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 Δ*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 Δ*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 Δ*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 (~ 10^8 - 10^9 CFU/g tissue) and high to moderate infection in the liver and spleen (~ 10^5 - 10^6 CFU/g tissue). The small intestine was relatively poorly colonised (~ 10^4 CFU/g). The main *Salmonella*-containing tissues in the intestinal tract showed significantly lower levels of colonisation by the SL1344 Δ*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).

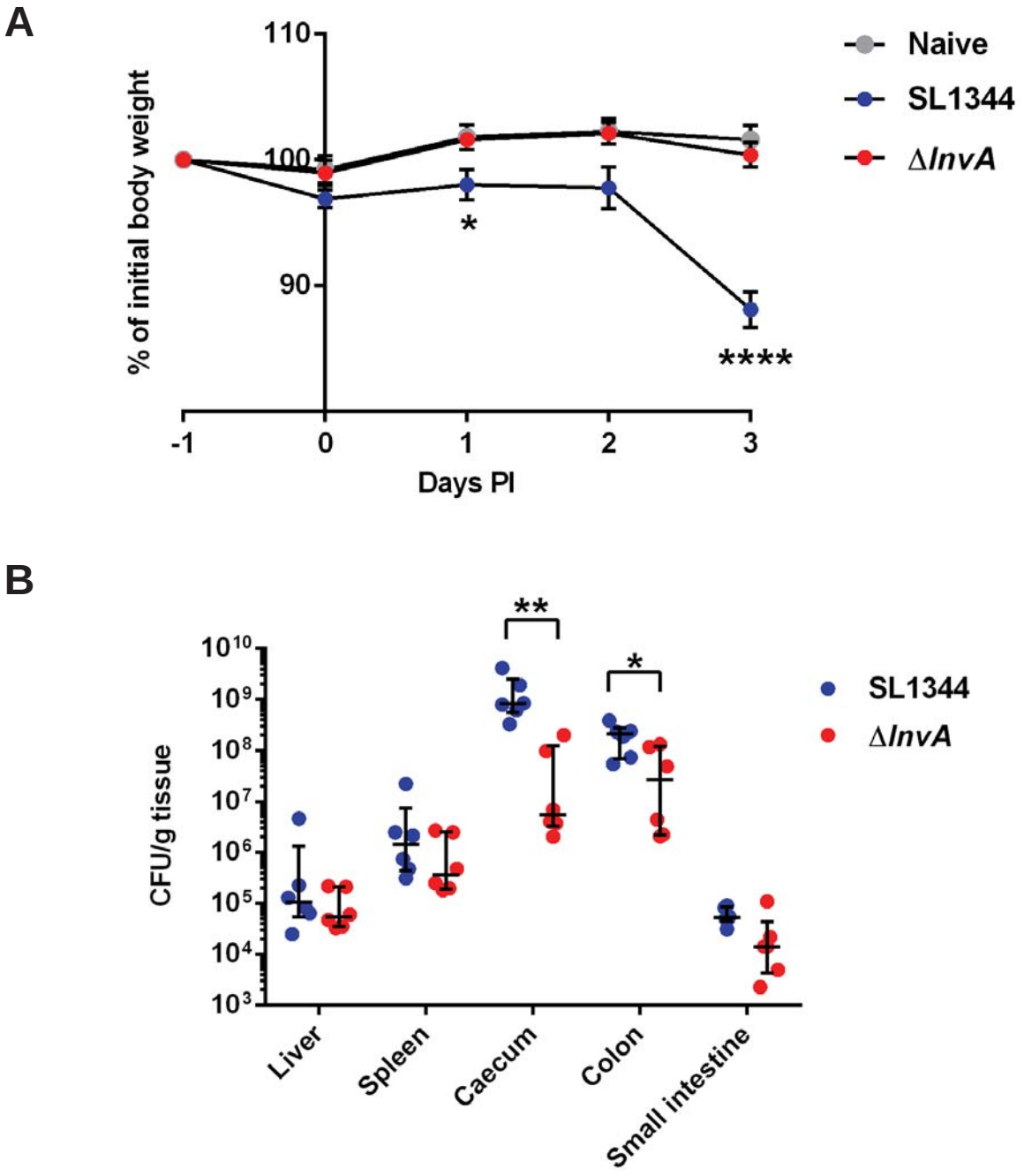


Figure 3.1. Weight loss and tissue colonisation in the streptomycin mouse model. (A) Representative weight curve of C57BL/6N mice treated with streptomycin prior to oral delivery of PBS (naïve controls), wild type *S.* Typhimurium SL1344 (9 x 10^4 CFU), or *S.* Typhimurium SL1344 Δ*InvA* (1.6 x 10^5 CFU). Data shown are mean ± 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 Δ*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 Δ*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 μm 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 Δ*InvA* were highly similar in appearance to those infected with wild type SL1344.

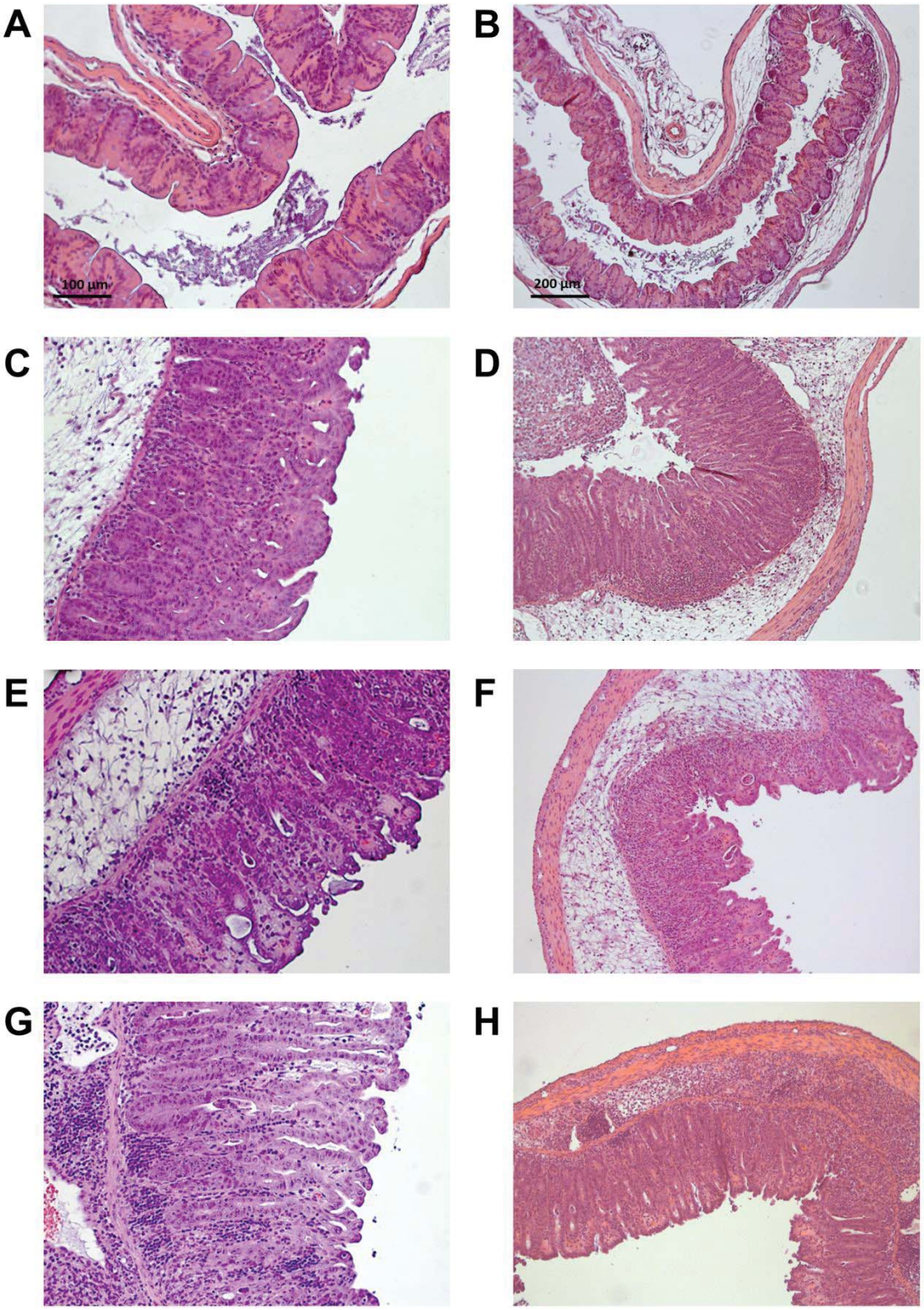


Figure 3.3. Inflammatory changes in *S.* Typhimurium-infected caecal tissue. Representative hematoxylin and eosin stained caecal tissue sections from: (A) and (B) naïve control mice at day 4 PI, (C) and (D) *S*. Typhimurium SL1344-infected mice at day 2 PI, (E) and (F) *S*. Typhimurium SL1344-infected 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% ± 34.3%, caecal tissue: 81.2% ± 5.7%), whereas in the colon *Salmonella* accounted for less than half of bacteria (colon content: 43.4% ± 11.9%) (mean ± 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% ± 2.3%, caecal tissue: 2.9% ± 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% ± 3.8%, caecal tissue: 3.5% ± 3.9%, colon content: 4.9% ± 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% ± 6.4% in samples from naïve control mice at day 4, and 7.6% ± 11.4% in samples from infected mice.

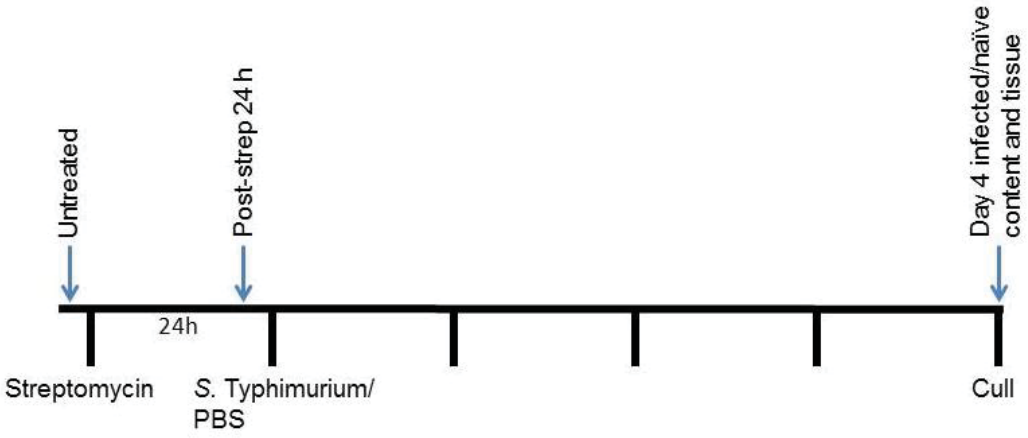
The genus *Olsenella* of the family *Coriobacteriaceae* displays an unusual sample distribution. In all untreated faecal samples and samples taken at day 4 from both naïve control and infected mice, *Olsenella* comprise < 1% of the sample microbiota. However in samples taken at 24 h post-streptomycin treatment *Olsenella* comprise a significant proportion of the microbiota (32.7% ± 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.

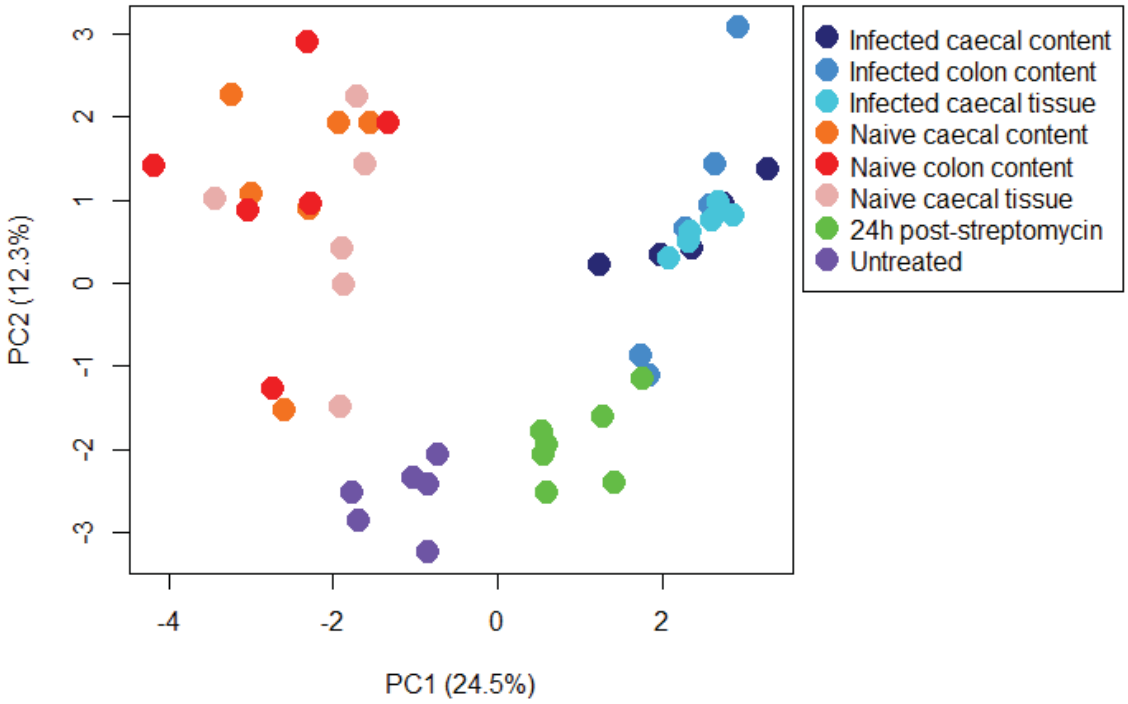
3.3.4.5 Effects of streptomycin treatment and *S.* Typhimurium infection upon microbial diversity

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.

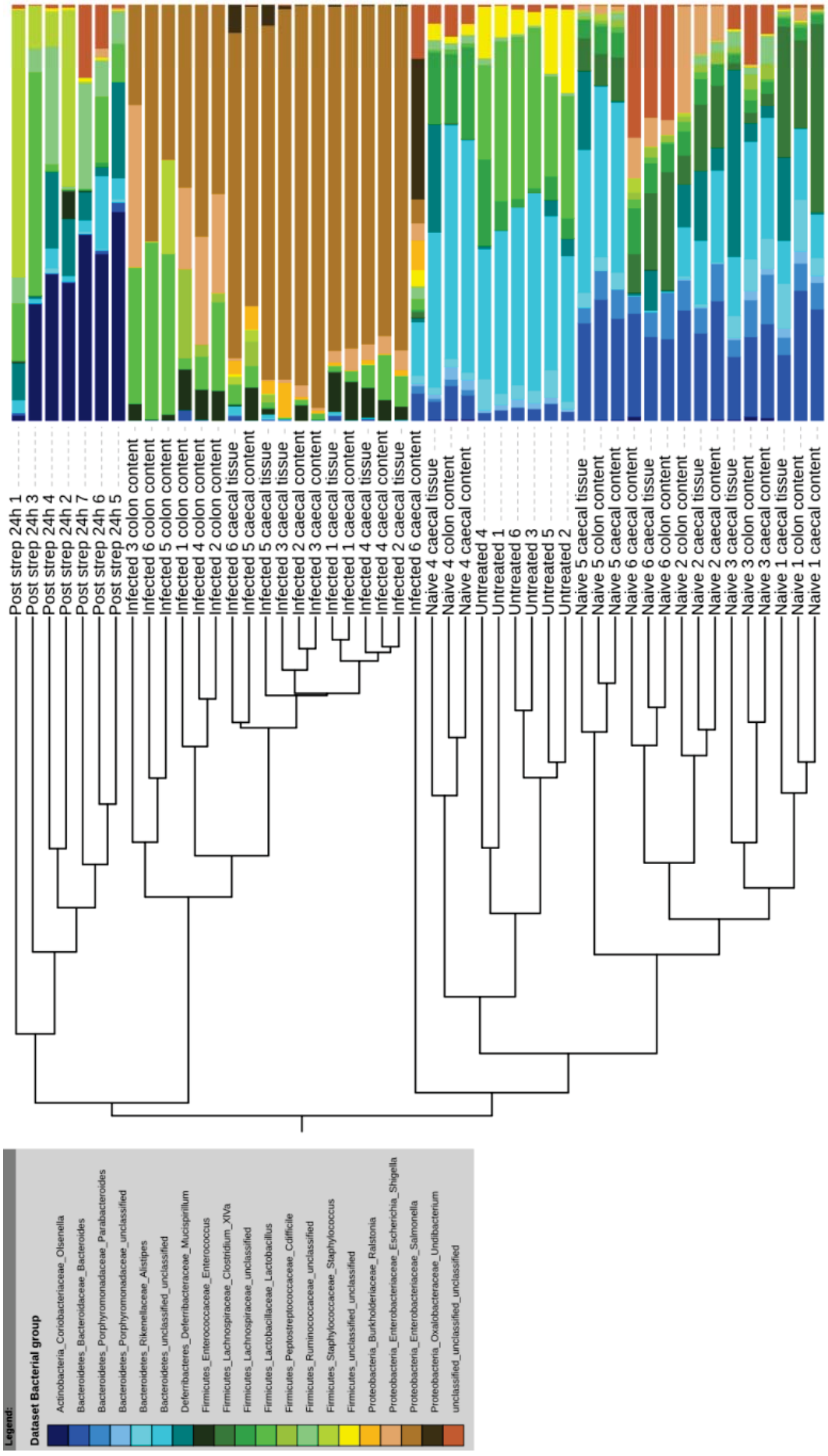
A


Untreated
Post-strep 24 h
Day 4 infected/naïve content and tissue
24h
Streptomycin
S. Typhimurium/ PBS
Cull

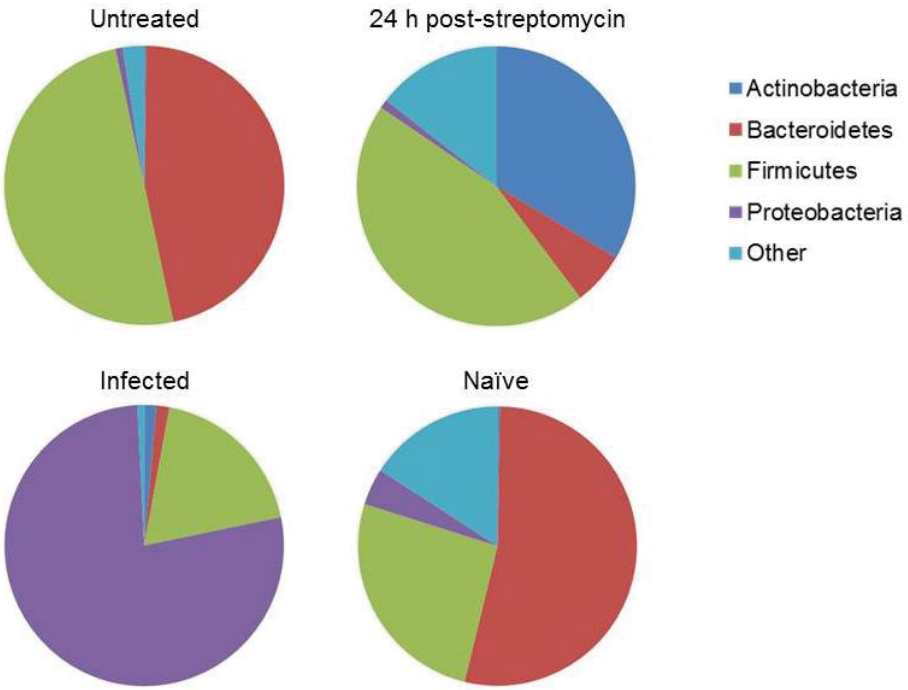
B


PC2 (12.3%)
PC1 (24.5%)
-4
-2
0
2
-3
-2
-1
0
1
2
3
Infected caecal content
Infected colon content
Infected caecal tissue
Naive caecal content
Naive colon content
Naive caecal tissue
24h post-streptomycin
Untreated

C


Post strep 24h 1
Post strep 24h 3
Post strep 24h 4
Post strep 24h 2
Post strep 24h 7
Post strep 24h 6
Post strep 24h 5
Infected 3 colon content
Infected 6 colon content
Infected 5 colon content
Infected 1 colon content
Infected 4 colon content
Infected 2 colon content
Infected 6 caecal tissue
Infected 5 caecal content
Infected 5 caecal tissue
Infected 3 caecal tissue
Infected 2 caecal content
Infected 3 caecal content
Infected 1 caecal tissue
Infected 1 caecal content
Infected 4 caecal tissue
Infected 4 caecal content
Infected 2 caecal tissue
Infected 6 caecal content
Naive 4 caecal tissue
Naive 4 colon content
Naive 4 caecal content
Untreated 4
Untreated 1
Untreated 6
Untreated 3
Untreated 5
Untreated 2
Naive 5 caecal tissue
Naive 5 colon content
Naive 5 caecal content
Naive 6 caecal content
Naive 6 caecal tissue
Naive 6 colon content
Naive 2 colon content
Naive 2 caecal tissue
Naive 2 caecal content
Naive 3 caecal tissue
Naive 3 colon content
Naive 3 caecal content
Naive 1 caecal tissue
Naive 1 colon content
Naive 1 caecal content
Legend:
Dataset Bacterial group
Actinobacteria_Coriobacteriaceae_Olsenella
Bacteroidetes_Bacteroidaceae_Bacteroides
Bacteroidetes_Porphyromonadaceae_Parabacteroides
Bacteroidetes_Porphyromonadaceae_unclassified
Bacteroidetes_Rikenellaceae_Alistipes
Bacteroidetes_unclassified_unclassified
Deferribacteres_Deferribacteraceae_Mucispirillum
Firmicutes_Enterococcaceae_Enterococcus
Firmicutes_Lachnospiraceae_Clostridium_XlVa
Firmicutes_Lachnospiraceae_unclassified
Firmicutes_Lactobacillaceae_Lactobacillus
Firmicutes_Peptostreptococcaceae_Cdifficile
Firmicutes_Ruminococcaceae_unclassified
Firmicutes_Staphylococcaceae_Staphylococcus
Firmicutes_unclassified_unclassified
Proteobacteria_Burkholderiaceae_Ralstonia
Proteobacteria_Enterobacteriaceae_Escherichia_Shigella
Proteobacteria_Enterobacteriaceae_Salmonella
Proteobacteria_Oxalobacteraceae_Undibacterium
unclassified_unclassified_unclassified

D


Untreated
24 h post-streptomycin
Infected
Naïve
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
Other

E

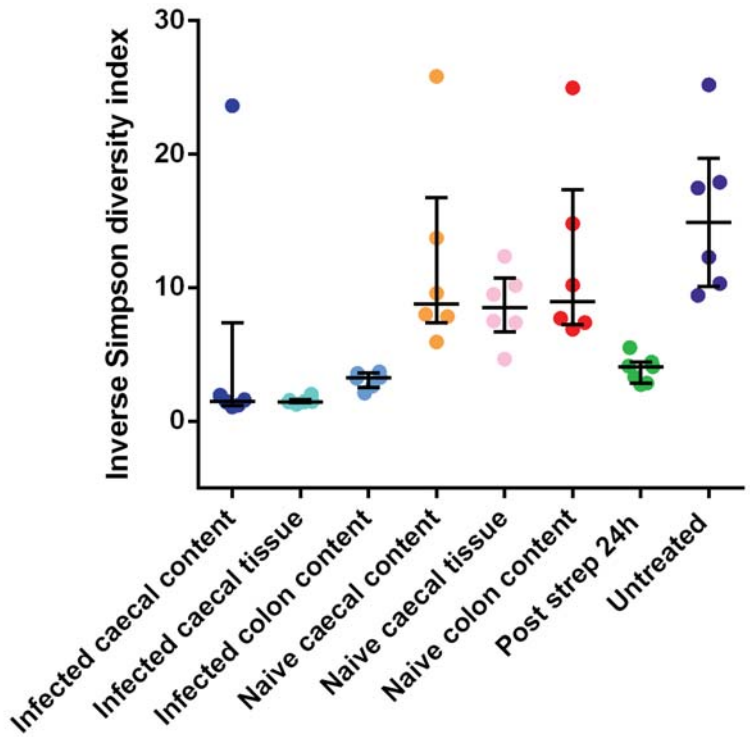
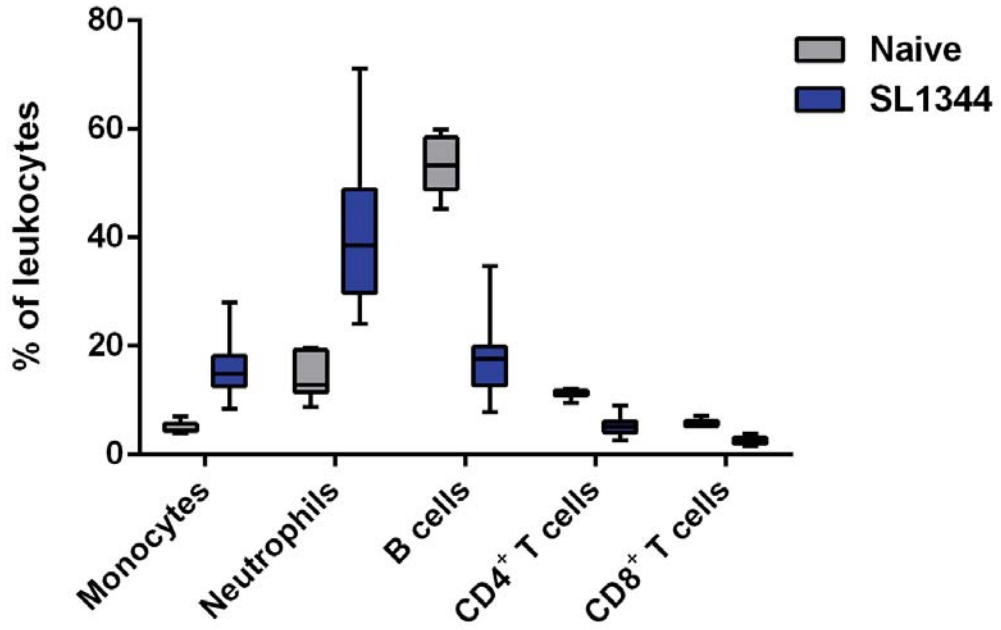

Inverse Simpson diversity index
30
20
10
0
Infected caecal content
Infected caecal tissue
Infected colon content
Naive caecal content
Naive caecal tissue
Naive colon content
Post strep 24h
Untreated

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


80
60
40
20
0
% of leukocytes
Naive
SL1344
Monocytes
Neutrophils
B cells
CD4+ T cells
CD8+ T cells

B

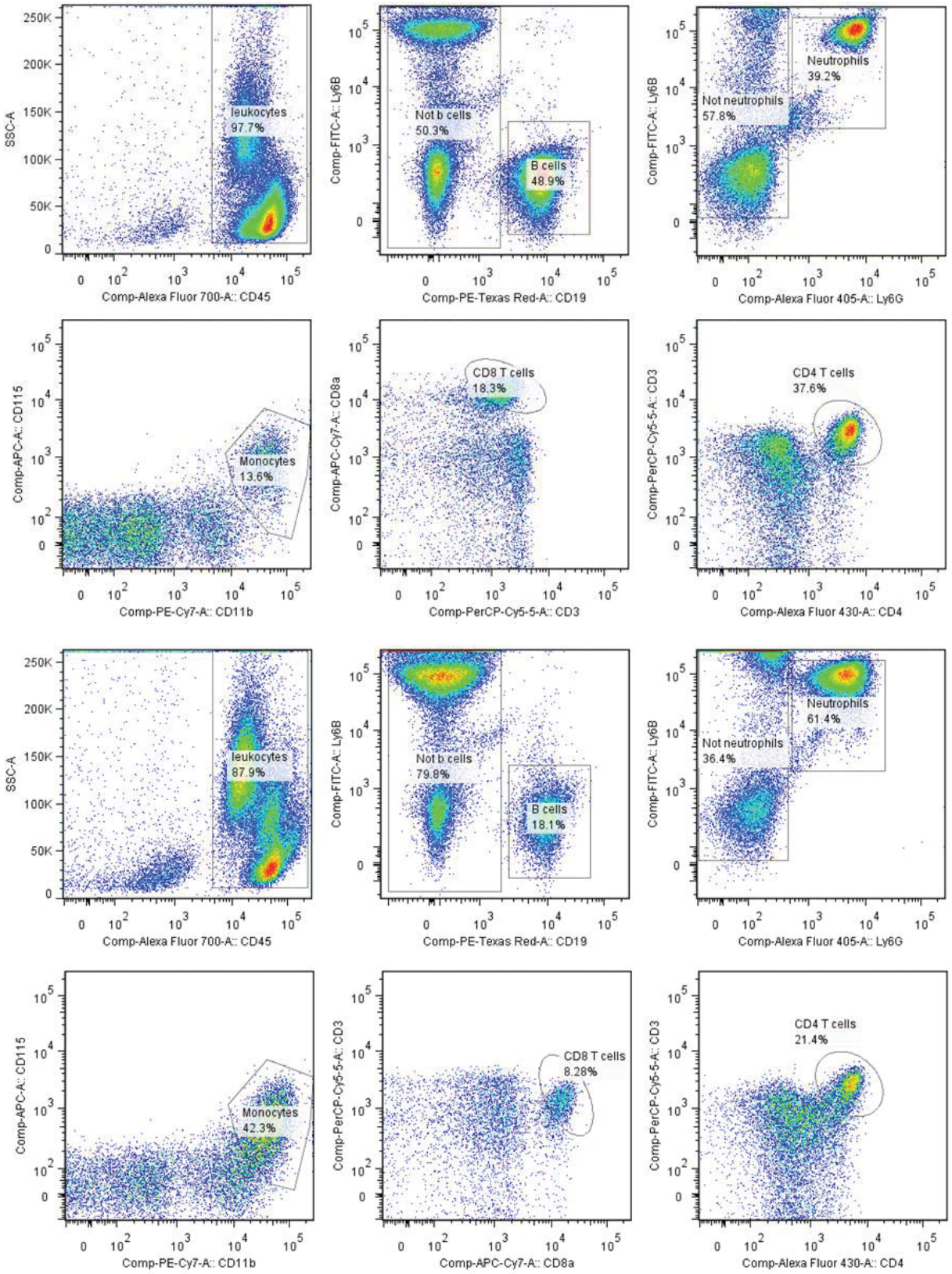

SSC-A
Comp-Alexa Fluor 700-A:: CD45
leukocytes
97.7%
Comp-FITC-A:: Ly6B
Comp-PE-Texas Red-A:: CD19
Not b cells
50.3%
B cells
48.9%
Comp-FITC-A:: Ly6B
Comp-Alexa Fluor 405-A:: Ly6G
Neutrophils
39.2%
Not neutrophils
57.8%
Comp-APC-A:: CD115
Comp-PE-Cy7-A:: CD11b
Monocytes
13.6%
Comp-APC-Cy7-A:: CD8a
Comp-PerCP-Cy5-5-A:: CD3
CD8 T cells
18.3%
Comp-PerCP-Cy5-5-A:: CD3
Comp-Alexa Fluor 430-A:: CD4
CD4 T cells
37.6%
SSC-A
Comp-Alexa Fluor 700-A:: CD45
leukocytes
87.9%
Comp-FITC-A:: Ly6B
Comp-PE-Texas Red-A:: CD19
Not b cells
79.8%
B cells
18.1%
Comp-FITC-A:: Ly6B
Comp-Alexa Fluor 405-A:: Ly6G
Neutrophils
61.4%
Not neutrophils
36.4%
Comp-APC-A:: CD115
Comp-PE-Cy7-A:: CD11b
Monocytes
42.3%
Comp-PerCP-Cy5-5-A:: CD3
Comp-APC-Cy7-A:: CD8a
CD8 T cells
8.28%
Comp-PerCP-Cy5-5-A:: CD3
Comp-Alexa Fluor 430-A:: CD4
CD4 T cells
21.4%

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 Δ*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).

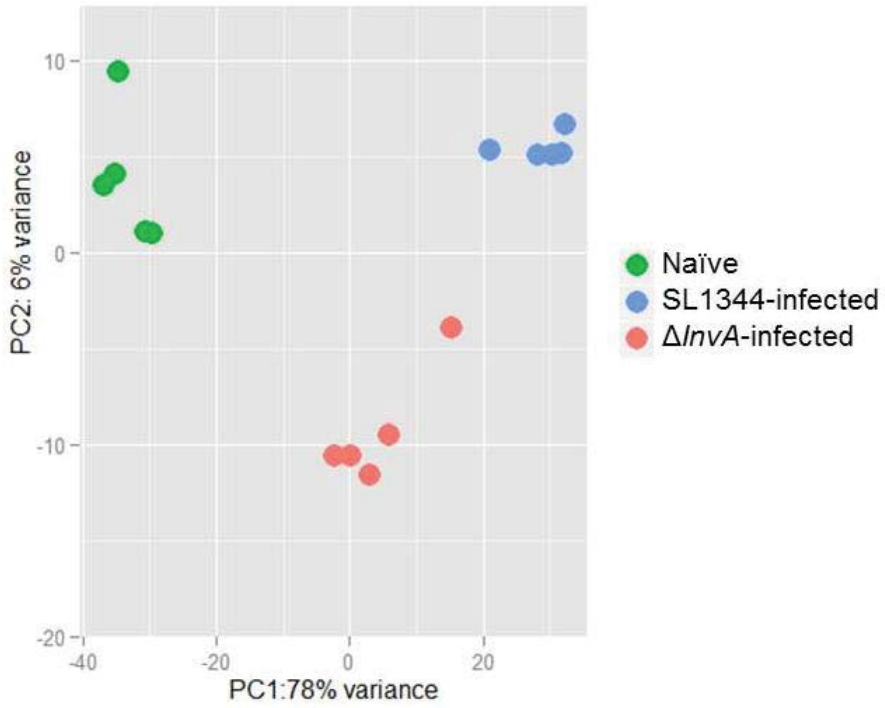


Figure 3.6. Principal component analysis of caecal transcriptional profiles. Points represent profiles from individual mice in the three treatment groups; naïve controls, *S*. Typhimurium SL1344-infected mice, and SL1344 Δ*InvA S*. Typhimurium-infected mice. Principal component analysis was performed 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 Δ*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 Δ*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.

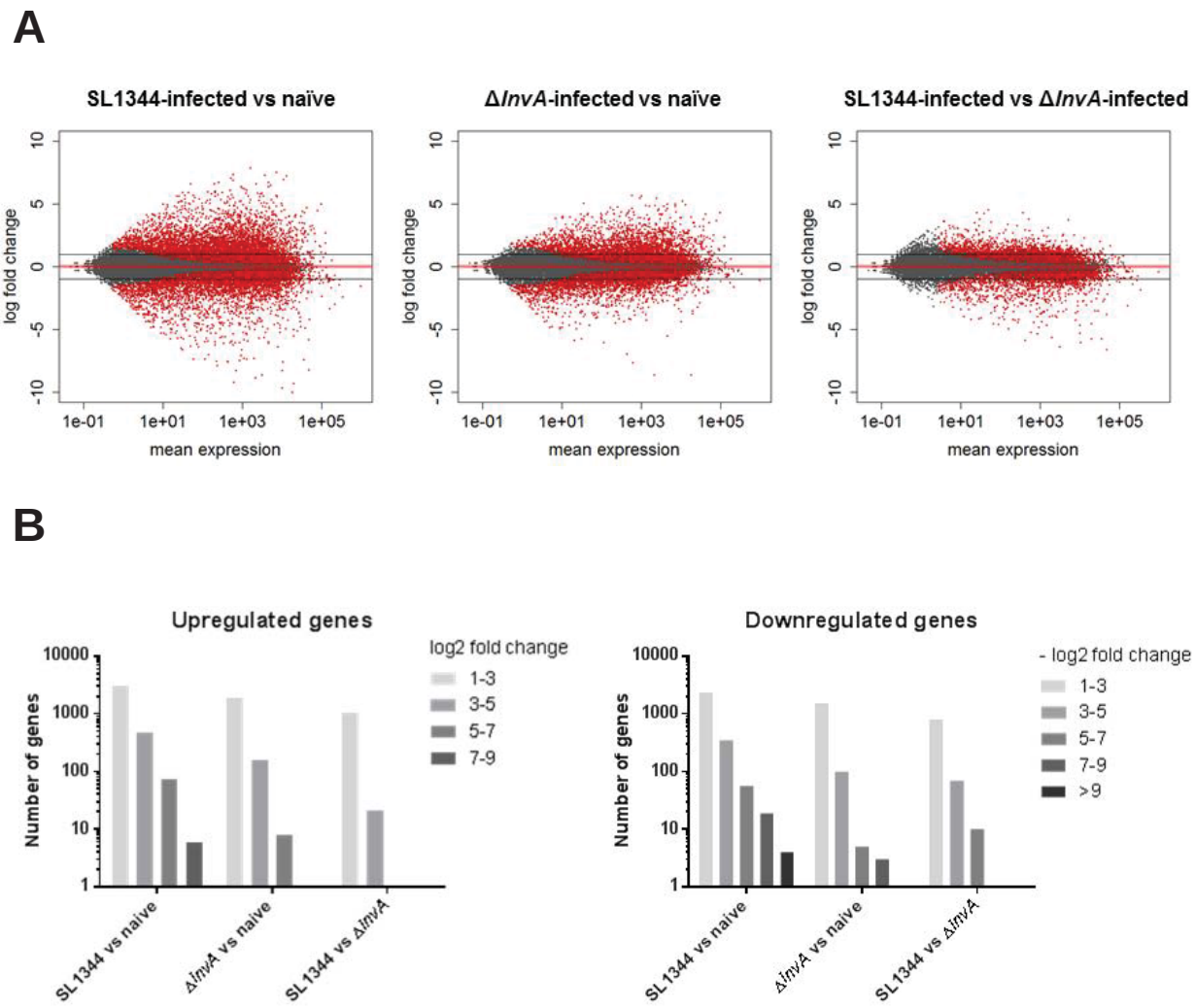


Figure 3.7. Infection with *S.* Typhimurium remodels the transcriptome in the mouse caecum. (A) MA plots providing an overview of transcriptome-scale differences in gene expression. Log2 fold change is plotted against normalised transcript abundance; each point represents a single gene. Points coloured 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 metallopeptidase 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 metallopeptidase 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.

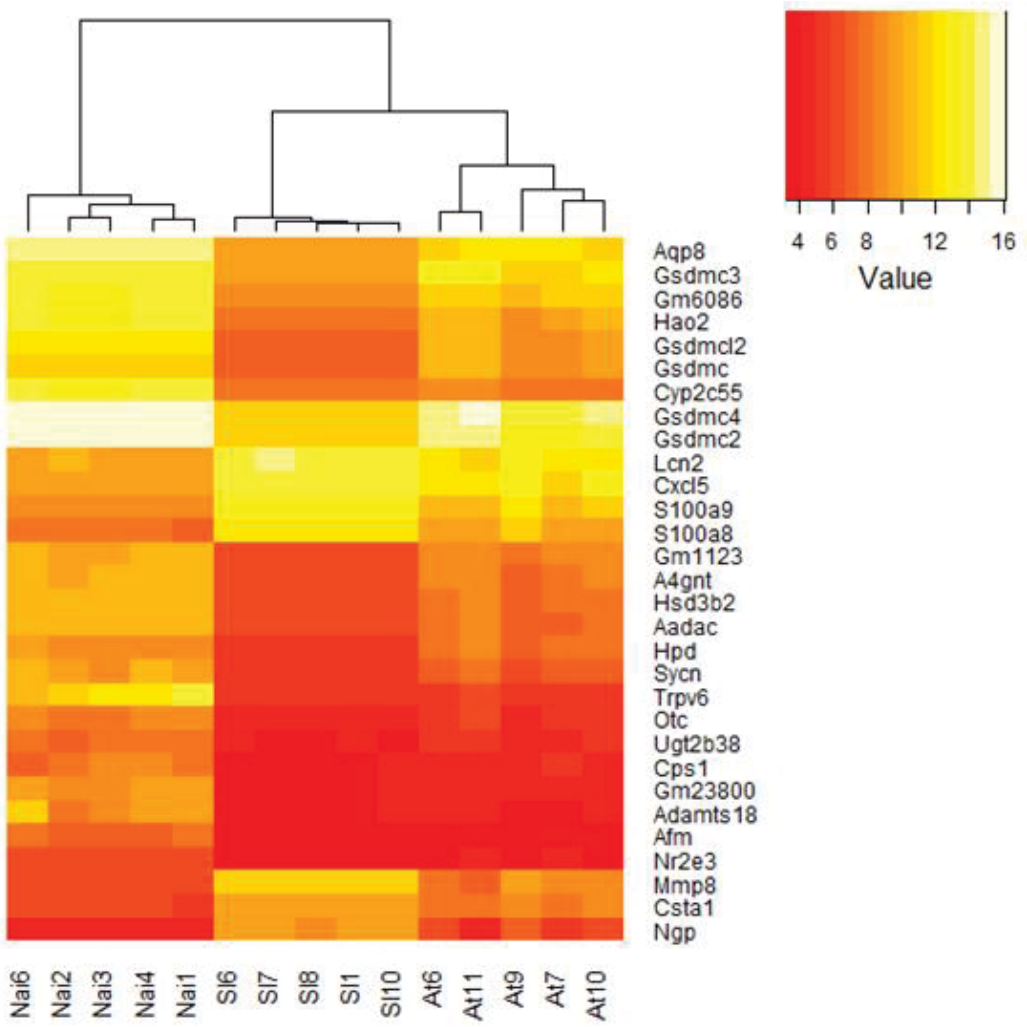


Figure 3.8. Expression profiles of the 30 most differentially expressed genes during *S.* Typhimurium infection of caecum. Heatmap of normalised transcript abundance in caeca of naïve control mice (samples with prefix 'Nai'), *S.* Typhimurium SL1344-infected mice (prefix 'Sl') and mice infected with the attenuated SL1344 Δ*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 (log2 fold change > 1, adjusted p-value < 0.05) and downregulated genes (log2 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.

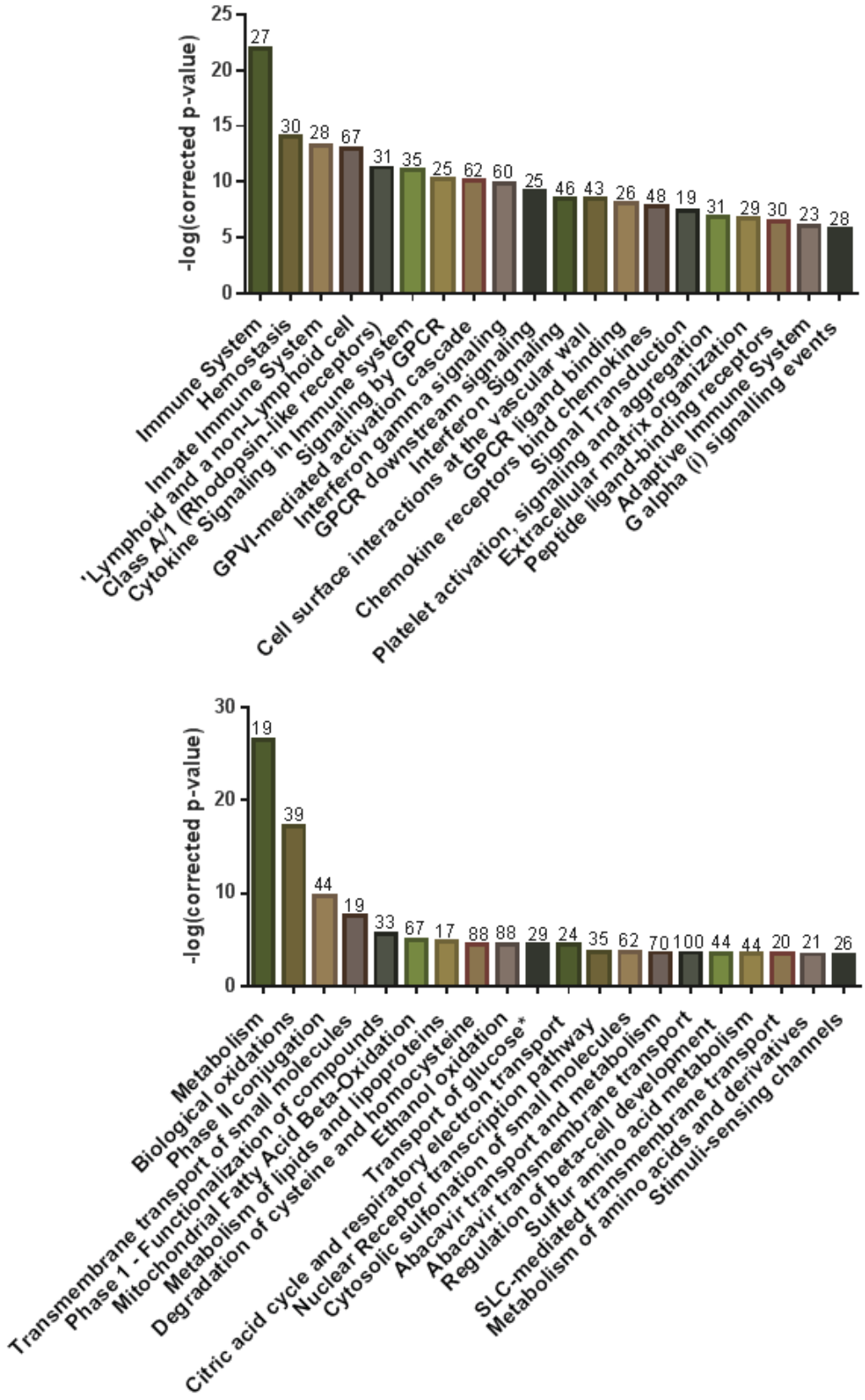


Figure 3.9. Pathway ORA of genes regulated in *S.* Typhimurium infection. The InnateDB pathway analysis tool was used to identify pathways annotated with genes regulated in infection, and to perform statistical

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 Δ*InvA*, was able to disseminate to extra-intestinal sites with efficiency similar to the wild type SL1344. However, mice infected with SL1344 Δ*InvA* displayed reduced intestinal organ counts with minimal weight loss. Histopathological analysis determined that SL1344 Δ*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 Δ*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 Δ*invG* and wild type infection, however lower loads of SL1344 Δ*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 Δ*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 Δ*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 Δ*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 Δ*InvA*-infected mice resulting in relatively lower numbers of *Salmonella* in the lumen. 16S rRNA gene sequencing of SL1344 Δ*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 ± 20,647 clusters generating an average of 1,969 ± 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 Δ*InvA*-infected sample groups, and least of all between the two infection groups. The smaller effect of SL1344 Δ*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 Δ*InvA*-infected sample groups compared with differentially expressed transcripts between the two infection groups indicates that the wild type- and SL1344 Δ*InvA*-infected tissues do share many transcriptional responses to *Salmonella*. This is in line with the observation that SL1344 Δ*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 NFκB 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 IFNɣ’ and ‘positive regulation of IFNɣ production’, in addition to ‘negative regulation of IFNɣ 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.