mutually exclusive possibilities might contribute to this difference. For example, despite the reduced level of *IL10rb* expression in mutant mice the leaky tm1a allele may give rise to sufficient IL10rb activity to avert development of obvious intestinal inflammation. As demonstrated in Chapter 4 correlation between transcript and protein abundance is modest, and therefore the dramatic reduction in transcript levels may have a smaller effect on the level of IL10rb protein. In addition, although a large reduction in *IL10rb* expression relative to wild type controls was demonstrated for isolated macrophages and *Salmonella*-infected caecum from *IL10rb*^{tm1a/tm1a} mice, expression in naïve caecum or colon was not investigated here.

Further, both the IL10rb- and IL10-deficient lines described previously were in a different mouse background strain to the *IL10rb*^{tm1a/tm1a} mice described in this thesis [334, 335]. Background strain can cause wide variations in the effect of inactivating mutations in mice, and therefore differences between the strains might be a factor in the different phenotypes observed here [361]. Finally, differences in the conditions under which mice were housed may also have a role. IL10-deficient mice maintained under SPF conditions were reported to show reduced levels of colitis compared to conventionally-housed mice [335]. A study in which IL10-deficient mice were shown to be protected from spontaneous colitis following *C. rodentium* infection also demonstrates the importance of environmental exposure to different bacteria in the development of colitis [338].

Previous studies addressing the effects of IL10 signalling upon outcomes in *Salmonella* infection support a detrimental impact of IL10 upon host defences [337, 341]. However, no significant difference in intestinal colonisation or systemic dissemination by *S. Typhimurium* was observed between wild type and *IL10rb*^{tm1a/tm1a} mice. As relatively small numbers of mice were used in these experiments and bacterial counts displayed substantial variability this limited our ability to detect a phenotype. However similar to the absence of spontaneous colitis in *IL10rb*^{tm1a/tm1a} mice it is again possible that leaky expression of *IL10rb* may have maintained sufficient signalling receptor activity to prevent detection of a phenotype. Alternatively, IL10-induced suppression of inflammatory and antimicrobial

pathways may not have a substantial effect on the course of this infection. Also important to consider is the role of IL10rb in the receptor complex activated by other IL10 family cytokines. While reduced IL10 signalling in isolation may be protective in *Salmonella* infection, the combined reduction in signalling by the entire array of IL10 cytokine family members may lead to more complex effects on the host response.

In contrast to *IL10rb*^{tm1a/tm1a} mice for which qPCR and RNAseq supported a substantial reduction in target gene expression but no phenotype was observed in *S. Typhimurium* infection, in *IL22ra1*^{tm1a/tm1a} mice the reverse was observed. Analysis of *IL22ra1* transcript abundances in naïve and *Salmonella*-infected caecal tissue suggested that whilst *IL22ra1* expression is significantly reduced in the caecum of naïve *IL22ra1*^{tm1a/tm1a} mice relative to wild type controls, at day 4 PI mutant expression is comparable to wild type levels. Although *IL22ra1* transcript levels in naïve caecal tissue were just 3-fold lower than wild type levels this seemingly modest reduction must be sufficient to give rise to the phenotypic and molecular differences observed here in *S. Typhimurium* infection, and the susceptibility phenotype of *IL22ra1*^{tm1a/tm1a} mice in infection with *C. rodentium* [329]. While the RNAseq evidence suggests that *IL22ra1* expression in the caecum is restored to wild type levels upon infection a delay in restoration during the earlier stages of infection may give rise to the observed differences. An alternative possibility is that whilst splicing in the tm1a allele successfully excludes the inserted reporter cassette, the full-length *IL22ra1* transcript produced in *IL22ra1*^{tm1a/tm1a} mice contains a mutation, which either partially or completely inactivates the function of the encoded protein. To investigate this possibility an end-point reverse transcription (RT) PCR should be performed to amplify the whole transcript, followed by PCR and sequencing.

The detection of reduced systemic dissemination in *IL22ra1*^{tm1a/tm1a} mice, despite the absence of significant differences in intestinal pathology and colonisation, is an interesting finding. As IL22 is reported to be important for tissue repair and epithelial barrier maintenance, mice deficient in the cytokine receptor might be expected to display increased systemic *Salmonella* dissemination. However a study of *IL22*^{-/-} mice in the streptomycin model published last year demonstrated a significant reduction in *Salmonella* colonisation of the colon. The reduction in intestinal *Salmonella* was attributed to the failure of antimicrobial peptide induction in these mice, to the benefit of commensal bacteria susceptible to these molecules [152]. Although we did not detect a significant difference in *Salmonella*

colonisation of the intestine we observed reduced CFU in the liver and spleen. One possibility for the difference between our findings and the *IL22*^{-/-} study is that accurate enumeration of tissue-associated bacterial counts is more difficult in the intestinal tissues than in the liver and spleen due to wide-ranging levels of contamination with adherent faecal material. Therefore reduced *Salmonella* in the gut may become apparent only as reduced systemic dissemination of *Salmonella* to the liver and spleen. Indeed interquartile ranges in CFU were observed to be much smaller in liver and spleen than the intestinal organs. To further investigate the possibility of increased colonisation resistance in *IL22ra1*^{tm1a/tm1a} mice it would be interesting to perform 16S rRNA gene sequencing of intestinal content during *S. Typhimurium* infection similar to the *IL22*^{-/-} study [152].

The possibility of IL22 effects on systemic control of bacteria should also be considered. However the finding that *IL22ra1*^{tm1a/tm1a} mice display reduced *Salmonella* in the liver and spleen lies at odds with the finding that IL22 induces hepatic complement protein production resulting in enhanced opsonization of systemic bacteria [346].

Far fewer differences between the transcriptome profiles of *IL22ra1*^{tm1a/tm1a} and wild type caecum were observed in naïve mice compared with mice infected with *S. Typhimurium*. Although fewer naïve mice were sequenced this is unlikely to account for the dramatic difference observed. The greater differences in the transcriptomes of infected mice were observed despite a significant difference in *IL22ra1* expression exclusively in the naïve condition. A likely explanation is that in the naïve state IL22ra1 signalling activity is limited in comparison to the infected state. Therefore during infection defective IL22 signalling is less well tolerated, resulting in larger effects on genes downstream of this cytokine. It was surprising to find that just one quarter of DE genes in naïve *IL22ra1*^{tm1a/tm1a} mice were also DE during *S. Typhimurium* infection. This suggests that the combinatorial effect of activation of other pathways alongside IL22 signalling during infection results in effects on genes more broadly and different to those in the naïve condition where less immune signalling pathways are active.

Based on the broader role of IL10rb compared with IL22ra1, IL22ra1-deficient mice might be predicted to display a subset of the gene expression differences observed in IL10rb-deficient mice during infection. However we observed a greater number of DE genes in the *IL22ra1*^{tm1a/tm1a} mutant mice. A complex interplay between cytokine signalling pathways during infection might account for this result. For further investigation of the individual

effects of IL10 and IL22 signalling during gastrointestinal infection with *S. Typhimurium* it would be interesting to repeat infections with knockout mice for the cytokines themselves.

The failure to identify differential expression of specific genes reported to be regulated by IL10 in *S. Typhimurium*-infected *IL10rb*^{tm1a/tm1a} mice was potentially concerning, although again the difference may be a consequence of the broader role of IL10rb in signalling by other cytokines. However finding downregulation of several genes activated by IL22 in these mice was a reassurance that the DE genes we detected included those resulting specifically from IL10rb deficiency. The finding of upregulation of many genes relating to activation of T cells through the TCR complex in *S. Typhimurium*-infected *IL10rb*^{tm1a/tm1a} mice was interesting and is worthy of more detailed investigation in the context of current knowledge on the effects of IL10 on T cells. The first step in this investigation would be to profile the transcriptome of naïve *IL10rb*^{tm1a/tm1a} mouse caecum in order to determine whether this effect is present in the absence of *Salmonella*.

In summary the transcriptome analysis of mutant mice successfully identified genes known to be involved in the targeted pathways, demonstrating the approach can effectively identify genes regulated in infection by these targeted genes. Also detected were many genes with no obvious link to IL10 and IL22 pathways, and many predicted genes and genes of unknown function. These results implicate these genes as potential targets of the IL10 and IL22 pathways, although these may be indirect, and provide a starting point for further investigation of these links. Further work could involve more complex analysis to consider the individual and combined effects of *Salmonella*-infection and host genotype upon gene expression.

Detection of an infection susceptibility phenotype in *BC017643*^{tm1a/tm1a} mice demonstrates the utility of screening approach for gaining insight into genes of unknown function. Prior work identified an important role for the product of gene *BC017643* in the control of systemic *S. Typhimurium* infection. The results presented in this chapter confirm this finding and suggest the gene is also important for control of *S. Typhimurium* in caecal tissue. *BC017643* has been shown to positively regulate proteins involved in the phagocyte NADPH oxidase complex, with the highest levels of expression in human blood cell subsets displayed by neutrophils and monocytes. As neutrophil-mediated killing of *Salmonella* in the intestinal mucosa is one of the major arms of defence at this site the increased *Salmonella* CFU in intestinal tissue of *BC017643*^{tm1a/tm1a} is in keeping with the proposed role for the

BC017643 protein. The observation of more dramatic differences in CFU between *BC017643*^{tm1a/tm1a} and wild type mice in the liver and spleen is interesting. On the one hand this might indicate that the activity of BC017643 is of greater importance in the control of disseminated bacteria. Alternatively this might reflect a complex relationship between control of *Salmonella* colonisation of the gut and systemic dissemination.

7 Final discussion

The aim of the work presented in this thesis was to produce a detailed description of the streptomycin mouse model of *S. Typhimurium* gastroenteritis, and investigate the potential of this model for the identification of host genes involved in infection susceptibility. Characterisation of the model involved classical phenotyping approaches, such as enumeration of bacterial counts and flow cytometric analysis of immune cell populations, in addition to ‘omics’-based approaches to profile changes in the microbiota, and regulation of the caecal tissue transcriptome and proteome during *S. Typhimurium* infection.

7.1 New opportunities presented by high throughput technologies

The arrival of high throughput ‘omics’ technologies for biological research has provided new approaches to understanding large scale biological effects, such as the response of a host to infection, at the levels of molecules and pathways. High throughput analyses provide an unbiased approach to investigating a system; rather than using a pre-selected focus, the measurable repertoire of molecules present in the materials is interrogated. Signatures of known processes serve to validate the quality of the data, and novel processes for which a role in the system under investigation was not previously appreciated can be identified.

Whilst the advantages of high throughput research are plentiful there are also problems demanding attention. The challenge of extracting biological meaning from large volumes of data has been tackled by the development of analysis tools for identifying pathways and processes enriched in molecular profiles. However the coverage of pathway annotations, and accuracy with which models reflect true biology, limits the insight these tools generate. In addition, established tools for the integration of multiple data types are not widely available, yet integration will be central to the future understanding whole systems, and the multiple levels of regulation within.

An area of rapidly-paced progress, proteomics has advanced far beyond the identification of individual proteins to quantitative profiling of thousands of proteins present in complex mixtures. Although the sensitivity limit of MS precludes the quantification of low abundance proteins the captured abundant fraction can be used to validate pathways identified as regulated in more comprehensive transcriptomic profiles. In this work we selected the overlap of regulated genes at the transcript and protein level as ‘functionally validated

regulated transcripts' and used these to identify pathways regulated in *S. Typhimurium*-infected caecal tissue at the stage where *Salmonella* numbers are high and inflammation extensive. An alternative approach might use changes in the transcriptome for identification of pathways which are regulated and accept that a set degree of support at the protein level (for example a minimum of two proteins within the pathway regulated in the direction supported by the transcriptome) confirms the regulation of these pathways; thereby reducing the bias toward greater significance for pathways which contain more abundant proteins.

Unfortunately work to profile changes in the metabolome of caecal tissue and serum in response to *S. Typhimurium* infection was not completed soon enough for inclusion in this report. Metabolomics data will provide a further level of information on the effects of *S. Typhimurium* infection on the host, complementary to the transcriptome and proteome. A study of the host metabolic response to *Salmonella* infection in the murine typhoid fever model reported infection had 'a profound impact' on host metabolism [228]. Suggesting that dramatic changes in metabolism are also a feature of gastrointestinal infection, a large proportion of transcripts and proteins which were downregulated in infection of the caecum in our work are annotated to metabolic pathways. The metabolomics data will allow us to observe how changes in the regulation of metabolic genes are realised in the metabolic state of the caecum during *S. Typhimurium*-induced inflammation. This data is currently being collected ready for analysis through collaboration.

The transcriptomic and proteomic analyses reported in this thesis were performed upon whole caecal tissue, the samples analysed consisting of a large number of different sub-regions and cell types. By averaging all signals across the whole tissue information is lost; abundance changes in rare cell types are diluted to insignificance and where a gene is up- or down-regulated in different cell types these effects are combined in one overall fold change. Applying a multi-'omic' approach on a more tissue or cell-focussed level would allow us to dissect the overall changes reported here in order to uncover their contributions. Approaches might include dissociation of the epithelium from the mouse caecum, or potentially infection of mouse intestinal organoids; three-dimensional crypt structures which contain a well-defined collection of cell types [362].

7.2 A role for complement in gastrointestinal *S. Typhimurium* infection

In Chapter 5 we outlined results which implicate the complement protein cascade in the host response to mucosal *S. Typhimurium* infection. These results suggest that during *S. Typhimurium* infection complement production and activation might occur locally in intestinal tissue. However further work is needed to dissect the contributions of hepatic and locally-produced complement in the caecum. Previous studies have reported that complement activity is important for protection against *C. rodentium* though these reports made no attempt to describe the source of protective complement proteins [309, 310]. It would therefore be interesting to compare both the extent of regulation of complement proteins and their sources in infections with luminal-resident bacterial pathogens such as *C. rodentium*, and invasive intestinal bacteria including *Salmonella*.

Future experiments should also characterise the roles of complement in the intestinal mucosa during *S. Typhimurium* infection. Mice deficient in the central complement protein C3 could be used to investigate both the sources of C3 and functional importance of complement activation. Comparison of the relative levels of C3 in the blood and the intestinal mucosa of wild type and $C3^{-/-}$ mice with systemic C3 replacement by serum transfer during *S. Typhimurium* infection will provide insight into the source of C3 in the mucosa. The functional role of systemic and locally-derived complement could be examined by testing the susceptibility of $C3^{-/-}$ mice to infection in the streptomycin mouse model, and determining whether transfer of serum from wild type mice to $C3^{-/-}$ mice has an effect on infection outcome. Oral delivery of a targeted inhibitor of complement activation has been shown to ameliorate intestinal injury during DSS-induced colitis, presenting an interesting potential approach to investigate the effects of intestinal complement activation in our model [363].

The processes of local complement activation in the intestinal mucosa and the infiltration of products from systemic complement activation may have overlapping roles and outcomes in infection. It would be interesting to determine the specific effector pathways of the complement cascade to which these processes contribute. Previous work has suggested that MAC-mediated killing of *Salmonella* is of minor functional relevance compared with opsonisation of *Salmonella* by complement activation fragments in mouse serum [287]. It would therefore be of interest to investigate whether this is also true of complement-mediated effects on *Salmonella* within the intestinal mucosa.

In this work we detected a dramatic increase in the complement activation product C3d in the plasma of mice at day 4 PI with *S. Typhimurium* in the streptomycin mouse model. At this time point numbers of systemic *Salmonella* are considerable, and it is therefore possible that much of the C3d detected was generated outside of the gut. However, it would be interesting to determine whether complement activation fragments can be detected in the blood during self-limiting gastrointestinal infection with *S. Typhimurium* in humans, and whether these fragments might potentially serve to differentiate gastroenteritis caused by *Salmonella* from diarrheal disease resulting from other intestinal pathogens. We are planning to perform such experiments through collaboration.

7.3 Insight into genes involved in the host defence to infection from the streptomycin mouse model

The work reported in Chapter 6 demonstrates that characterisation of responses of mutant mice to infection in the streptomycin mouse model can provide valuable insight into the involvement of host genes in defence against *Salmonella*. For example in the case of *BC017643* our data suggested that defects in this gene affect host susceptibility in both the mucosal and systemic stages of *Salmonella* infection. *IL22ra1* mutant mice displayed a contrasting susceptibility profile with our data suggesting a reduction in the cytokine receptor component this gene encodes has little influence on mucosal infection, yet a protective effect upon systemic infection. The outcome of infection in *BC017643* mutant mice is consistent with the recently discovered role of the encoded protein in the generation of phagocyte reactive oxygen species, and the major role of phagocytes, in particular neutrophils recruited to the intestine, in *Salmonella* killing. The observations in *IL22ra1* mutants however call for deeper investigation of the mechanisms involved.

Amongst the findings described in Chapter 6 the apparent detection of ‘leaky’ gene expression from the targeted gene in two *tm1a* mutant mouse lines tested suggests caution may be required in interpretation of results from experiments using *tm1a* mutants. Regrettably we failed to determine the difference between mutant and wild type mice at the level of the target protein, the functional molecule. Ideally immunoprecipitation followed by Western blotting would be used in order to determine whether mutant mice produce the protein of interest, and if applicable whether the amino acid sequence of the target protein is identical in mutant and wild type mice.

Though likely resulting in an underestimation of the number of ‘hits’, incomplete knock down of target gene expression is not hugely problematic for primary phenotyping, where the aim is simply to establish a link between a gene and a particular process. However variable levels of target gene expression in hypomorphs prevent solid conclusions on the involvement of the target gene in particular pathways being drawn, representing a major barrier to secondary phenotyping efforts. For genes of particular interest, or where a phenotype was observed in the primary pipeline, tm1b lines should automatically be generated to facilitate secondary phenotyping.

RNAseq analysis of *Salmonella*-infected caecal tissue from *IL22ra1* and *IL10rb* mutant mice identified collections of genes whose expression was affected, either directly or indirectly, in response to reduced expression of just these single target genes. The sizeable overlap in genes with altered expression in these two mutants is in line with the known relationship of *IL22ra1* and *IL10rb* in IL22 cytokine signalling. The observation of a T cell signature in genes more highly expressed in *IL10rb* mutants relative to wild type controls is broadly consistent with the immunomodulatory effects of IL10, and further study of T cell activation in the caecum during *Salmonella* infection and the effects of *IL10rb* and *IL10* mutations would be interesting.

Whilst the value of the streptomycin model in uncovering aspects of the host response to infection is evident, there are a number of features of the model which require more extensive consideration and testing in order to maximise the information which might be gained in these experiments, as follows.

In the murine typhoid model differences in *Salmonella* counts recovered from the liver and spleen of wild type and mutant animals are thought to be determined predominantly by the involvement of the gene of interest in systemic host defences. However in the streptomycin model it seems likely that systemic *Salmonella* counts also depend upon the control of bacteria in intestinal tissue, especially the passage of *Salmonella* from the intestinal mucosa both into the circulation and lymphatic system. Therefore whilst differences in intestinal counts between wild type mice and a mutant line inarguably indicate a defect in intestinal defence mechanisms, differences in counts in the liver and spleen need further investigation. The complementary murine typhoid model might be used to investigate these mechanisms. In the case of *IL22ra1* mutant mice no difference in liver and spleen *Salmonella* CFU following IP delivery of *S. Typhimurium* in the MGP phenotyping pipeline was detected

in comparison with wild type controls. Though the number of mice tested in the MGP pipeline was small this result might indicate that the systemic protection observed in the streptomycin model has a founding at the level of intestinal control.

Of the three mutant lines examined in this work none displayed a strong phenotype with respect to intestinal bacterial counts; in *BC017643* mutant mice a difference in intestinal counts relative to wild type mice was observed, but compared with the liver and spleen the effect here was much smaller. The failure to detect a strong phenotype in intestinal counts in this work despite reproducible phenotypes in *Salmonella* numbers in the liver and spleen is likely due to the small number of lines tested. Nonetheless it would be valuable to confirm that the challenge as it was carried out in this work finds a strong difference in intestinal counts in a mutant mouse with a known phenotype in susceptibility to intestinal colonisation. As alluded to above, and potentially supported by the findings in *IL22ra1* mutant mice, the possibility exists that despite the absence of *Salmonella* colonisation of liver and spleen in the majority of human NTS infections, the extent of colonisation of these sites may act as better indicators of defects in the intestinal immune response to *Salmonella* in the streptomycin mouse model. The large range in intestinal *Salmonella* CFU reported here may have been a barrier to the detection of a phenotype in intestinal colonisation. Further experimentation should be carried out to determine whether this range truly reflects naturally occurring variability, or whether variability might be a consequence of faecal contamination or methodological problems such as insufficient homogenisation of intestinal tissue.

The infection time point selected as the major focus of analysis in this work was day four, at which stage intestinal tissues are extensively colonised by *Salmonella*, intestinal inflammation is severe, and considerable systemic spread of *Salmonella* has occurred. Many published studies which investigated effects of host and bacterial genotype on the outcomes of infection in the streptomycin mouse model focused their attention or at least included analysis of earlier time points [130, 152]. Whilst the results described in this thesis demonstrate it is possible to detect phenotypic differences in infections with both bacterial and host mutants (Chapters 3 and 6 respectively) at day 4 PI, it is possible that differences may be more significant at earlier time points. In addition the use of earlier time points where systemic infection is limited might be useful in distinguishing intestinal and systemic processes, as discussed in the case of complement activation. A more detailed characterisation of earlier time points is warranted to investigate variability in infection parameters throughout the

course of infection, and how the interval between infection and experimental characterisation impacts upon effects of host and bacterial genotype.

8 References

1. Jones, K.E., et al., *Global trends in emerging infectious diseases*. Nature, 2008. **451**(7181): p. 990-993.
2. Bhutta, Z.A., et al., *Global burden, distribution, and interventions for infectious diseases of poverty*. Infectious Diseases of Poverty, 2014. **3**: p. 21-21.
3. Patz, J.A., et al., *Effects of environmental change on emerging parasitic diseases*. International Journal for Parasitology, 2000. **30**(12–13): p. 1395-1405.
4. Starr, J., *Hospital acquired infection*. BMJ : British Medical Journal, 2007. **334**(7596): p. 708-708.
5. Griffith, D.C., L.A. Kelly-Hope, and M.A. Miller, *Review of Reported Cholera Outbreaks Worldwide, 1995–2005*. The American Journal of Tropical Medicine and Hygiene, 2006. **75**(5): p. 973-977.
6. Crump, J.A., S.P. Luby, and E.D. Mintz, *The global burden of typhoid fever*. Bulletin of the World Health Organization, 2004. **82**: p. 346-353.
7. Cliff, A. and P. Haggett, *Time, travel and infection*. British Medical Bulletin, 2004. **69**(1): p. 87-99.
8. Hufnagel, L., D. Brockmann, and T. Geisel, *Forecast and control of epidemics in a globalized world*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(42): p. 15124-15129.
9. Levy, S.B. and B. Marshall, *Antibacterial resistance worldwide: causes, challenges and responses*. Nature Medicine, 2004.
10. Tamboli, C.P., et al., *Dysbiosis in inflammatory bowel disease*. Gut, 2004. **53**(1): p. 1-4.
11. Stecher, B., L. Maier, and W.-D. Hardt, *'Blooming' in the gut: how dysbiosis might contribute to pathogen evolution*. Nature Reviews Microbiology, 2013. **11**(4): p. 277-284.
12. Levine, M.M., et al., *The Global Enteric Multicenter Study (GEMS): Impetus, Rationale, and Genesis*. Clinical Infectious Diseases, 2012. **55**(suppl 4): p. S215-S224.
13. Lanata, C.F., et al., *Global Causes of Diarrheal Disease Mortality in Children <5 Years of Age: A Systematic Review*. PLoS ONE, 2013. **8**(9): p. e72788.
14. Walker, C.L.F., et al., *Global burden of childhood pneumonia and diarrhoea*. The Lancet. **381**(9875): p. 1405-1416.
15. Guerrant, R.L., et al., *Malnutrition as an enteric infectious disease with long-term effects on child development*. Nutrition Reviews, 2008. **66**(9): p. 487-505.
16. Liu, L., et al., *Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000*. The Lancet. **379**(9832): p. 2151-2161.
17. Walker, C.L.F., et al., *Scaling Up Diarrhea Prevention and Treatment Interventions: A Lives Saved Tool Analysis*. PLoS Medicine, 2011. **8**(3): p. e1000428.

18. Jiang, V., et al., *Performance of rotavirus vaccines in developed and developing countries*. Human Vaccines, 2010. **6**(7): p. 532-542.
19. Kotloff, K.L., et al., *Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study*. The Lancet. **382**(9888): p. 209-222.
20. Locking, M.E., et al., *Risk factors for sporadic cases of Escherichia coli O157 infection: the importance of contact with animal excreta*. Epidemiology and Infection, 2001. **127**(02): p. 215-220.
21. Rangel, J., et al., *Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982–2002* Emerging Infectious Diseases, 2004. **11**(4).
22. Kyne, L., et al., *Health Care Costs and Mortality Associated with Nosocomial Diarrhea Due to Clostridium difficile*. Clinical Infectious Diseases, 2002. **34**(3): p. 346-353.
23. MacCannell, T., et al., *Guideline for the Prevention and Control of Norovirus Gastroenteritis Outbreaks in Healthcare Settings*. Infection Control and Hospital Epidemiology, 2011. **32**(10): p. 939-969.
24. Kuijper, E.J., et al., *Emergence of Clostridium difficile-associated disease in North America and Europe*. Clinical Microbiology and Infection, 2006. **12**: p. 2-18.
25. Peterson, L.W. and D. Artis, *Intestinal epithelial cells: regulators of barrier function and immune homeostasis*. Nature Reviews Immunology, 2014. **14**(3): p. 141-153.
26. Leblond, C. and B.E. Walker, *Renewal of cell populations*. Physiological Reviews, 1956. **36**(2): p. 255-276.
27. Barker, N., *Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration*. Nature Reviews Molecular Cell Biology, 2014. **15**(1): p. 19-33.
28. Mowat, A.M. and W.W. Agace, *Regional specialization within the intestinal immune system*. Nature Reviews Immunology, 2014. **14**(10): p. 667-685.
29. Bevins, C.L. and N.H. Salzman, *Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis*. Nature Reviews Microbiology, 2011. **9**(5): p. 356-368.
30. Specian, R.D. and M.G. Oliver, *Functional biology of intestinal goblet cells*. American Journal of Physiology-Cell Physiology, 1991. **260**(2): p. C183-C193.
31. Gunawardene, A.R., B.M. Corfe, and C.A. Staton, *Classification and functions of enteroendocrine cells of the lower gastrointestinal tract*. International Journal of Experimental Pathology, 2011. **92**(4): p. 219-231.
32. Mellitzer, G. and G. Gradwohl, *Enteroendocrine cells and lipid absorption*. Current Opinion in Lipidology, 2011. **22**(3): p. 171-175.
33. Eberl, G. and M. Lochner, *The development of intestinal lymphoid tissues at the interface of self and microbiota*. Mucosal Immunology, 2009. **2**(6): p. 478-485.

34. Spencer, J., et al., *The development of gut associated lymphoid tissue in the terminal ileum of fetal human intestine*. *Clinical and Experimental Immunology*, 1986. **64**(3): p. 536-543.
35. Lelouard, H., et al., *Peyer's Patch Dendritic Cells Sample Antigens by Extending Dendrites Through M Cell-Specific Transcellular Pores*. *Gastroenterology*, 2012. **142**(3): p. 592-601.e3.
36. Mabbott, N.A., et al., *Microfold (M) cells: important immunosurveillance posts in the intestinal epithelium*. *Mucosal Immunology*, 2013. **6**(4): p. 666-677.
37. Neutra, M.R., N.J. Mantis, and J.-P. Kraehenbuhl, *Collaboration of epithelial cells with organized mucosal lymphoid tissues*. *Nature Immunology*, 2001. **2**(11): p. 1004-1009.
38. Baptista, A.P., et al., *Colonic patch and colonic SILT development are independent and differentially regulated events*. *Mucosal Immunology*, 2013. **6**(3): p. 511-521.
39. Cornes, J.S., *Number, size, and distribution of Peyer's patches in the human small intestine: Part I The development of Peyer's patches*. *Gut*, 1965. **6**(3): p. 225-229.
40. Pabst, O., et al., *Cryptopatches and isolated lymphoid follicles: dynamic lymphoid tissues dispensable for the generation of intraepithelial lymphocytes*. *European Journal of Immunology*, 2005. **35**(1): p. 98-107.
41. Lorenz, R.G. and R.D. Newberry, *Isolated Lymphoid Follicles Can Function as Sites for Induction of Mucosal Immune Responses*. *Annals of the New York Academy of Sciences*, 2004. **1029**(1): p. 44-57.
42. Bouskra, D., et al., *Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis*. *Nature*, 2008. **456**(7221): p. 507-510.
43. Ferguson, A., *Intraepithelial lymphocytes of the small intestine*. *Gut*, 1977. **18**(11): p. 921-937.
44. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing*. *Nature*, 2010. **464**(7285): p. 59-65.
45. Sommer, F. and F. Backhed, *The gut microbiota - masters of host development and physiology*. *Nature Reviews Microbiology*, 2013. **11**(4): p. 227-238.
46. O'Hara, A.M. and F. Shanahan, *The gut flora as a forgotten organ*. *EMBO Reports*, 2006. **7**(7): p. 688-693.
47. Arumugam, M., et al., *Enterotypes of the human gut microbiome*. *Nature*, 2011. **473**(7346): p. 174-180.
48. Clemente, Jose C., et al., *The Impact of the Gut Microbiota on Human Health: An Integrative View*. *Cell*, 2012. **148**(6): p. 1258-1270.
49. Stecher, B., et al., *Like Will to Like: Abundances of Closely Related Species Can Predict Susceptibility to Intestinal Colonization by Pathogenic and Commensal Bacteria*. *PLoS Pathogens*, 2010. **6**(1): p. e1000711.
50. Endt, K., et al., *The Microbiota Mediates Pathogen Clearance from the Gut Lumen after Non-Typhoidal Salmonella Diarrhea*. *PLoS Pathogens*, 2010. **6**(9): p. e1001097.

51. Panda, S., et al., *Short-Term Effect of Antibiotics on Human Gut Microbiota*. PLoS ONE, 2014. **9**(4): p. e95476.
52. Pérez-Cobas, A.E., et al., *Gut microbiota disturbance during antibiotic therapy: a multi-omic approach*. Gut, 2012.
53. Kato, H., et al., *Colonisation and transmission of Clostridium difficile in healthy individuals examined by PCR ribotyping and pulsed-field gel electrophoresis*. Journal of Medical Microbiology, 2001. **50**(8): p. 720-727.
54. Poutanen, S.M. and A.E. Simor, *Clostridium difficile-associated diarrhea in adults*. Canadian Medical Association Journal, 2004. **171**(1): p. 51-58.
55. Karadsheh, Z. and S. Sule, *Fecal Transplantation for the Treatment of Recurrent Clostridium Difficile Infection*. North American Journal of Medical Sciences, 2013. **5**(6): p. 339-343.
56. Wostmann, B.S., *The Germfree Animal in Nutritional Studies*. Annual Review of Nutrition, 1981. **1**(1): p. 257-279.
57. Abrams, G.D., H. Bauer, and H. Sprinz, *Influence of the normal flora on mucosal morphology and cellular renewal in the ileum. A comparison of germ-free and conventional mice*, 1962, DTIC Document.
58. Renz, H., P. Brandtzaeg, and M. Hornef, *The impact of perinatal immune development on mucosal homeostasis and chronic inflammation*. Nature Reviews Immunology, 2012. **12**(1): p. 9-23.
59. Uematsu, S., et al., *Regulation of humoral and cellular gut immunity by lamina propria dendritic cells expressing Toll-like receptor 5*. Nature Immunology, 2008. **9**(7): p. 769.
60. Suzuki, K., et al., *Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(7): p. 1981-1986.
61. Sanos, S.L., et al., *ROR γ t and commensal microflora are required for the differentiation of mucosal interleukin 22-producing NKp46(+) cells*. Nature Immunology, 2009. **10**(1): p. 83-91.
62. Sharma, R., et al., *Rat intestinal mucosal responses to a microbial flora and different diets*. Gut, 1995. **36**(2): p. 209-214.
63. Anderson, R., et al., *Lactobacillus plantarum MB452 enhances the function of the intestinal barrier by increasing the expression levels of genes involved in tight junction formation*. BMC Microbiology, 2010. **10**(1): p. 316.
64. Zelante, T., et al., *Tryptophan Catabolites from Microbiota Engage Aryl Hydrocarbon Receptor and Balance Mucosal Reactivity via Interleukin-22*. Immunity, 2013. **39**(2): p. 372-385.
65. Atarashi, K., et al., *Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota*. Nature, 2013. **500**(7461): p. 232-236.

66. Sjögren, K., et al., *The gut microbiota regulates bone mass in mice*. Journal of Bone and Mineral Research, 2012. **27**(6): p. 1357-1367.
67. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest*. Nature, 2006. **444**(7122): p. 1027-131.
68. Pop, M., et al., *Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition*. Genome Biology, 2014. **15**(6): p. R76.
69. Lupp, C., et al., *Host-Mediated Inflammation Disrupts the Intestinal Microbiota and Promotes the Overgrowth of Enterobacteriaceae*. Cell Host and Microbe, 2007. **2**(2): p. 119-129.
70. Molloy, Michael J., et al., *Intraluminal Containment of Commensal Outgrowth in the Gut during Infection-Induced Dysbiosis*. Cell Host & Microbe, 2013. **14**(3): p. 318-328.
71. Frank, D.N., et al., *Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(34): p. 13780-13785.
72. Winter, S.E., et al., *Host-Derived Nitrate Boosts Growth of E. coli in the Inflamed Gut*. Science, 2013. **339**(6120): p. 708-711.
73. Sartor, R.B. and S.K. Mazmanian, *Intestinal Microbes in Inflammatory Bowel Diseases*. American Journal of Gastroenterology Supplement, 2012. **1**(1): p. 15-21.
74. Bäumler, A.J., *The record of horizontal gene transfer in Salmonella*. Trends in Microbiology, 1997. **5**(8): p. 318-322.
75. Tindall, B.J., et al., *Nomenclature and taxonomy of the genus Salmonella*. International Journal of Systematic and Evolutionary Microbiology, 2005. **55**(1): p. 521-524.
76. Boyd, E.F., et al., *Molecular genetic relationships of the salmonellae*. Applied and Environmental Microbiology, 1996. **62**(3): p. 804-8.
77. McQuiston, J.R., et al., *Molecular Phylogeny of the Salmonellae: Relationships among Salmonella Species and Subspecies Determined from Four Housekeeping Genes and Evidence of Lateral Gene Transfer Events*. Journal of Bacteriology, 2008. **190**(21): p. 7060-7067.
78. Aleksic, S., F. Heinzerling, and J. Bockemühl, *Human Infection Caused by Salmonellae of Subspecies II to VI in Germany, 1977–1992*. Zentralblatt für Bakteriologie, 1996. **283**(3): p. 391-398.
79. U.S. Department of Health and Human Services, A., GA, Centers for Disease Control and Prevention, *Salmonella: annual summary, 2004*. 2005.
80. Achtman, M., et al., *Multilocus Sequence Typing as a Replacement for Serotyping in Salmonella enterica*. PLoS Pathogens, 2012. **8**(6): p. e1002776.
81. Galanis, E., et al., *Web-based Surveillance and Global Salmonella Distribution, 2000–2002*. Emerging Infectious Diseases, 2006. **12**(3): p. 381-388.
82. Hendriksen, R.S., et al., *Global Monitoring of Salmonella Serovar Distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank*:

- Results of Quality Assured Laboratories from 2001 to 2007*. Foodborne Pathogens and Disease, 2011. **8**(8): p. 887-900.
83. Santos, R.L., et al., *Animal models of Salmonella infections: enteritis versus typhoid fever*. Microbes and Infection, 2001. **3**(14–15): p. 1335-1344.
 84. Bhutta, Z., et al., *Background Paper on Vaccination against Typhoid Fever using New-Generation Vaccines - presented at the SAGE November 2007 meeting*.
 85. Majowicz, S.E., et al., *The Global Burden of Nontyphoidal Salmonella Gastroenteritis*. Clinical Infectious Diseases, 2010. **50**(6): p. 882-889.
 86. Mead, P.S., et al., *Food-related illness and death in the United States*. Emerging Infectious Diseases, 1999. **5**(5): p. 607-625.
 87. Gordon, M.A., *Salmonella infections in immunocompromised adults*. Journal of Infection, 2008. **56**(6): p. 413-422.
 88. Feasey, N.A., et al., *Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa*. Lancet, 2012. **379**(9835): p. 2489-2499.
 89. Okoro, C.K., et al., *Signatures of Adaptation in Human Invasive Salmonella Typhimurium ST313 Populations from Sub-Saharan Africa*. PLoS Neglected Tropical Diseases, 2015. **9**(3): p. e0003611.
 90. Raffatellu, M., et al., *Simian immunodeficiency virus-induced mucosal interleukin-17 deficiency promotes Salmonella dissemination from the gut*. Nature Medicine, 2008. **14**(4): p. 421-428.
 91. Murase, T., et al., *Fecal excretion of Salmonella enterica serovar Typhimurium following a food-borne outbreak*. Journal of Clinical Microbiology, 2000. **38**(9): p. 3495-3497.
 92. Levine, M.M., R.E. Black, and C. Lanata, *Precise estimation of the number of chronic carriers of Salmonella typhi in Santiago, Chile, an endemic area*. Journal of Infectious Diseases, 1982. **146**(6): p. 724-726.
 93. Gonzalez-Escobedo, G., J.M. Marshall, and J.S. Gunn, *Chronic and acute infection of the gall bladder by Salmonella Typhi: Understanding the carrier state*. Nature Reviews Microbiology, 2011. **9**(1): p. 9-14.
 94. Schiøler, H., et al., *Biliary calculi in chronic salmonella carriers and healthy controls. A controlled study*. Scandinavian Journal of Infectious Diseases, 1983. **15**(1): p. 17-19.
 95. Prouty, A.M., W.H. Schwesinger, and J.S. Gunn, *Biofilm formation and interaction with the surfaces of gallstones by Salmonella spp*. Infection and Immunity, 2002. **70**(5): p. 2640-2649.
 96. Roumagnac, P., et al., *Evolutionary History of Salmonella Typhi*. Science, 2006. **314**(5803): p. 1301-1304.
 97. Baker, S. and G. Dougan, *The Genome of Salmonella enterica Serovar Typhi*. Clinical Infectious Diseases, 2007. **45**(Supplement 1): p. S29-S33.

98. Wong, V.K., et al., *Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events*. Nature Genetics, 2015. **47**(6): p. 632-639.
99. McClelland, M., et al., *Complete genome sequence of Salmonella enterica serovar Typhimurium LT2*. Nature, 2001. **413**(6858): p. 852-856.
100. Deng, W., et al., *Comparative Genomics of Salmonella enterica Serovar Typhi Strains Ty2 and CT18*. Journal of Bacteriology, 2003. **185**(7): p. 2330-2337.
101. Cole, S.T., et al., *Massive gene decay in the leprosy bacillus*. Nature, 2001. **409**(6823): p. 1007-1011.
102. Thomson, N.R., et al., *The Complete Genome Sequence and Comparative Genome Analysis of the High Pathogenicity Yersinia enterocolitica Strain 8081*. PLoS Genetics, 2006. **2**(12): p. e206.
103. Pickard, D., et al., *Composition, Acquisition, and Distribution of the Vi Exopolysaccharide-Encoding Salmonella enterica Pathogenicity Island SPI-7*. Journal of Bacteriology, 2003. **185**(17): p. 5055-5065.
104. Raffatellu, M., et al., *The Capsule Encoding the viaB Locus Reduces Interleukin-17 Expression and Mucosal Innate Responses in the Bovine Intestinal Mucosa during Infection with Salmonella enterica Serotype Typhi*. Infection and Immunity, 2007. **75**(9): p. 4342-4350.
105. Garai, P., D.P. Gnanadhas, and D. Chakravorty, *Salmonella enterica serovars Typhimurium and Typhi as model organisms*. Virulence, 2012. **3**(4): p. 377-388.
106. Khoramian-Falsafi, T., et al., *Effect of motility and chemotaxis on the invasion of Salmonella typhimurium into HeLa cells*. Microbial Pathogenesis, 1990. **9**(1): p. 47-53.
107. Barman, M., et al., *Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract*. Infection and Immunity, 2008. **76**(3): p. 907-915.
108. Stecher, B., et al., *Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota*. PLoS Biology, 2007. **5**(10): p. 2177-2189.
109. Tsolis, R.M., et al., *From bench to bedside: stealth of enteroinvasive pathogens*. Nature Reviews Microbiology, 2008. **6**(12): p. 883-892.
110. Spees, A.M., et al., *Streptomycin-Induced Inflammation Enhances Escherichia coli Gut Colonization Through Nitrate Respiration*. mBio, 2013. **4**(4).
111. Shea, J.E., et al., *Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium*. Proceedings of the National Academy of Sciences, 1996. **93**(6): p. 2593-2597.
112. Kubori, T., et al., *Supramolecular Structure of the Salmonella typhimurium Type III Protein Secretion System*. Science, 1998. **280**(5363): p. 602-605.

113. Steele-Mortimer, O., et al., *The invasion-associated type III secretion system of Salmonella enterica serovar Typhimurium is necessary for intracellular proliferation and vacuole biogenesis in epithelial cells*. Cellular Microbiology, 2002. **4**(1): p. 43-54.
114. Galán, J.E. and R. Curtiss, *Cloning and molecular characterization of genes whose products allow Salmonella typhimurium to penetrate tissue culture cells*. Proceedings of the National Academy of Sciences, 1989. **86**(16): p. 6383-6387.
115. Finlay, B.B., S. Ruschkowski, and S. Dedhar, *Cytoskeletal rearrangements accompanying Salmonella entry into epithelial cells*. Journal of Cell Science, 1991. **99**(2): p. 283-296.
116. Zhou, D. and J. Galán, *Salmonella entry into host cells: The work in concert of type III secreted effector proteins*. Microbes and Infection, 2001. **3**(14-15): p. 1293-1298.
117. Zhou, D., M.S. Mooseker, and J.E. Galán, *Role of the S. typhimurium Actin-Binding Protein SipA in Bacterial Internalization*. Science, 1999. **283**(5410): p. 2092-2095.
118. Hayward, R.D. and V. Koronakis, *Direct nucleation and bundling of actin by the SipC protein of invasive Salmonella*. The EMBO Journal, 1999. **18**(18): p. 4926-4934.
119. Francis, C.L., et al., *Ruffles induced by Salmonella and other stimuli direct macropinocytosis of bacteria*. Nature, 1993. **364**(6438): p. 639-642.
120. Fu, Y. and J.E. Galán, *A Salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion*. Nature, 1999. **401**(6750): p. 293-297.
121. Ochman, H., et al., *Identification of a pathogenicity island required for Salmonella survival in host cells*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(15): p. 7800-7804.
122. Abrahams, G.L. and M. Hensel, *Manipulating cellular transport and immune responses: Dynamic interactions between intracellular Salmonella enterica and its host cells*. Cellular Microbiology, 2006. **8**(5): p. 728-737.
123. Hassett, D.J. and M.S. Cohen, *Bacterial adaptation to oxidative stress: implications for pathogenesis and interaction with phagocytic cells*. The FASEB Journal, 1989. **3**(14): p. 2574-82.
124. Alpuche-Aranda, C.M., et al., *Salmonella stimulate macrophage macropinocytosis and persist within spacious phagosomes*. The Journal of Experimental Medicine, 1994. **179**(2): p. 601-608.
125. Swanson, J.A. and C. Watts, *Macropinocytosis*. Trends in Cell Biology, 1995. **5**(11): p. 424-428.
126. Coburn, B., et al., *Salmonella enterica serovar Typhimurium pathogenicity island 2 is necessary for complete virulence in a mouse model of infectious enterocolitis*. Infection and Immunity, 2005. **73**(6): p. 3219-3227.

127. Kuhle, V., D. Jäckel, and M. Hensel, *Effector Proteins Encoded by Salmonella Pathogenicity Island 2 Interfere with the Microtubule Cytoskeleton after Translocation into Host Cells*. *Traffic*, 2004. **5**(5): p. 356-370.
128. Grant, A.J., et al., *Attenuated Salmonella Typhimurium Lacking the Pathogenicity Island-2 Type 3 Secretion System Grow to High Bacterial Numbers inside Phagocytes in Mice*. *PLoS Pathogens*, 2012. **8**(12): p. e1003070.
129. Cheminay, C., A. Möhlenbrink, and M. Hensel, *Intracellular Salmonella Inhibit Antigen Presentation by Dendritic Cells*. *The Journal of Immunology*, 2005. **174**(5): p. 2892-2899.
130. Hapfelmeier, S., et al., *The Salmonella Pathogenicity Island (SPI)-2 and SPI-1 Type III Secretion Systems Allow Salmonella Serovar typhimurium to Trigger Colitis via MyD88-Dependent and MyD88-Independent Mechanisms*. *The Journal of Immunology*, 2005. **174**(3): p. 1675-1685.
131. Rohmer, L., D. Hocquet, and S.I. Miller, *Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis*. *Trends in Microbiology*, 2011. **19**(7): p. 341-348.
132. Stecher, B., et al., *Motility allows S. Typhimurium to benefit from the mucosal defence*. *Cellular Microbiology*, 2008. **10**(5): p. 1166-1180.
133. Winter, S.E., et al., *Gut inflammation provides a respiratory electron acceptor for Salmonella*. *Nature*, 2010. **467**(7314): p. 426-429.
134. Flo, T.H., et al., *Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron*. *Nature*, 2004. **432**(7019): p. 917-921.
135. Raffatellu, M., et al., *Lipocalin-2 Resistance Confers an Advantage to Salmonella enterica Serotype Typhimurium for Growth and Survival in the Inflamed Intestine*. *Cell Host and Microbe*, 2009. **5**(5): p. 476-486.
136. Guo, L., et al., *Regulation of Lipid A Modifications by Salmonella typhimurium Virulence Genes phoP-phoQ*. *Science*, 1997. **276**(5310): p. 250-253.
137. Bader, M.W., et al., *Recognition of antimicrobial peptides by a bacterial sensor kinase*. *Cell*, 2005. **122**(3): p. 461-472.
138. Ernst, R.K., T. Guina, and S.I. Miller, *How Intracellular Bacteria Survive: Surface Modifications That Promote Resistance to Host Innate Immune Responses*. *Journal of Infectious Diseases*, 1999. **179**(Supplement 2): p. S326-S330.
139. Kaiser, P., et al., *The streptomycin mouse model for Salmonella diarrhea: functional analysis of the microbiota, the pathogen's virulence factors, and the host's mucosal immune response*. *Immunological Reviews*, 2012. **245**(1): p. 56-83.
140. Khan, S.A., et al., *A lethal role for lipid A in Salmonella infections*. *Molecular Microbiology*, 1998. **29**(2): p. 571-579.

141. Stecher, B., et al., *Chronic Salmonella enterica Serovar Typhimurium-Induced Colitis and Cholangitis in Streptomycin-Pretreated Nramp1+/+ Mice*. *Infection and Immunity*, 2006. **74**(9): p. 5047-5057.
142. Monack, D.M., D.M. Bouley, and S. Falkow, *Salmonella typhimurium Persists within Macrophages in the Mesenteric Lymph Nodes of Chronically Infected Nramp1(+)(/)(+) Mice and Can Be Reactivated by IFN γ Neutralization*. *The Journal of Experimental Medicine*, 2004. **199**(2): p. 231-241.
143. Vidal, S., et al., *The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene*. *The Journal of Experimental Medicine*, 1995. **182**(3): p. 655-666.
144. Vidal, S.M., et al., *Natural resistance to intracellular infections: Nramp1 encodes a membrane phosphoglycoprotein absent in macrophages from susceptible (Nramp1 D169) mouse strains*. *The Journal of Immunology*, 1996. **157**(8): p. 3559-68.
145. Canonne-Hergaux, F., et al., *The Nramp1 Protein and Its Role in Resistance to Infection and Macrophage Function*. *Proceedings of the Association of American Physicians*, 1999. **111**(4): p. 283-289.
146. Frost, A.J., A.P. Bland, and T.S. Wallis, *The Early Dynamic Response of the Calf Ileal Epithelium to Salmonella typhimurium*. *Veterinary Pathology Online*, 1997. **34**(5): p. 369-386.
147. Coombes, B.K., et al., *Analysis of the Contribution of Salmonella Pathogenicity Islands 1 and 2 to Enteric Disease Progression Using a Novel Bovine Ileal Loop Model and a Murine Model of Infectious Enterocolitis*. *Infection and Immunity*, 2005. **73**(11): p. 7161-7169.
148. Miller, C.P. and M. Bohnhoff, *Changes in the Mouse's Enteric Microflora Associated with Enhanced Susceptibility to Salmonella Infection following Streptomycin Treatment*. *The Journal of Infectious Diseases*, 1963. **113**(1): p. 59-66.
149. Barthel, M., et al., *Pretreatment of mice with streptomycin provides a Salmonella enterica serovar Typhimurium colitis model that allows analysis of both pathogen and host*. *Infection and Immunity*, 2003. **71**(5): p. 2839-2858.
150. Hapfelmeier, S., et al., *Role of the Salmonella Pathogenicity Island 1 Effector Proteins SipA, SopB, SopE, and SopE2 in Salmonella enterica Subspecies 1 Serovar Typhimurium Colitis in Streptomycin-Pretreated Mice*. *Infection and Immunity*, 2004. **72**(2): p. 795-809.
151. Songhet, P., et al., *Stromal IFN- γ R-Signaling Modulates Goblet Cell Function During Salmonella Typhimurium Infection*. *PLoS ONE*, 2011. **6**(7): p. e22459.
152. Behnsen, J., et al., *The Cytokine IL-22 Promotes Pathogen Colonization by Suppressing Related Commensal Bacteria*. *Immunity*, 2014. **40**(2): p. 262-273.
153. Loftus Jr, E.V., *Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences*. *Gastroenterology*, 2004. **126**(6): p. 1504-1517.

154. Wiles, S., et al., *Organ specificity, colonization and clearance dynamics in vivo following oral challenges with the murine pathogen Citrobacter rodentium*. Cellular Microbiology, 2004. **6**(10): p. 963-972.
155. Garmendia, J., G. Frankel, and V.F. Crepin, *Enteropathogenic and Enterohemorrhagic Escherichia coli Infections: Translocation, Translocation, Translocation*. Infection and Immunity, 2005. **73**(5): p. 2573-2585.
156. Simmons, C.P., et al., *Impaired Resistance and Enhanced Pathology During Infection with a Noninvasive, Attaching-Effacing Enteric Bacterial Pathogen, Citrobacter rodentium, in Mice Lacking IL-12 or IFN- γ* . The Journal of Immunology, 2002. **168**(4): p. 1804-1812.
157. Simmons, C.P., et al., *Central Role for B Lymphocytes and CD4+ T Cells in Immunity to Infection by the Attaching and Effacing Pathogen Citrobacter rodentium*. Infection and Immunity, 2003. **71**(9): p. 5077-5086.
158. Bry, L. and M.B. Brenner, *Critical Role of T Cell-Dependent Serum Antibody, but Not the Gut-Associated Lymphoid Tissue, for Surviving Acute Mucosal Infection with Citrobacter rodentium, an Attaching and Effacing Pathogen*. The Journal of Immunology, 2004. **172**(1): p. 433-441.
159. Maaser, C., et al., *Clearance of Citrobacter rodentium Requires B Cells but Not Secretory Immunoglobulin A (IgA) or IgM Antibodies*. Infection and Immunity, 2004. **72**(6): p. 3315-3324.
160. Mundy, R., et al., *Citrobacter rodentium of mice and man*. Cellular Microbiology, 2005. **7**(12): p. 1697-1706.
161. Yan, Y., et al., *Temporal and Spatial Analysis of Clinical and Molecular Parameters in Dextran Sodium Sulfate Induced Colitis*. PLoS ONE, 2009. **4**(6): p. e6073.
162. Hans, W., et al., *The role of the resident intestinal flora in acute and chronic dextran sulfate sodium-induced colitis in mice*. European Journal of Gastroenterology and Hepatology, 2000. **12**(3): p. 267-273.
163. Hudcovic, T., et al., *The role of microflora in the development of intestinal inflammation: acute and chronic colitis induced by dextran sulfate in germ-free and conventionally reared immunocompetent and immunodeficient mice*. Folia Microbiologica, 2001. **46**(6): p. 565-572.
164. Mannick, E.E., et al., *Altered phenotype of dextran sulfate sodium colitis in interferon regulatory factor-1 knock-out mice*. Journal of Gastroenterology and Hepatology, 2005. **20**(3): p. 371-380.
165. Araki, A., et al., *MyD88-deficient mice develop severe intestinal inflammation in dextran sodium sulfate colitis*. Journal of Gastroenterology, 2005. **40**(1): p. 16-23.
166. Fukata, M., et al., *Toll-like receptor-4 is required for intestinal response to epithelial injury and limiting bacterial translocation in a murine model of acute colitis*. American Journal of Physiology-Gastrointestinal and Liver Physiology, 2005. **288**(5): p. G1055-G1065.

167. Vance, R.E., R.R. Isberg, and D.A. Portnoy, *Patterns of Pathogenesis: Discrimination of Pathogenic and Nonpathogenic Microbes by the Innate Immune System*. Cell Host and Microbe, 2009. **6**(1): p. 10-21.
168. Godinez, I., et al., *T cells help to amplify inflammatory responses induced by Salmonella enterica serotype Typhimurium in the intestinal mucosa*. Infection and Immunity, 2008. **76**(5): p. 2008-2017.
169. Haas, P.J. and J. Van Strijp, *Anaphylatoxins: Their role in bacterial infection and inflammation*. Immunologic Research, 2007. **37**(3): p. 161-175.
170. Gewirtz, A.T., et al., *Cutting edge: Bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression*. Journal of Immunology, 2001. **167**(4): p. 1882-1885.
171. Tükel, Ç., et al., *Toll-like receptors 1 and 2 cooperatively mediate immune responses to curli, a common amyloid from enterobacterial biofilms*. Cellular Microbiology, 2010. **12**(10): p. 1495-1505.
172. MacLennan, C., et al., *Interleukin (IL)-12 and IL-23 Are Key Cytokines for Immunity against Salmonella in Humans*. Journal of Infectious Diseases, 2004. **190**(10): p. 1755-1757.
173. Mastroeni, P. and A. Grant, *Dynamics of spread of Salmonella enterica in the systemic compartment*. Microbes and Infection, 2013. **15**(13): p. 849-857.
174. Chen, L.M., K. Kaniga, and J.E. Galán, *Salmonella spp. are cytotoxic for cultured macrophages*. Molecular Microbiology, 1996. **21**(5): p. 1101-1115.
175. Miao, E.A., et al., *Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria*. Nature Immunology, 2010. **11**(12): p. 1136-1142.
176. Mittrücker, H.W. and S.H. Kaufmann, *Immune response to infection with Salmonella typhimurium in mice*. Journal of Leukocyte Biology, 2000. **67**(4): p. 457-63.
177. Geddes, K., et al., *Identification of an innate T helper type 17 response to intestinal bacterial pathogens*. Nature Medicine, 2011. **17**(7): p. 837-844.
178. Noriega, L.M., et al., *Salmonella infections in a cancer center*. Supportive Care in Cancer, 1994. **2**(2): p. 116-122.
179. Nunes, J.S., et al., *Morphologic and cytokine profile characterization of salmonella enterica serovar typhimurium infection in calves with bovine leukocyte adhesion deficiency*. Veterinary Pathology, 2010. **47**(2): p. 322-333.
180. Godinez, I., et al., *Interleukin-23 orchestrates mucosal responses to Salmonella enterica serotype typhimurium in the intestine*. Infection and Immunity, 2009. **77**(1): p. 387-398.
181. Gautreaux, M.D., E.A. Dietch, and R.D. Berg, *T lymphocytes in host defense against bacterial translocation from the gastrointestinal tract*. Infection and Immunity, 1994. **62**(7): p. 2874-2884.

182. Thiennimitr, P., S.E. Winter, and A.J. Bäumlner, *Salmonella, the host and its microbiota*. Current Opinion in Microbiology, 2012. **15**(1): p. 108-114.
183. Nauciel, C., *Role of CD4+ T cells and T-independent mechanisms in acquired resistance to Salmonella typhimurium infection*. The Journal of Immunology, 1990. **145**(4): p. 1265-9.
184. Mittrücker, H.-W., A. Köhler, and S.H.E. Kaufmann, *Characterization of the Murine T-Lymphocyte Response to Salmonella enterica Serovar Typhimurium Infection*. Infection and Immunity, 2002. **70**(1): p. 199-203.
185. Kupz, A., S. Bedoui, and R.A. Strugnell, *Cellular Requirements for Systemic Control of Salmonella enterica Serovar Typhimurium Infections in Mice*. Infection and Immunity, 2014. **82**(12): p. 4997-5004.
186. Brown, A. and C.E. Hormaeche, *The antibody response to salmonellae in mice and humans studied by immunoblots and ELISA*. Microbial Pathogenesis, 1989. **6**(6): p. 445-454.
187. Eisenstein, T.K., L.M. Killar, and B.M. Sultzer, *Immunity to Infection with Salmonella typhimurium: Mouse-Strain Differences in Vaccine- and Serum-Mediated Protection*. Journal of Infectious Diseases, 1984. **150**(3): p. 425-435.
188. Guzman, C.A., et al., *Vaccines against typhoid fever*. Vaccine, 2006. **24**(18): p. 3804-3811.
189. Martinoli, C., A. Chiavelli, and M. Rescigno, *Entry Route of Salmonella typhimurium Directs the Type of Induced Immune Response*. Immunity, 2007. **27**(6): p. 975-984.
190. Patel, S.Y., et al., *Genetically determined susceptibility to mycobacterial infection*. Journal of Clinical Pathology, 2008. **61**(9): p. 1006-1012.
191. Ottenhoff, T.H.M., et al., *Genetics, cytokines and human infectious disease: lessons from weakly pathogenic mycobacteria and salmonellae*. Nature Genetics, 2002. **32**(1): p. 97-105.
192. Jostins, L., et al., *Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease*. Nature, 2012. **491**(7422): p. 119-124.
193. Lees, C.W., et al., *New IBD genetics: common pathways with other diseases*. Gut, 2011. **60**(12): p. 1739-1753.
194. Travassos, L.H., et al., *Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry*. Nature Immunology, 2010. **11**(1): p. 55-62.
195. Kuczynski, J., et al., *Experimental and analytical tools for studying the human microbiome*. Nature Reviews Genetics, 2012. **13**(1): p. 47-58.
196. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions*. Genome Biology, 2013. **14**(4): p. R36.
197. Love, M., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biology, 2014. **15**(12): p. 550.
198. Kozich, J.J., et al., *Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform*. Applied and Environmental Microbiology, 2013. **79**(17): p. 5112-5120.

199. Letunic, I. and P. Bork, *Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation*. *Bioinformatics*, 2007. **23**(1): p. 127-128.
200. Wisniewski, J.R., et al., *Universal sample preparation method for proteome analysis*. *Nature Methods*, 2009. **6**(5): p. 359-362.
201. Oliveros, J.C., *Venny. An interactive tool for comparing lists with Venn's diagrams*. 2007-2015.
202. Hulsen, T., J. de Vlieg, and W. Alkema, *BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams*. *BMC Genomics*, 2008. **9**(1): p. 488.
203. Que, J.U. and D.J. Hentges, *Effect of streptomycin administration on colonization resistance to Salmonella typhimurium in mice*. *Infection and Immunity*, 1985. **48**(1): p. 169-174.
204. Bohnhoff, M. and C.P. Miller, *Enhanced Susceptibility to Salmonella Infection in Streptomycin-Treated Mice*. *The Journal of Infectious Diseases*, 1962. **111**(2): p. 117-127.
205. Bohnhoff, M., B.L. Drake, and C.P. Miller, *Effect of Streptomycin on Susceptibility of Intestinal Tract to Experimental Salmonella Infection*. *Experimental Biology and Medicine*, 1954. **86**(1): p. 132-137.
206. Sekirov, I., et al., *Antibiotic-Induced Perturbations of the Intestinal Microbiota Alter Host Susceptibility to Enteric Infection*. *Infection and Immunity*, 2008. **76**(10): p. 4726-4736.
207. Su, Z., et al., *Comparing Next-Generation Sequencing and Microarray Technologies in a Toxicological Study of the Effects of Aristolochic Acid on Rat Kidneys*. *Chemical Research in Toxicology*, 2011. **24**(9): p. 1486-1493.
208. Guida, A., et al., *Using RNA-seq to determine the transcriptional landscape and the hypoxic response of the pathogenic yeast Candida parapsilosis*. *BMC Genomics*, 2011. **12**(1): p. 628.
209. Nalpas, N., et al., *Whole-transcriptome, high-throughput RNA sequence analysis of the bovine macrophage response to Mycobacterium bovis infection in vitro*. *BMC Genomics*, 2013. **14**(1): p. 230.
210. Chen, H., et al., *Genome-Wide Gene Expression Profiling of Nucleus Accumbens Neurons Projecting to Ventral Pallidum Using both Microarray and Transcriptome Sequencing*. *Frontiers in Neuroscience*, 2011. **5**: p. 98.
211. Stockhammer, O.W., et al., *Transcriptome Profiling and Functional Analyses of the Zebrafish Embryonic Innate Immune Response to Salmonella Infection*. *The Journal of Immunology*, 2009. **182**(9): p. 5641-5653.
212. Schreiber, F., et al., *The Human Transcriptome During Nontyphoid Salmonella and HIV Coinfection Reveals Attenuated NFκB-Mediated Inflammation and Persistent Cell Cycle Disruption*. *Journal of Infectious Diseases*, 2011. **204**(8): p. 1237-1245.

213. Liu, X., et al., *Global analysis of the eukaryotic pathways and networks regulated by Salmonella typhimurium in mouse intestinal infection in vivo*. BMC Genomics, 2010. **11**(1): p. 722.
214. Khatri, P., M. Sirota, and A.J. Butte, *Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges*. PLoS Computational Biology, 2012. **8**(2): p. e1002375.
215. Lynn, D.J., et al., *InnateDB: facilitating systems-level analyses of the mammalian innate immune response*. Molecular Systems Biology, 2008. **4**(1): p. n/a-n/a.
216. Breuer, K., et al., *InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation*. Nucleic Acids Research, 2013. **41**(D1): p. D1228-D1233.
217. Lynn, D., et al., *Curating the innate immunity interactome*. BMC Systems Biology, 2010. **4**(1): p. 117.
218. Everest, P., et al., *Evaluation of Salmonella typhimurium Mutants in a Model of Experimental Gastroenteritis*. Infection and Immunity, 1999. **67**(6): p. 2815-2821.
219. Sabat, G., et al., *Selective and Sensitive Method for PCR Amplification of Escherichia coli 16S rRNA Genes in Soil*. Applied and Environmental Microbiology, 2000. **66**(2): p. 844-849.
220. Ferreira, R.B.R., et al., *The Intestinal Microbiota Plays a Role in Salmonella-Induced Colitis Independent of Pathogen Colonization*. PLoS ONE, 2011. **6**(5): p. e20338.
221. Parks, W.C., C.L. Wilson, and Y.S. Lopez-Boado, *Matrix metalloproteinases as modulators of inflammation and innate immunity*. Nature Reviews Immunology, 2004. **4**(8): p. 617-629.
222. Elkington, P.T.G., C.M. O'Kane, and J.S. Friedland, *The paradox of matrix metalloproteinases in infectious disease*. Clinical and Experimental Immunology, 2005. **142**(1): p. 12-20.
223. Jenner, R.G. and R.A. Young, *Insights into host responses against pathogens from transcriptional profiling*. Nature Reviews Microbiology, 2005. **3**(4): p. 281-294.
224. Hardin, J.A., et al., *Aquaporin expression is downregulated in a murine model of colitis and in patients with ulcerative colitis, Crohn's disease and infectious colitis*. Cell and Tissue Research, 2004. **318**(2): p. 313-323.
225. Guttman, J.A., et al., *Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens*. Cellular Microbiology, 2007. **9**(1): p. 131-141.
226. Saeki, N., et al., *Gasdermin (Gsdm) localizing to mouse Chromosome 11 is predominantly expressed in upper gastrointestinal tract but significantly suppressed in human gastric cancer cells*. Mammalian Genome, 2000. **11**(9): p. 718-724.
227. Saeki, N. and H. Sasaki, *Gasdermin Superfamily: A Novel Gene Family Functioning in Epithelial Cells*. Endothelium and Epithelium2012: Nova Science Publishers Inc.
228. Antunes, L.C.M., et al., *Impact of Salmonella Infection on Host Hormone Metabolism Revealed by Metabolomics*. Infection and Immunity, 2011. **79**(4): p. 1759-1769.
229. Beisel, W.R., et al., *Metabolic effects of intracellular infections in man*. Annals of Internal Medicine, 1967. **67**(4): p. 744-779.

230. Hardardottir, I., C. Grunfeld, and K. Feingold, *Effects of endotoxin on lipid metabolism*. Biochemical Society Transactions, 1995. **23**(4): p. 1013-1018.
231. Khovidhunkit, W., et al., *Effects of infection and inflammation on lipid and lipoprotein metabolism: mechanisms and consequences to the host*. The Journal of Lipid Research, 2004. **45**(7): p. 1169-1196.
232. Wasinger, V.C., M. Zeng, and Y. Yau, *Current Status and Advances in Quantitative Proteomic Mass Spectrometry*. International Journal of Proteomics, 2013. **2013**: p. 12.
233. Zhu, W., J.W. Smith, and C.-M. Huang, *Mass Spectrometry-Based Label-Free Quantitative Proteomics*. Journal of Biomedicine and Biotechnology, 2010. **2010**: p. 6.
234. Guo, T., et al., *Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps*. Nature Medicine, 2015. **21**(4): p. 407-413.
235. Zhang, B., et al., *Proteogenomic characterization of human colon and rectal cancer*. Nature, 2014. **513**(7518): p. 382-387.
236. Levin, Y., E. Hradetzky, and S. Bahn, *Quantification of proteins using data-independent analysis (MSE) in simple and complex samples: A systematic evaluation*. Proteomics, 2011. **11**(16): p. 3273-3287.
237. Distler, U., et al., *Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics*. Nature Methods, 2014. **11**(2): p. 167-170.
238. Distler, U., et al., *In-depth protein profiling of the postsynaptic density from mouse hippocampus using data-independent acquisition proteomics*. Proteomics, 2014. **14**(21-22): p. 2607-2613.
239. Yang, Y., et al., *Mass spectrometry-based proteomic approaches to study pathogenic bacteria-host interactions*. Protein & Cell, 2015. **6**(4): p. 265-274.
240. Shi, L., et al., *Proteome of Salmonella Enterica Serotype Typhimurium Grown in a Low Mg(2+)/pH Medium*. Journal of proteomics & bioinformatics, 2009. **2**: p. 388-397.
241. Becker, D., et al., *Robust Salmonella metabolism limits possibilities for new antimicrobials*. Nature, 2006. **440**(7082): p. 303-307.
242. Shi, L., et al., *Proteomic Investigation of the Time Course Responses of RAW 264.7 Macrophages to Infection with Salmonella enterica*. Infection and Immunity, 2009. **77**(8): p. 3227-3233.
243. Hardwidge, P.R., et al., *Proteomic Analysis of the Intestinal Epithelial Cell Response to Enteropathogenic Escherichia coli*. Journal of Biological Chemistry, 2004. **279**(19): p. 20127-20136.
244. Meissner, F., et al., *Direct Proteomic Quantification of the Secretome of Activated Immune Cells*. Science, 2013. **340**(6131): p. 475-478.
245. Meissner, F. and M. Mann, *Quantitative shotgun proteomics: considerations for a high-quality workflow in immunology*. Nature Immunology, 2014. **15**(2): p. 112-117.

246. Filipowicz, W., S.N. Bhattacharyya, and N. Sonenberg, *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nature Reviews Genetics, 2008. **9**(2): p. 102-114.
247. Anderson, P. and N. Kedersha, *RNA granules: post-transcriptional and epigenetic modulators of gene expression.* Nature Reviews Molecular Cell Biology, 2009. **10**(6): p. 430-436.
248. Anderson, P. and N. Kedersha, *Stressful initiations.* Journal of Cell Science, 2002. **115**(16): p. 3227-3234.
249. Adeli, K., *Translational control mechanisms in metabolic regulation: critical role of RNA binding proteins, microRNAs, and cytoplasmic RNA granules.* Vol. 301. 2011. E1051-E1064.
250. Sun, N., et al., *Quantitative Proteome and Transcriptome Analysis of the Archaeon Thermoplasma acidophilum Cultured under Aerobic and Anaerobic Conditions.* Journal of Proteome Research, 2010. **9**(9): p. 4839-4850.
251. Dressaire, C., et al., *Transcriptome and Proteome Exploration to Model Translation Efficiency and Protein Stability in Lactococcus lactis.* PLoS Computational Biology, 2009. **5**(12): p. e1000606.
252. Gygi, S.P., et al., *Correlation between protein and mRNA abundance in yeast.* Molecular and Cellular Biology, 1999. **19**(3): p. 1720-1730.
253. Anderson, L. and J. Seilhamer, *A comparison of selected mRNA and protein abundances in human liver.* Electrophoresis, 1997. **18**(3-4): p. 533-537.
254. Ideker, T., et al., *Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network.* Science, 2001. **292**(5518): p. 929-934.
255. Chen, G., et al., *Discordant Protein and mRNA Expression in Lung Adenocarcinomas.* Molecular & Cellular Proteomics, 2002. **1**(4): p. 304-313.
256. Hegde, P.S., I.R. White, and C. Debouck, *Interplay of transcriptomics and proteomics.* Current Opinion in Biotechnology, 2003. **14**: p. 647-651.
257. Nie, L., et al., *Integrative Analysis of Transcriptomic and Proteomic Data: Challenges, Solutions and Applications.* Critical Reviews in Biotechnology, 2007. **27**(2): p. 63-75.
258. Kamisoglu, K., et al., *Tandem Analysis of Transcriptome and Proteome Changes after a Single Dose of Corticosteroid: A Systems Approach to Liver Function in Pharmacogenomics.* OMICS: A Journal of Integrative Biology, 2015. **19**(2): p. 80-91.
259. Jovanovic, M., et al., *Dynamic profiling of the protein life cycle in response to pathogens.* Science, 2015. **347**(6226).
260. Zong, Q., et al., *Messenger RNA translation state: The second dimension of high-throughput expression screening.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(19): p. 10632-10636.
261. Ingolia, N.T., *Ribosome profiling: new views of translation, from single codons to genome scale.* Nature Reviews Genetics, 2014. **15**(3): p. 205-213.

262. Wrzodek, C., et al., *InCroMAP: integrated analysis of cross-platform microarray and pathway data*. Bioinformatics, 2013. **29**(4): p. 506-508.
263. Zhang, W., F. Li, and L. Nie, *Integrating multiple 'omics' analysis for microbial biology: application and methodologies*. Microbiology, 2010. **156**(2): p. 287-301.
264. Joyce, A.R. and B.Ø. Palsson, *The model organism as a system: integrating 'omics' data sets*. Nature Reviews Molecular Cell Biology, 2006. **7**(3): p. 198-210.
265. Ning, K., D. Fermin, and A.I. Nesvizhskii, *Comparative Analysis of Different Label-Free Mass Spectrometry Based Protein Abundance Estimates and Their Correlation with RNA-Seq Gene Expression Data*. Journal of Proteome Research, 2012. **11**(4): p. 2261-2271.
266. Uhlén, M., et al., *Tissue-based map of the human proteome*. Science, 2015. **347**(6220).
267. Zubarev, R.A., *The challenge of the proteome dynamic range and its implications for in-depth proteomics*. PROTEOMICS, 2013. **13**(5): p. 723-726.
268. Schwanhäusser, B., et al., *Global quantification of mammalian gene expression control*. Nature, 2011. **473**(7347): p. 337-342.
269. Bruderer, R., et al., *Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues*. Molecular & Cellular Proteomics, 2015. **14**(5): p. 1400-1410.
270. Richards, A.L., A.E. Merrill, and J.J. Coon, *Proteome sequencing goes deep*. Current Opinion in Chemical Biology, 2015. **24**: p. 11-17.
271. Le Roch, K.G., et al., *Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle*. Genome Research, 2004. **14**(11): p. 2308-2318.
272. Koussounadis, A., et al., *Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system*. Scientific Reports, 2015. **5**: p. 10775.
273. Kamburov, A., et al., *Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA*. Bioinformatics, 2011. **27**(20): p. 2917-2918.
274. Nesargikar, P.N., B. Spiller, and R. Chavez, *The complement system: history, pathways, cascade and inhibitors*. European Journal of Microbiology & Immunology, 2012. **2**(2): p. 103-111.
275. Nielsen, C.H. and R.G.Q. Leslie, *Complement's participation in acquired immunity*. Journal of Leukocyte Biology, 2002. **72**(2): p. 249-261.
276. Harboe, M. and T.E. Mollnes, *The alternative complement pathway revisited*. Journal of Cellular and Molecular Medicine, 2008. **12**(4): p. 1074-1084.
277. Pepys, M.B., *Role of complement in induction of antibody production in vivo: effect of cobra factor and other C3-reactive agents on thymus-dependent and thymus independent antibody responses*. The Journal of Experimental Medicine, 1974. **140**(1): p. 126-145.
278. Carter, R. and D. Fearon, *CD19: lowering the threshold for antigen receptor stimulation of B lymphocytes*. Science, 1992. **256**(5053): p. 105-107.

279. Kopf, M., et al., *Complement component C3 promotes T-cell priming and lung migration to control acute influenza virus infection*. *Nature Medicine*, 2002. **8**(4): p. 373-378.
280. Dunkelberger, J.R. and W.-C. Song, *Complement and its role in innate and adaptive immune responses*. *Cell Research*, 2009. **20**(1): p. 34-50.
281. Wagner, E. and M.M. Frank, *Therapeutic potential of complement modulation*. *Nature Reviews Drug Discovery*, 2010. **9**(1): p. 43-56.
282. Siggins, M.K., et al., *Differential timing of antibody-mediated phagocytosis and cell-free killing of invasive African Salmonella allows immune evasion*. *European Journal of Immunology*, 2014. **44**(4): p. 1093-1098.
283. MacLennan, C.A., et al., *The neglected role of antibody in protection against bacteremia caused by nontyphoidal strains of Salmonella in African children*. *The Journal of Clinical Investigation*, 2008. **118**(4): p. 1553-1562.
284. Mäkelä, P.H., et al., *Salmonella, complement and mouse macrophages*. *Immunology Letters*, 1988. **19**(3): p. 217-222.
285. Saxén, H., I. Reima, and P.H. Mäkelä, *Alternative complement pathway activation by Salmonella O polysaccharide as a virulence determinant in the mouse*. *Microbial Pathogenesis*, 1987. **2**(1): p. 15-28.
286. Marcus, S., D.W. Esplin, and D.M. Donaldson, *Lack of Bactericidal Effect of Mouse Serum on a Number of Common Microorganisms*. *Science*, 1954. **119**(3103): p. 877-877.
287. Siggins, M.K., et al., *Absent Bactericidal Activity of Mouse Serum against Invasive African Nontyphoidal Salmonella Results from Impaired Complement Function but Not a Lack of Antibody*. *The Journal of Immunology*, 2011. **186**(4): p. 2365-2371.
288. Warren, J., et al., *Increased Susceptibility of C1q-Deficient Mice to Salmonella enterica Serovar Typhimurium Infection*. *Infection and Immunity*, 2002. **70**(2): p. 551-557.
289. Nakano, A., E. Kita, and S. Kashiba, *Different sensitivity of complement to Salmonella typhimurium accounts for the difference in natural resistance to murine typhoid between A/J and C57BL/6 mice*. *Microbiology and Immunology*, 1995. **39**(2): p. 95-103.
290. Alper, C.A., et al., *Human C'3: Evidence for the Liver as the Primary Site of Synthesis*. *Science*, 1969. **163**(3864): p. 286-288.
291. Hetland, G., et al., *Synthesis of complement components C5, C6, C7, C8 and C9 in vitro by human monocytes and assembly of the terminal complement complex*. *Scandinavian Journal of Immunology*, 1986. **24**(4): p. 421-428.
292. Singhrao, S.K., et al., *Role of Complement in the Aetiology of Pick's Disease?* *Journal of Neuropathology and Experimental Neurology*, 1996. **55**(5): p. 578-593.
293. Sacks, S.H., et al., *Endogenous complement C3 synthesis in immune complex nephritis*. *The Lancet*, 1993. **342**(8882): p. 1273-1274.

294. Neumann, E., et al., *Local production of complement proteins in rheumatoid arthritis synovium*. *Arthritis and Rheumatism*, 2002. **46**(4): p. 934-945.
295. Kopp, Z.A., et al., *Do Antimicrobial Peptides and Complement Collaborate in the Intestinal Mucosa?* *Frontiers in Immunology*, 2015. **6**: p. 17.
296. Laufer, J., et al., *Cellular localization of complement C3 and C4 transcripts in intestinal specimens from patients with Crohn's disease*. *Clinical and Experimental Immunology*, 2000. **120**(1): p. 30-37.
297. Sugihara, T., et al., *The increased mucosal mRNA expressions of complement C3 and interleukin-17 in inflammatory bowel disease*. *Clinical and Experimental Immunology*, 2010. **160**(3): p. 386-393.
298. Uemura, K., et al., *L-MBP Is Expressed in Epithelial Cells of Mouse Small Intestine*. *The Journal of Immunology*, 2002. **169**(12): p. 6945-6950.
299. Halstensen, T.S., et al., *Surface epithelium related activation of complement differs in Crohn's disease and ulcerative colitis*. *Gut*, 1992. **33**(7): p. 902-908.
300. Duncombe, V., et al., *Local and Systemic Complement Activity in Small Intestinal Bacterial Overgrowth*. *Digestive Diseases and Sciences*, 1997. **42**(6): p. 1128-1136.
301. Yasojima, K., et al., *Complement Gene Expression by Rabbit Heart: Upregulation by Ischemia and Reperfusion*. *Circulation Research*, 1998. **82**(11): p. 1224-1230.
302. Wirthmueller, U., et al., *Properdin, a positive regulator of complement activation, is released from secondary granules of stimulated peripheral blood neutrophils*. *The Journal of Immunology*, 1997. **158**(9): p. 4444-51.
303. Lu, F., S.M. Fernandes, and A.E. Davis, *The role of the complement and contact systems in the dextran sulfate sodium-induced colitis model: the effect of C1 inhibitor in inflammatory bowel disease*. Vol. 298. 2010. G878-G883.
304. Wende, E., et al., *The Complement Anaphylatoxin C3a Receptor (C3aR) Contributes to the Inflammatory Response in Dextran Sulfate Sodium (DSS)-Induced Colitis in Mice*. *PLoS ONE*, 2013. **8**(4): p. e62257.
305. Johswich, K., et al., *Role of the C5a receptor (C5aR) in acute and chronic dextran sulfate-induced models of inflammatory bowel disease*. *Inflammatory Bowel Diseases*, 2009. **15**(12): p. 1812-1823.
306. Deguchi, Y., et al., *Development of dextran sulfate sodium-induced colitis is aggravated in mice genetically deficient for complement C5*. *International Journal of Molecular Medicine*, 2005. **16**(4): p. 605-608.
307. Wen, L., J.P. Atkinson, and P.C. Giclas, *Clinical and laboratory evaluation of complement deficiency*. *Journal of Allergy and Clinical Immunology*, 2004. **113**(4): p. 585-593.

308. Østvik, A.E., et al., *Mucosal Toll-like Receptor 3-dependent Synthesis of Complement Factor B and Systemic Complement Activation in Inflammatory Bowel Disease*. *Inflammatory Bowel Diseases*, 2014. **20**(6): p. 995-1003.
309. Belzer, C., et al., *The role of specific IgG and complement in combating a primary mucosal infection of the gut epithelium*. *European Journal of Microbiology and Immunology*, 2011. **1**(4): p. 311-318.
310. Jain, U., et al., *Properdin Provides Protection from Citrobacter rodentium–Induced Intestinal Inflammation in a C5a/IL-6–Dependent Manner*. *The Journal of Immunology*, 2015. **194**(7): p. 3414-3421.
311. Jain, U., et al., *The Complement System in Inflammatory Bowel Disease*. *Inflammatory Bowel Diseases*, 2014. **20**(9): p. 1628-1637.
312. Botto, M., et al., *Biosynthesis and secretion of complement component (C3) by activated human polymorphonuclear leukocytes*. *The Journal of Immunology*, 1992. **149**(4): p. 1348-55.
313. Zhou, W., et al., *Role of dendritic cell synthesis of complement in the allospecific T cell response*. *Molecular Immunology*, 2007. **44**(1–3): p. 57-63.
314. Roche, J.K., *Isolation of a purified epithelial cell population from human colon*, in *Colorectal Cancer2001*, Springer. p. 15-20.
315. Weigmann, B., et al., *Isolation and subsequent analysis of murine lamina propria mononuclear cells from colonic tissue*. *Nature protocols*, 2007. **2**(10): p. 2307-2311.
316. Fritzinger, D., et al., *Complement Depletion with Humanized Cobra Venom Factor in a Mouse Model of Age-Related Macular Degeneration*, in *Inflammation and Retinal Disease: Complement Biology and Pathology*, J.D. Lambris and A.P. Adamis, Editors. 2010, Springer New York. p. 151-162.
317. Wang, S.-Y., et al., *Depletion of the C3 component of complement enhances the ability of rituximab-coated target cells to activate human NK cells and improves the efficacy of monoclonal antibody therapy in an in vivo model*. *Blood*, 2009. **114**(26): p. 5322-5330.
318. Angel, C.S., M. Ruzek, and M.K. Hostetter, *Degradation of C3 by Streptococcus pneumoniae*. *The Journal of Infectious Diseases*, 1994. **170**(3): p. 600-608.
319. Laarman, A.J., et al., *Staphylococcus aureus metalloprotease aureolysin cleaves complement C3 to mediate immune evasion*. *The Journal of Immunology*, 2011. **186**(11): p. 6445-6453.
320. Kieslich, C.A. and D. Morikis, *The Two Sides of Complement C3d: Evolution of Electrostatics in a Link between Innate and Adaptive Immunity*. *PLoS Computational Biology*, 2012. **8**(12): p. e1002840.
321. Bradley, A., et al., *The mammalian gene function resource: the international knockout mouse consortium*. *Mammalian Genome*, 2012. **23**(9-10): p. 580-586.
322. Skarnes, W.C., et al., *A conditional knockout resource for the genome-wide study of mouse gene function*. *Nature*, 2011. **474**(7351): p. 337-342.

323. White, Jacqueline K., et al., *Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes*. Cell, 2013. **154**(2): p. 452-464.
324. Lebeis, S.L., et al., *TLR Signaling Mediated by MyD88 Is Required for a Protective Innate Immune Response by Neutrophils to Citrobacter rodentium*. The Journal of Immunology, 2007. **179**(1): p. 566-577.
325. Bonizzi, G. and M. Karin, *The two NF- κ B activation pathways and their role in innate and adaptive immunity*. Trends in Immunology, 2004. **25**(6): p. 280-288.
326. Nijnik, A., et al., *The critical role of histone H2A-deubiquitinase Mysm1 in hematopoiesis and lymphocyte differentiation*. Vol. 119. 2012. 1370-1379.
327. Roman-Garcia, P., et al., *Vitamin B12-dependent taurine synthesis regulates growth and bone mass*. The Journal of Clinical Investigation, 2014. **124**(7): p. 2988-3002.
328. Franchi, L., et al., *NLRC4-driven production of IL-1[beta] discriminates between pathogenic and commensal bacteria and promotes host intestinal defense*. Nature Immunology, 2012. **13**(5): p. 449-456.
329. Pham, Tu Anh N., et al., *Epithelial IL-22RA1-Mediated Fucosylation Promotes Intestinal Colonization Resistance to an Opportunistic Pathogen*. Cell Host & Microbe, 2014. **16**(4): p. 504-516.
330. Sabat, R., *IL-10 family of cytokines*. Cytokine and Growth Factor Reviews, 2010. **21**(5): p. 315-324.
331. Glocker, E.-O., et al., *IL-10 and IL-10 receptor defects in humans*. Annals of the New York Academy of Sciences, 2011. **1246**(1): p. 102-107.
332. Moore, K.W., et al., *Interleukin-10 and the Interleukin-10 receptor*. Annual Review of Immunology, 2001. **19**(1): p. 683-765.
333. Kotlarz, D., et al., *Loss of Interleukin-10 Signaling and Infantile Inflammatory Bowel Disease: Implications for Diagnosis and Therapy*. Gastroenterology, 2012. **143**(2): p. 347-355.
334. Spencer, S.D., et al., *The Orphan Receptor CRF2-4 Is an Essential Subunit of the Interleukin 10 Receptor*. The Journal of Experimental Medicine, 1998. **187**(4): p. 571-578.
335. Kühn, R., et al., *Interleukin-10-deficient mice develop chronic enterocolitis*. Cell, 1993. **75**(2): p. 263-274.
336. Dai, W.J., G. Köhler, and F. Brombacher, *Both innate and acquired immunity to Listeria monocytogenes infection are increased in IL-10-deficient mice*. The Journal of Immunology, 1997. **158**(5): p. 2259-67.
337. Arai, T., et al., *Effects of in vivo administration of anti-IL-10 monoclonal antibody on the host defence mechanism against murine Salmonella infection*. Immunology, 1995. **85**(3): p. 381-388.
338. Dann, S.M., et al., *Attenuation of Intestinal Inflammation in Interleukin-10-Deficient Mice Infected with Citrobacter rodentium*. Infection and Immunity, 2014. **82**(5): p. 1949-1958.

339. Schopf, L.R., et al., *IL-10 Is Critical for Host Resistance and Survival During Gastrointestinal Helminth Infection*. The Journal of Immunology, 2002. **168**(5): p. 2383-2392.
340. Kullberg, M.C., et al., *Induction of colitis by a CD4+ T cell clone specific for a bacterial epitope*. Proceedings of the National Academy of Sciences, 2003. **100**(26): p. 15830-15835.
341. Lokken, K.L., et al., *Malaria Parasite Infection Compromises Control of Concurrent Systemic Non-typhoidal Salmonella Infection via IL-10-Mediated Alteration of Myeloid Cell Function*. PLoS Pathogens, 2014. **10**(5): p. e1004049.
342. Sonnenberg, G.F., et al., *Innate Lymphoid Cells Promote Anatomical Containment of Lymphoid-Resident Commensal Bacteria*. Science, 2012. **336**(6086): p. 1321-1325.
343. Rutz, S., C. Eidenschenk, and W. Ouyang, *IL-22, not simply a Th17 cytokine*. Immunological Reviews, 2013. **252**(1): p. 116-132.
344. Pickard, J.M., et al., *Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness*. Nature, 2014. **514**(7524): p. 638-641.
345. Zheng, Y., et al., *Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens*. Nature Medicine, 2008. **14**(3): p. 282-289.
346. Hasegawa, M., et al., *Interleukin-22 regulates the complement system to promote resistance against pathobionts after pathogen-induced intestinal damage*. Immunity, 2014. **41**(4): p. 620-632.
347. Yamamoto, H. and C. Kemper, *Complement and IL-22: Partnering Up for Border Patrol*. Immunity, 2014. **41**(4): p. 511-513.
348. Wolk, K., et al., *IL-22 regulates the expression of genes responsible for antimicrobial defense, cellular differentiation, and mobility in keratinocytes: a potential role in psoriasis*. European Journal of Immunology, 2006. **36**(5): p. 1309-1323.
349. Zenewicz, L.A. and R.A. Flavell, *Recent advances in IL-22 biology*. International Immunology, 2011. **23**(3): p. 159-163.
350. Sugimoto, K., et al., *IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis*. The Journal of Clinical Investigation, 2008. **118**(2): p. 534-544.
351. Nagalakshmi, M.L., et al., *Interleukin-22 activates STAT3 and induces IL-10 by colon epithelial cells*. International Immunopharmacology, 2004. **4**(5): p. 679-691.
352. Pickert, G., et al., *STAT3 links IL-22 signaling in intestinal epithelial cells to mucosal wound healing*. The Journal of Experimental Medicine, 2009. **206**(7): p. 1465-1472.
353. Sestito, R., et al., *STAT3-dependent effects of IL-22 in human keratinocytes are counterregulated by sirtuin 1 through a direct inhibition of STAT3 acetylation*. The FASEB Journal, 2011. **25**(3): p. 916-927.
354. Laine, V.J., D.S. Grass, and T.J. Nevalainen, *Protection by group II phospholipase A2 against Staphylococcus aureus*. The Journal of Immunology, 1999. **162**(12): p. 7402-7408.

355. Antoniv, T.T. and L.B. Ivashkiv, *Interleukin-10-induced gene expression and suppressive function are selectively modulated by the PI3K-Akt-GSK3 pathway*. Immunology, 2011. **132**(4): p. 567-577.
356. Williams, L., et al., *IL-10 expression profiling in human monocytes*. Journal of Leukocyte Biology, 2002. **72**(4): p. 800-809.
357. Donnelly, R.P., H. Dickensheets, and D.S. Finbloom, *The Interleukin-10 Signal Transduction Pathway and Regulation of Gene Expression in Mononuclear Phagocytes*. Journal of Interferon and Cytokine Research, 1999. **19**(6): p. 563-573.
358. Jung, M., et al., *Expression profiling of IL-10-regulated genes in human monocytes and peripheral blood mononuclear cells from psoriatic patients during IL-10 therapy*. European Journal of Immunology, 2004. **34**(2): p. 481-493.
359. Fiorentino, D.F., et al., *IL-10 acts on the antigen-presenting cell to inhibit cytokine production by Th1 cells*. The Journal of Immunology, 1991. **146**(10): p. 3444-51.
360. Asseman, C., et al., *An Essential Role for Interleukin 10 in the Function of Regulatory T Cells That Inhibit Intestinal Inflammation*. The Journal of Experimental Medicine, 1999. **190**(7): p. 995-1004.
361. Montagutelli, X., *Effect of the Genetic Background on the Phenotype of Mouse Mutations*. Journal of the American Society of Nephrology, 2000. **11**(suppl 2): p. S101-S105.
362. Sato, T., et al., *Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche*. Nature, 2009. **459**(7244): p. 262-265.
363. Elvington, M., et al., *A novel protocol allowing oral delivery of a protein complement inhibitor that subsequently targets to inflamed colon mucosa and ameliorates murine colitis*. Clinical and Experimental Immunology, 2014. **177**(2): p. 500-508.

Appendices

Appendix 1. Reactome pathways significantly associated with transcripts upregulated in wild type *S. Typhimurium* infection.

Pathway name	Pathway uploaded gene count	Genes in InnateDB for this entity	% of pathway genes upregulated	Pathway p-value (corrected)
Immune system	236	884	27	1.14E-22
Hemostasis	121	407	30	9.79E-15
Innate immune system	128	450	28	5.29E-14
Immunoregulatory interactions between a lymphoid and a non-lymphoid cell	30	45	67	1.14E-13
Class A/1 (Rhodopsin-like receptors)	87	277	31	5.42E-12
Cytokine signalling in immune system	68	194	35	9.02E-12
Signalling by GPCR	151	616	25	6.06E-11
GPVI-mediated activation cascade	26	42	62	6.84E-11
Interferon gamma signalling	26	43	60	1.27E-10
GPCR downstream signalling	130	524	25	7.25E-10
Interferon signalling	33	72	46	3.09E-09
Cell surface interactions at the vascular wall	36	83	43	3.16E-09
GPCR ligand binding	99	379	26	9.00E-09
Chemokine receptors bind chemokines	28	58	48	1.69E-08
Signal transduction	287	1490	19	4.18E-08
Platelet activation, signalling and aggregation	57	186	31	1.28E-07
Extracellular matrix organisation	63	216	29	1.69E-07
Peptide ligand-binding receptors	55	181	30	3.17E-07
Adaptive immune system	112	484	23	9.01E-07
G alpha (i) signalling events	60	214	28	1.63E-06
PD-1 signalling	11	14	79	2.48E-06
Generation of second messenger molecules	14	22	64	2.91E-06
Toll-like receptor cascades	41	127	32	3.02E-06
DAPI2 interactions	43	138	31	4.45E-06
Signalling by interleukins	34	100	34	8.42E-06
Phosphorylation of CD3 and TCR zeta chains	9	11	82	1.88E-05
G-protein beta:gamma signalling	15	28	54	2.04E-05
G beta:gamma signalling through PI3Kgamma	14	25	56	2.27E-05
Translocation of ZAP-70 to immunological synapse	8	9	89	2.37E-05
TCR signalling	21	52	40	5.48E-05
Other semaphorin interactions	11	18	61	9.49E-05
Interleukin-2 signalling	17	39	44	1.32E-04
Endosomal/vacuolar pathway	7	8	88	1.32E-04
Interleukin receptor SHC signalling	12	22	55	1.66E-04
Co-stimulation by the CD28 family	19	49	39	2.84E-04
DAPI2 signalling	37	131	28	2.89E-04
Initial triggering of complement	10	17	59	3.60E-04
Latent infection of <i>Homo sapiens</i> with <i>Mycobacterium tuberculosis</i>	14	32	44	6.77E-04
Phagosomal maturation (early endosomal stage)	14	32	44	6.77E-04
FCGR activation	8	12	67	6.83E-04

Cross-presentation of particulate exogenous antigens (phagosomes)	6	7	86	6.98E-04
Gastrin-CREB signalling pathway via PKC and MAPK	45	180	25	9.76E-04
Toll-like receptor 4 (TLR4) cascade	31	109	28	1.03E-03
G alpha (q) signalling events	40	156	26	1.29E-03
Semaphorin interactions	20	59	34	1.36E-03
Transport of inorganic cations/anions and amino acids/oligopeptides	23	75	31	2.35E-03
Degradation of the extracellular matrix	28	100	28	2.78E-03
Axon guidance	59	268	22	3.46E-03
Activation of C3 and C5	5	6	83	3.73E-03
Integrin cell surface interactions	19	59	32	3.98E-03
Collagen degradation	17	51	33	4.92E-03
Complement cascade	12	30	40	5.18E-03
Trafficking and processing of endosomal TLR	7	12	58	5.26E-03
Antigen presentation: folding, assembly and peptide loading of class I MHC	9	19	47	6.05E-03
Interleukin-3, 5 and GM-CSF signalling	14	39	36	6.35E-03
FCERI mediated Ca ²⁺ mobilisation	12	31	39	6.91E-03
Binding and uptake of ligands by scavenger receptors	13	36	36	8.92E-03
Formation of fibrin clot (clotting cascade)	12	32	38	9.45E-03
Basigin interactions	10	24	42	9.64E-03
Platelet homeostasis	14	41	34	1.03E-02
Collagen formation	19	65	29	1.22E-02
Signalling by VEGF	26	100	26	1.23E-02
Amino acid transport across the plasma membrane	11	29	38	1.29E-02
Nucleotide-like (purinergic) receptors	7	14	50	1.49E-02
Antigen processing-cross presentation	21	76	28	1.51E-02
VEGFA-VEGFR2 pathway	24	92	26	1.69E-02
Scavenging by class A receptors	8	18	44	1.73E-02
Fc epsilon receptor (FCERI) signalling	34	146	23	1.81E-02
Calcitonin-like ligand receptors	5	8	63	1.99E-02
Downstream TCR signalling	12	35	34	1.99E-02
Activated TLR4 signalling	25	99	25	2.09E-02
Formyl peptide receptors bind formyl peptides and many other ligands	7	15	47	2.27E-02
Transmembrane transport of small molecules	96	518	19	2.34E-02
MyD88:Mal cascade initiated on plasma membrane	21	81	26	3.05E-02
Toll-like receptor 2 (TLR2) cascade	21	81	26	3.05E-02
Toll-like receptor TLR1:TLR2 cascade	21	81	26	3.05E-02
Toll-like receptor TLR6:TLR2 cascade	21	81	26	3.05E-02
Activation of matrix metalloproteinases	12	37	32	3.13E-02
Activation, myristoylation of BID and translocation to mitochondria	5	9	56	3.49E-02
Dissolution of fibrin clot	5	9	56	3.49E-02
Removal of amino-terminal pro-peptides from gamma-carboxylated proteins	5	9	56	3.49E-02
Signal regulatory protein (SIRP) family interactions	5	9	56	3.49E-02
Antigen activates B Cell Receptor (BCR) leading to generation of second messengers	11	33	33	3.55E-02
Classical antibody-mediated complement activation	4	6	67	3.78E-02

Death receptor signalling	6	13	46	4.25E-02
Eicosanoid ligand-binding receptors	6	13	46	4.25E-02
Extrinsic pathway for apoptosis	6	13	46	4.25E-02
Platelet adhesion to exposed collagen	6	13	46	4.25E-02
Synthesis of leukotrienes (LT) and eoxins (EX)	6	13	46	4.25E-02
G alpha (s) signalling events	22	89	25	4.29E-02
Signalling by PDGF	32	145	22	4.65E-02
Signalling by SCF-KIT	26	112	23	4.87E-02

Appendix 2. Reactome pathways significantly associated with transcripts downregulated in wild type *S. Typhimurium* infection.

Pathway name	Pathway uploaded gene count	Genes in InnateDB for this entity	% of pathway genes downregulated	Pathway p-value (corrected)
Metabolism	275	1414	19	3.19E-27
Biological oxidations	57	146	39	5.49E-18
Phase II conjugation	28	63	44	2.19E-10
Transmembrane transport of small molecules	99	518	19	2.64E-08
Phase 1 - Functionalization of compounds	26	80	33	2.49E-06
Mitochondrial fatty acid beta-oxidation	10	15	67	1.08E-05
Metabolism of lipids and lipoproteins	90	519	17	1.44E-05
Degradation of cysteine and homocysteine	7	8	88	2.98E-05
Ethanol oxidation	7	8	88	2.98E-05
Transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds	25	86	29	3.33E-05
The citric acid (TCA) cycle and respiratory electron transport	34	139	24	3.55E-05
Nuclear receptor transcription pathway	16	46	35	1.89E-04
Cytosolic sulfonation of small molecules	8	13	62	2.44E-04
Abacavir transport and metabolism	7	10	70	2.62E-04
Abacavir transmembrane transport	5	5	100	2.66E-04
Regulation of beta-cell development	11	25	44	3.22E-04
Sulfur amino acid metabolism	11	25	44	3.22E-04
SLC-mediated transmembrane transport	44	223	20	3.30E-04
Metabolism of amino acids and derivatives	38	183	21	3.68E-04
Stimuli-sensing channels	23	88	26	3.74E-04
Organic cation transport	6	8	75	5.19E-04
Respiratory electron transport	21	79	27	6.04E-04
Glutathione conjugation	11	27	41	6.41E-04
Fatty acid, triacylglycerol, and ketone body metabolism	34	165	21	9.83E-04
Reversible hydration of carbon dioxide	7	12	58	9.95E-04
Glucuronidation	5	6	83	1.02E-03

Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.	24	101	24	1.08E-03
Regulation of gene expression in endocrine-committed (NEUROG3+) progenitor cells	4	4	100	1.59E-03
Synthesis of ketone bodies	4	4	100	1.59E-03
Incretin synthesis, secretion, and inactivation	9	21	43	1.65E-03
Na ⁺ /Cl ⁻ dependent neurotransmitter transporters	8	17	47	1.77E-03
Xenobiotics	10	26	38	1.95E-03
Cytochrome P450 - arranged by substrate type	17	64	27	2.36E-03
Ion channel transport	28	137	20	3.56E-03
Synthesis, secretion, and inactivation of Glucagon-like Peptide-1 (GLP-1)	8	19	42	3.89E-03
Class A/1 (Rhodopsin-like receptors)	47	277	17	4.51E-03
Ketone body metabolism	4	5	80	5.69E-03
Propionyl-CoA catabolism	4	5	80	5.69E-03
Sulfide oxidation to sulfate	4	5	80	5.69E-03
Organic cation/anion/zwitterion transport	6	12	50	6.46E-03
Bile acid and bile salt metabolism	11	36	31	7.11E-03
Pyruvate metabolism and Citric Acid (TCA) cycle	12	42	29	7.77E-03
ABC-family proteins mediated transport	9	26	35	7.82E-03
Beta oxidation of butanoyl-CoA to acetyl-CoA	3	3	100	1.00E-02
GRB7 events in ERBB2 signalling	3	3	100	1.00E-02
Mitochondrial fatty acid beta-oxidation of unsaturated fatty acids	3	3	100	1.00E-02
GPCR ligand binding	58	379	15	1.16E-02
Erythrocytes take up carbon dioxide and release oxygen	4	6	67	1.22E-02
O ₂ /CO ₂ exchange in erythrocytes	4	6	67	1.22E-02
Transport and synthesis of PAPS	4	6	67	1.22E-02
Regulation of gene expression in beta cells	6	14	43	1.46E-02
Retinoid metabolism and transport	11	40	28	1.55E-02
Defective AMN causes hereditary megaloblastic anemia 1	16	72	22	1.58E-02
Defective BTM causes biotinidase deficiency	16	72	22	1.58E-02
Defective CD320 causes methylmalonic aciduria	16	72	22	1.58E-02
Defective CUBN causes hereditary megaloblastic anemia 1	16	72	22	1.58E-02
Defective GIF causes intrinsic factor deficiency	16	72	22	1.58E-02
Defective HLCS causes multiple carboxylase deficiency	16	72	22	1.58E-02
Defective LMBRD1 causes methylmalonic aciduria and homocystinuria type cblF	16	72	22	1.58E-02
Defective MMAA causes methylmalonic aciduria type cblA	16	72	22	1.58E-02
Defective MMAB causes methylmalonic aciduria type cblB	16	72	22	1.58E-02

Defective MMACHC causes methylmalonic aciduria and homocystinuria type cblC	16	72	22	1.58E-02
Defective MMADHC causes methylmalonic aciduria and homocystinuria type cblD	16	72	22	1.58E-02
Defective MTR causes methylmalonic aciduria and homocystinuria type cblG	16	72	22	1.58E-02
Defective MTRR causes methylmalonic aciduria and homocystinuria type cblE	16	72	22	1.58E-02
Defective MUT causes methylmalonic aciduria mut type	16	72	22	1.58E-02
Defective TCN2 causes hereditary megaloblastic anemia	16	72	22	1.58E-02
Defects in biotin (B7) metabolism	16	72	22	1.58E-02
Defects in cobalamin (B12) metabolism	16	72	22	1.58E-02
Defects in vitamin and cofactor metabolism	16	72	22	1.58E-02
Metabolism of vitamins and cofactors	16	72	22	1.58E-02
Metabolism of water-soluble vitamins and cofactors	16	72	22	1.58E-02
Branched-chain amino acid catabolism	6	15	40	1.77E-02
Sialic acid metabolism	9	31	29	2.09E-02
Peptide ligand-binding receptors	31	181	17	2.09E-02
Synthesis, secretion, and inactivation of Glucose-dependent Insulinotropic Polypeptide (GIP)	5	11	45	2.10E-02
Synthesis of bile acids and bile salts	8	26	31	2.34E-02
Aflatoxin activation and detoxification	6	16	38	2.48E-02
O-linked glycosylation of mucins	10	38	26	2.67E-02
Alpha-linolenic (omega3) and linoleic (omega6) acid metabolism	5	12	42	3.07E-02
Alpha-linolenic acid (ALA) metabolism	5	12	42	3.07E-02
Peroxisomal lipid metabolism	7	22	32	3.15E-02
Synthesis of bile acids and bile salts via 7alpha-hydroxycholesterol	7	22	32	3.15E-02
Mitochondrial fatty acid beta-oxidation of saturated fatty acids	4	8	50	3.39E-02
Pyruvate metabolism	7	23	30	4.07E-02
G alpha (i) signalling events	34	214	16	4.12E-02
Inositol phosphate metabolism	10	41	24	4.42E-02

Appendix 3. GO terms significantly associated with transcripts regulated in wild type *S. Typhimurium* infection.

Pathway name	Pathway uploaded gene count	Genes in InnateDB for this entity	Pathway p-value (corrected)
--------------	-----------------------------	-----------------------------------	-----------------------------

Upregulated

Biological process

innate immune response	302	718	8.28E-86
inflammatory response	129	249	3.21E-47
immune response	110	237	1.79E-34
response to lipopolysaccharide	71	162	5.72E-20
defence response to Gram-positive bacterium	36	54	2.61E-17
chemotaxis	40	77	5.70E-14
cellular response to lipopolysaccharide	44	93	1.65E-13
cell surface receptor signalling pathway	54	134	4.15E-13
neutrophil chemotaxis	27	41	1.40E-12
cell adhesion	95	329	1.46E-12

Molecular function

protein binding	1433	9494	2.15E-18
transmembrane signalling receptor activity	44	105	2.51E-11
carbohydrate binding	65	203	1.21E-10
receptor binding	81	288	3.75E-10
cytokine activity	56	171	1.36E-09
cytokine receptor activity	18	28	4.14E-08
calcium ion binding	137	642	5.45E-08
integrin binding	31	76	1.20E-07
cysteine-type endopeptidase inhibitor activity	20	40	1.43E-06
chemokine activity	24	55	1.50E-06

Cellular component

external side of plasma membrane	120	237	1.50E-42
extracellular space	298	1104	1.89E-36
extracellular region	282	1217	1.86E-22
cell surface	128	409	1.77E-20
extracellular matrix	68	214	6.07E-11
membrane raft	57	170	3.50E-10
integral component of plasma membrane	125	546	3.34E-09
proteinaceous extracellular matrix	61	217	1.32E-07
plasma membrane	581	3699	2.14E-07
neuronal cell body	74	302	1.45E-06

Downregulated

Biological process

metabolic process	230	1034	3.60E-31
oxidation-reduction process	175	770	2.56E-24
lipid glycosylation	21	28	2.49E-13
transmembrane transport	105	513	5.39E-11

steroid hormone mediated signalling pathway	20	57	4.80E-05
fatty acid biosynthetic process	18	51	1.62E-04
drug transmembrane transport	10	17	1.77E-04
sodium ion transport	27	103	2.65E-04
acyl-CoA metabolic process	11	22	3.75E-04
fatty acid metabolic process	20	65	3.84E-04
response to starvation	14	35	3.93E-04

Molecular function

oxidoreductase activity	90	327	3.21E-17
catalytic activity	114	646	8.54E-08
heme binding	45	181	1.09E-06
acyl-CoA dehydrogenase activity	13	20	1.14E-06
transferase activity, transferring hexosyl groups	22	57	2.06E-06
carboxylic ester hydrolase activity	16	32	2.67E-06
oxidoreductase activity, acting on the CH-CH group of donors	14	25	3.24E-06
glutathione transferase activity	15	29	3.84E-06
steroid hormone receptor activity	20	55	2.56E-05
iron ion binding	45	209	6.09E-05

Cellular component

mitochondrion	247	1510	3.20E-14
mitochondrial inner membrane	80	306	4.93E-14
apical plasma membrane	62	231	4.26E-11
peroxisome	36	111	1.84E-08
mitochondrial matrix	39	147	1.49E-06
brush border membrane	15	45	1.75E-03
extracellular vesicular exosome	257	2080	2.08E-03
endoplasmic reticulum membrane	88	584	2.43E-03
basolateral plasma membrane	31	151	4.96E-03
brush border	11	30	6.58E-03

Appendix 4. Reactome pathways significantly associated with genes upregulated at the RNA and protein level in wild type *S. Typhimurium* infection.

Pathway name	Upregulated gene count	% of pathway genes upregulated	Pathway p-value (corrected)	Gene Symbols
Regulation of complement cascade	4	22	3.77E-05	C3, C4b, Cfb, Cfh
Activation of C3 and C5	3	50	4.12E-05	C3, C4b, Cfb
Endosomal/vacuolar pathway	3	38	9.17E-05	B2m, Ctss, H2-K1
Complement cascade	4	13	1.62E-04	C3, C4b, Cfb, Cfh
Interferon signalling	5	7	2.42E-04	B2m, Gbp2, H2-K1, Isg15, Ptpn6
Antigen processing-cross presentation	5	7	2.52E-04	B2m, Ctss, H2-K1, Psmb8, Tapbp
Interferon gamma signalling	4	9	3.79E-04	B2m, Gbp2, H2-K1, Ptpn6
Initial triggering of complement	3	18	4.49E-04	C3, C4b, Cfb

Antigen presentation: folding, assembly and peptide loading of class I MHC	3	16	5.87E-04	B2m, H2-K1, Tapbp
Alternative complement activation	2	50	7.42E-04	C3, Cfb
Immune system	13	1	1.26E-03	B2m, C3, C4b, Cd74, Cfb, Cfh, Ctss, Gbp2, H2-K1, Isg15, Psmb8, Ptpn6, Tapbp
ER-phagosome pathway	4	6	1.32E-03	B2m, H2-K1, Psmb8, Tapbp
Amyloids	4	5	3.12E-03	B2m, Ltf, Lyz2, Tgfb1
Binding and uptake of ligands by scavenger receptors	3	8	3.12E-03	Apoe, Fth1, Hp
Iron uptake and transport	3	8	3.81E-03	Cp, Fth1, Lcn2
Immunoregulatory interactions between a lymphoid and a non-lymphoid cell	3	7	4.86E-03	B2m, C3, H2-K1
HDL-mediated lipid transport	2	14	6.93E-03	A2m, Apoe
Adaptive immune system	8	2	7.58E-03	B2m, C3, Cd74, Ctss, H2-K1, Psmb8, Ptpn6, Tapbp
Cytokine signaling in immune system	5	3	7.97E-03	B2m, Gbp2, H2-K1, Isg15, Ptpn6
Class I MHC mediated antigen processing & presentation	5	2	8.95E-03	B2m, Ctss, H2-K1, Psmb8, Tapbp
Scavenging by class A receptors	2	11	9.38E-03	Apoe, Fth1
Intrinsic pathway	2	10	1.04E-02	A2m, Kng1
Extracellular matrix organization	5	2	1.05E-02	A2m, Col18a1, Ctss, Itgad, Serpinh1
Collagen formation	3	5	1.05E-02	Col18a1, Ctss, Serpinh1
Platelet degranulation	3	4	1.36E-02	A2m, Kng1, Psap
Lipoprotein metabolism	2	8	1.40E-02	A2m, Apoe
Response to elevated platelet cytosolic Ca ²⁺	3	4	1.48E-02	A2m, Kng1, Psap
Formation of fibrin clot (clotting cascade)	2	6	2.31E-02	A2m, Kng1
Peptide ligand-binding receptors	4	2	2.53E-02	Anxa1, C3, Kng1, Psap
Degradation of the extracellular matrix	3	3	2.67E-02	A2m, Col18a1, Ctss
Assembly of collagen fibrils and other multimeric structures	2	5	3.17E-02	Col18a1, Ctss
Lipid digestion, mobilization, and transport	2	5	3.17E-02	A2m, Apoe
Collagen biosynthesis and modifying enzymes	2	4	4.04E-02	Col18a1, Serpinh1
Innate immune system	6	1	4.10E-02	B2m, C3, C4b, Cfb, Cfh, Ctss

Appendix 5. Reactome pathways significantly associated with genes downregulated at the RNA and protein level in wild type *S. Typhimurium* infection.

Pathway name	Down-regulated gene count	% of pathway genes upregulated	Pathway p-value (corrected)	Gene Symbols
Metabolism	50	4	3.08E-15	Abcb1a, Acads, Acat1, Acox1, Ahcy, Akr1c13, Aldh1a1, Aldh2, Atp5a1, Bdh1, Bsg, Cbr1, Cmpk1, Cox4i1, Cox6c, Cth, Cyc1, Cyp2c55, Echs1, Eci1, Eno3, Etfa, Etfb, Ethel, Gna11, Gstm1, Gstm2, Gstm3, Hadh, Hmgcs2, Hsd3b3, Maa, Mdh1, Mgst3, Papss2, Pfk1, Sdha, Slc16a1, Slc25a10, Sqrdl, Suclg2, Sult1b1, Sult1d1, Tst, Ugdh, Ugt1a1, Ugt1a7c, Uqcr10, Uqcr1, Uqcr2
Biological oxidations	15	10	3.50E-10	Ahcy, Aldh1a1, Aldh2, Cyp2c55, Gstm1, Gstm2, Gstm3, Maa, Mgst3, Papss2, Sult1b1, Sult1d1, Ugdh, Ugt1a1, Ugt1a7c
Phase II conjugation	11	17	5.10E-10	Ahcy, Gstm1, Gstm2, Gstm3, Mgst3, Papss2, Sult1b1, Sult1d1, Ugdh, Ugt1a1, Ugt1a7c
The citric acid (TCA) cycle and respiratory electron transport	13	9	1.83E-08	Atp5a1, Bsg, Cox4i1, Cox6c, Cyc1, Etfa, Etfb, Sdha, Slc16a1, Suclg2, Uqcr10, Uqcr1, Uqcr2
Degradation of cysteine and homocysteine	5	63	5.69E-08	Cth, Ethel, Slc25a10, Sqrdl, Tst
Sulfide oxidation to sulfate	4	80	4.51E-07	Ethel, Slc25a10, Sqrdl, Tst
Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.	10	10	5.77E-07	Atp5a1, Cox4i1, Cox6c, Cyc1, Etfa, Etfb, Sdha, Uqcr10, Uqcr1, Uqcr2
Respiratory electron transport	9	11	7.42E-07	Cox4i1, Cox6c, Cyc1, Etfa, Etfb, Sdha, Uqcr10, Uqcr1, Uqcr2
Sulfur amino acid metabolism	6	24	9.91E-07	Ahcy, Cth, Ethel, Slc25a10, Sqrdl, Tst
Beta oxidation of butanoyl-CoA to acetyl-CoA	3	100	5.98E-06	Acads, Echs1, Hadh
Synthesis of ketone bodies	3	75	2.03E-05	Acat1, Bdh1, Hmgcs2
Beta oxidation of hexanoyl-CoA to butanoyl-CoA	3	60	4.56E-05	Acads, Echs1, Hadh
Ketone body metabolism	3	60	4.56E-05	Acat1, Bdh1, Hmgcs2
Mitochondrial Fatty Acid Beta-Oxidation	4	27	5.51E-05	Acads, Echs1, Eci1, Hadh,

Glucuronidation	3	50	8.25E-05	Ugdh, Ugt1a1, Ugt1a7c
Fatty acid, triacylglycerol, and ketone body metabolism	9	5	1.87E-04	Acads, Acat1, Acox1, Bdh1, Echs1, Eci1, Hadh, Hmgcs2, Ugt1a7c
Mitochondrial fatty acid beta-oxidation of saturated fatty acids	3	38	2.04E-04	Acads, Echs1, Hadh
Metabolism of serotonin	2	100	3.42E-04	Aldh2, Maoa
Serotonin clearance from the synaptic cleft	2	100	3.42E-04	Aldh2, Maoa
Glutathione conjugation	4	15	4.94E-04	Gstm1, Gstm2, Gstm3, Mgst3
Cytosolic sulfonation of small molecules	3	23	8.45E-04	Papss2, Sult1b1, Sult1d1
Utilization of ketone bodies	2	67	9.07E-04	Acat1, Bdh1
Pyruvate metabolism and Citric Acid (TCA) cycle	4	10	2.45E-03	Bsg, Sdha, Slc16a1, Suclg2
Beta oxidation of decanoyl-CoA to octanoyl-CoA-CoA	2	40	2.64E-03	Echs1, Hadh
Beta oxidation of lauroyl-CoA to decanoyl-CoA-CoA	2	40	2.64E-03	Echs1, Hadh
Beta oxidation of octanoyl-CoA to hexanoyl-CoA	2	40	2.64E-03	Echs1, Hadh
Neurotransmitter clearance in the synaptic cleft	2	33	3.86E-03	Aldh2, Maoa
Basigin interactions	3	13	4.34E-03	Atp1b1, Bsg, Slc16a1
Ethanol oxidation	2	25	6.07E-03	Aldh1a1, Aldh2
Metabolism of amino acids and derivatives	7	4	6.14E-03	Acat1, Ahcy, Cth, Ethel, Slc25a10, Sqrdl, Tst
Gluconeogenesis	3	10	7.42E-03	Eno3, Mdh1, Slc25a10
Glucose metabolism	4	6	8.90E-03	Eno3, Mdh1, Pfkf, Slc25a10
Metabolism of lipids and lipoproteins	12	2	1.43E-02	Acads, Acat1, Acox1, Akr1c13, Bdh1, Cbr1, Echs1, Eci1, Hadh, Hmgcs2, Hsd3b3, Ugt1a7c
Phase I - Functionalization of compounds	4	5	1.79E-02	Aldh1a1, Aldh2, Cyp2c55, Maoa
Cell surface interactions at the vascular wall	4	5	2.01E-02	Atp1b1, Bsg, Ceacam1, Slc16a1
Synthesis of Prostaglandins (PG) and Thromboxanes (TX)	2	13	2.12E-02	Akr1c13, Cbr1
Citric acid cycle (TCA cycle)	2	12	2.29E-02	Sdha, Suclg2
Thrombin signalling through proteinase activated receptors (PARs)	2	12	2.29E-02	Gna11, Mapk3
Pyruvate metabolism	2	9	3.89E-02	Bsg, Slc16a1
Glycolysis	2	8	4.49E-02	Eno3, Pfkf

Appendix 6. Differentially expressed genes in caecal tissue from mutant versus wild type mice, as determined by RNAseq. Changes with an adjusted p-value < 0.05 were deemed statistically significant. Up to 50 genes with the largest increase and decrease in expression are listed (log2 fold change < -1 or > 1). Transcript abundance is the baseMean output from DESeq2 (the average of the normalised count values over all samples).

***IL10rb*^{tm1a/tm1a} (S. Typhimurium-infected)**

Gene symbol	Transcript abundance	log2 Fold change	Adjusted p-value
Upregulated			
EN2	21	2.02	4.65E-07
ERMN	111	1.95	3.75E-07
CD40LG	26	1.89	2.55E-07
SEMA3E	209	1.86	2.27E-08
HPX	56	1.71	9.92E-05
NLRP10	444	1.70	3.26E-05
TMPRSS11G	14	1.62	2.62E-04
CD200R2	39	1.61	4.13E-07
2210407C18RIK	9192	1.60	2.71E-14
ATP12A	81	1.56	5.44E-04
GAL3ST2	404	1.50	3.58E-05
HAPLN1	41	1.49	1.32E-03
RASSF10	71	1.47	1.18E-07
NMU	4	1.46	1.88E-03
IGLC2	754	1.46	2.31E-03
TRGJ1	24	1.45	2.00E-04
CYP2E1	17	1.43	2.52E-03
AQP3	27	1.42	1.03E-03
ST6GALNAC1	4	1.41	3.53E-03
TRBV15	19	1.40	2.09E-03
TFF1	4	1.39	3.45E-03
FGG	18	1.39	6.29E-04
LAT	438	1.38	1.54E-04
EAR1	11	1.37	4.06E-03
IGLV2	685	1.35	6.58E-03
SERPINA10	248	1.34	3.58E-05
ITIH4	14	1.33	8.07E-03
MCPT9	12	1.32	8.11E-03
ACTN2	58	1.31	3.00E-03
1700008I05RIK	7	1.31	1.02E-02
SERPINA1E	8	1.30	1.03E-02
KNDC1	17	1.29	7.58E-04
2010109A12RIK	7	1.29	6.38E-03

IL2RA	1628	1.29	1.06E-03
GM9994	264	1.28	3.64E-03
SLC17A4	1230	1.27	2.52E-03
TRGV2	91	1.27	2.66E-04
IGHG2B	867	1.26	1.49E-02
PMP22	8838	1.26	1.78E-06
9430041J12RIK	174	1.25	1.54E-03
TMPRSS3	2	1.24	1.61E-02
UBD	23136	1.24	5.05E-05
TCRG-V6	44	1.23	1.94E-02
TCRG-C1	237	1.23	3.66E-03
BLK	131	1.23	8.95E-03
UGT8A	39	1.21	2.42E-02
H19	79	1.21	2.49E-02
5830411N06RIK	65	1.21	3.15E-03
SERPINA1D	4	1.19	3.02E-02
IGKV4-68	433	1.19	2.94E-02

Downregulated

IL10RB	2873	-4.32	1.05E-243
PRSS27	178	-2.97	1.56E-19
GNA14	635	-2.93	8.03E-21
KRT36	166	-2.67	7.00E-35
GATA4	111	-2.57	4.75E-12
REG3G	2851	-2.41	1.06E-10
AQP4	891	-2.24	2.27E-08
B3GALT5	4551	-2.22	3.15E-10
MAL	478	-2.15	8.06E-10
BCL2L15	468	-2.11	3.53E-08
CHST4	464	-1.93	9.29E-07
PLA2G2A	158	-1.91	5.80E-06
BMP3	111	-1.86	2.88E-07
GM8540	58	-1.82	1.76E-05
PLA2G5	260	-1.72	1.16E-07
SP5	36	-1.71	9.36E-08
SEC1	158	-1.67	9.29E-07
M5C1000I18RIK	102	-1.65	3.49E-08
GM5689	47	-1.62	6.80E-05
CNKSR2	12	-1.62	1.89E-05
C2CD4B	149	-1.54	9.29E-07
PPEF1	48	-1.54	7.07E-04
FUT2	6901	-1.54	1.70E-05
9130204K15RIK	32	-1.53	1.06E-03
GM21860	21	-1.52	3.02E-04
CSTA	381	-1.49	5.60E-04

SLC5A9	402	-1.47	1.88E-04
EPGN	31	-1.46	1.65E-03
GM9458	64	-1.46	2.15E-03
Gm21742	56	-1.43	1.71E-03
SPRR2D	12	-1.43	1.53E-03
REG3B	4127	-1.38	4.23E-04
SAA1	2840	-1.36	2.12E-03
RDH18-PS	15	-1.35	1.97E-03
PLA2G12B	217	-1.35	1.46E-03
GGT1	461	-1.33	1.65E-03
SLC7A9	116	-1.33	8.07E-03
TAT	2330	-1.32	4.43E-04
GM26917	3027	-1.32	8.13E-03
GM21970	1905	-1.32	2.96E-16
LYPD3	27	-1.32	3.40E-03
GM21748	23	-1.30	1.09E-02
SPRR2F	33	-1.28	1.35E-02
1700007E05Rik	17	-1.27	6.20E-03
ST8SIA1	118	-1.27	3.29E-03
HOXD11	85	-1.26	1.49E-02
DMBT1	28655	-1.26	1.61E-02
PLA2G2E	67	-1.26	9.75E-03
2210011K15Rik	5	-1.25	1.49E-02
RDH16	466	-1.25	1.03E-02

***IL22ra1*^{tm1a/tm1a} (*S. Typhimurium*-infected)**

Gene symbol	Transcript abundance	log2 Fold change	Adjusted p-value
Upregulated			
CCL20	108	-3.03	1.49E-37
PRSS27	178	-2.55	1.97E-14
GNA14	635	-2.50	4.08E-15
PLA2G2A	158	-2.44	1.39E-10
REG3G	2851	-2.37	1.07E-10
PPEF1	48	-2.30	1.35E-09
KRT36	166	-2.25	7.91E-25
PLA2G5	260	-2.22	1.10E-13
GATA4	111	-2.22	2.58E-09
GM8540	58	-2.18	9.82E-09
SLC5A9	402	-1.92	1.82E-08
AQP4	891	-1.86	3.32E-06
B3GALT5	4551	-1.76	9.59E-07
GM2539	10	-1.75	2.01E-06

SLC1A1	211	-1.69	2.30E-06
CHI3L3	13158	-1.68	4.22E-05
ST8SIA1	118	-1.65	4.26E-06
CHI3L4	439	-1.64	6.54E-05
BCL2L15	468	-1.64	2.78E-05
ADCY2	212	-1.58	3.59E-07
CCIN	6	-1.56	1.50E-04
PLAT	1177	-1.56	5.96E-09
SEC1	158	-1.55	2.68E-06
SPRR2F	33	-1.52	2.89E-04
AOX3	21	-1.52	6.59E-05
CNKSR2	12	-1.52	2.14E-05
PLA2G2E	67	-1.51	1.51E-04
SLC12A3	10	-1.51	1.36E-04
MAL	478	-1.51	3.67E-05
GM5689	47	-1.51	8.74E-05
RS1	13	-1.47	3.64E-04
GM15726	78	-1.46	4.93E-04
GM9458	64	-1.45	5.84E-04
CSTA	381	-1.45	2.37E-04
MT3	64	-1.44	6.77E-04
NTN5	7	-1.43	4.74E-04
KCNA4	8	-1.43	8.28E-04
MT2	5723	-1.43	4.74E-04
EPGN	31	-1.42	6.21E-04
DMGDH	14	-1.42	9.14E-04
NPTX2	170	-1.42	3.18E-05
TMEM119	3443	-1.41	4.26E-04
MEG3_2	5	-1.41	9.38E-04
CHST4	464	-1.40	5.92E-04
MT1	4635	-1.40	2.09E-04
1500009C09RIK	92	-1.40	5.20E-05
ESM1	82	-1.40	4.74E-04
AI747448	53150	-1.38	1.97E-04
REG3B	4127	-1.37	1.13E-04
PLAGL1	402	-1.36	2.95E-04

Downregulated

SLC17A1	24	2.81	2.77E-16
SLC17A4	1230	2.80	1.29E-18
GAL3ST2	404	2.78	3.46E-20
SLC34A2	39	2.59	1.12E-13
GM9994	264	2.48	1.93E-13
ATP12A	81	2.34	5.23E-10
DIO1	161	2.32	1.33E-11

FABP2	3356	2.28	8.19E-10
EN2	21	2.27	1.53E-09
CPN1	251	2.20	9.50E-09
RBP2	315	2.11	3.10E-09
PMP22	8838	2.08	2.44E-19
PBP2	43	2.03	2.87E-08
2210407C18RIK	9192	2.01	2.48E-23
CNPY1	75	2.00	1.50E-10
NLRP10	444	1.98	8.47E-08
UPK1A	147	1.87	1.80E-08
MYO3A	29	1.87	1.02E-10
BCAS1	290	1.86	2.49E-06
TGM5	77	1.85	7.37E-08
IL1F8	42	1.83	5.36E-07
GM15368	18	1.81	2.26E-07
CYP2C55	152	1.81	3.36E-06
GM14259	50	1.80	7.09E-06
CLCN1	51	1.77	3.21E-09
SH3D21	415	1.76	1.96E-12
UPB1	32	1.71	3.73E-06
FAM189A2	986	1.70	2.67E-07
KCNU1	14	1.69	5.42E-07
SYT7	637	1.68	2.54E-11
HSD3B2	47	1.66	4.79E-05
9430041J12RIK	174	1.65	6.58E-07
REP15	81	1.64	1.57E-05
PPFIA3	1306	1.64	2.88E-09
TRIM58	11	1.64	4.99E-05
ABCB1A	2050	1.60	3.21E-09
ANG4	154	1.58	1.48E-04
1810065E05RIK	4819	1.56	1.77E-04
SLC9A2	4482	1.56	4.23E-06
GM13412	9	1.55	4.43E-05
GM20478	38	1.54	2.26E-06
SYTL2	738	1.54	1.50E-07
GM23800	7	1.53	1.25E-04
LGALS2	198	1.53	2.57E-04
MS4A18	256	1.53	2.84E-05
RAPGEFL1	407	1.52	4.87E-07
HSD11B2	734	1.52	2.68E-05
XPNPEP2	177	1.51	2.58E-08
BTNL7-PS	190	1.51	1.98E-06
GM26859	34	1.51	3.34E-05

***IL22ra1*^{tm1a/tm1a} (naïve)**

Gene symbol	Transcript abundance	log2 Fold change	Adjusted p-value
Upregulated			
RPRM	67	1.00	4.85E-03
AOC1	15530	1.01	3.58E-06
TFRC	26067	1.02	7.53E-04
CYP2E1	103	1.05	6.10E-04
GM26697	555	1.07	1.08E-04
B3GALT2	227	1.10	4.23E-06
SLC17A2	281	1.11	1.44E-04
IGKV5-48	989	1.25	2.06E-04
GRIN3A	118	1.26	1.79E-04
IGHV7-1	249	1.30	9.21E-05
MT2	6292	1.34	5.68E-14
CES2A	14332	1.38	2.03E-05
CYP4B1-PS1	78	1.48	2.21E-06
CYP4B1	5572	1.74	2.35E-10
CYP4B1-PS2	44	1.76	1.23E-10
Downregulated			
GM15726	98	-1.89	2.28E-10
MID1	886	-1.83	1.23E-10
MS4A10	677	-1.75	2.35E-11
GM21857	221	-1.62	6.32E-09
IL22RA1	1036	-1.61	1.59E-08
CYP2D9	629	-1.31	6.83E-06
PRSS16	75	-1.26	5.89E-11
RHOX5	67	-1.24	4.34E-05
MYOM3	52	-1.20	8.60E-05
GM8540	129	-1.19	5.47E-04
IGFBP2	1546	-1.18	4.13E-04
GM21742	73	-1.18	5.57E-04
3930402G23RIK	83	-1.18	1.56E-04
PCSK9	683	-1.18	1.79E-05
REG3B	52	-1.17	1.44E-04
MMP10	52	-1.15	9.47E-04
ASB11	115	-1.13	1.49E-04
EREG	747	-1.11	5.51E-04
FAM167B	40	-1.09	9.30E-04
ETV4	150	-1.02	5.06E-04