# 5 Assaying large *Salmonella* transposon mutant libraries in human macrophages

## *5.1 Introduction*

During human infection, Typhi traverses the gut epithelium and then invades, or is taken up by macrophages. Entry and survival inside both epithelial cells and macrophages are therefore key stages in the successful establishment of infection by Typhi. However, Typhimurium infection rarely progresses systemically in humans and so does not typically encounter macrophages, but this may be becoming a more common occurrence with some strains, particularly in Africa (Kingsley et al. 2009).

The growth conditions encountered by *Salmonella* in the macrophage have been the subject of studies performed using both serovars. Typhimurium has been demonstrated to generate and maintain *Salmonella*-containing vacuoles (SCVs) in this cell type, but data on Typhi is limited and largely extrapolated from Typhimurium (Drecktrah et al. 2007; Faucher et al. 2006; Oh et al. 1996). The mechanism of cell entry is also less than clear, with active invasion and uptake by the host both possible in this phagocytic cell type.

Successful Typhimurium infection of the murine host requires the ability to synthesise aromatic amino acids and nucleotides (Hautefort et al. 2008). Amino acids and purines appear to be limiting because the relevant mutant strains are impaired for intracellular survival and virulence in mice (Muñoz-Elías and McKinney 2006). Typhimurium SCVs of both macrophages and epithelial cells have been reported as limiting for aromatic amino acids, purines and pyrimidines (Hautefort et al. 2008). Proteomics from Typhimurium at various stages post infection of murine macrophages showed that the

physiological functions of recovered proteins were related to housekeeping roles, involving biosynthesis and metabolism of amino acids, carbohydrates, lipids, proteins and nucleotides, energy production and cellular processes (Shi et al. 2006). Genes encoding the transport of gluconate, glucoronate/galacturonate and galactonate have also been shown to be upregulated, alongside genes involved in the interconversion of these sugars to pyruvate and glyceraldehyde-3-phosphate (Eriksson et al. 2003). The Entner-Doudoroff pathway may also be acting as a rich source of NADPH for use in biosynthetic pathways and redox cycling (Eriksson et al. 2003). An overall picture has been uncovered, suggesting that the Typhimurium environment inside the macrophage is aerobic, low in phosphate, high in potassium, rich in amino acids and that gluconate and related carbohydrates may be the principal source of carbon for growth (Thompson et al. 2006).

Interestingly, results from studies on Typhi and human macrophages describe a slightly different story. One microarray study suggests iron is not limiting in the Typhi SCV and it may not be limiting for manganese or phosphate either, conclusions that were inferred from the non-induction of *sitABCD*, *phoN* and *pstAB* respectively (Faucher et al. 2006). Also, Typhi, but not Typhimurium, upregulates *aceA*, isocitrate lyase, which is involved in the metabolism of fatty acids and acetate as carbon sources via the glyoxylate shunt (Muñoz-Elías and McKinney 2006). This is probably because Typhi lacks the *dgo* operon involved in gluconate utilisation, and the associated transport proteins GntT and GntU were not up-regulated during intracellular growth. Other sugar transport systems (including fructose and hexose) were also repressed – suggesting that Typhi uses different carbon sources inside macrophages to Typhimurium (Faucher et al. 2006).

The environmental differences observed thus far in Typhi and Typhimurium macrophage conditions may be due in part to the different cell lines used. Hence, to ensure observed differences are genuine, both serovars need to be tested in the same cell line.

In this chapter, an assay of the genes required by Typhi and Typhimurium for infection of human macrophages was performed in a high throughput manner by making use of the one million mutant transposon libraries and Illumina sequencing. This picture was interpreted further by placing the relevant genes into a metabolic context, using StyCyc and StmCyc as references. Human macrophages represent a biologically relevant culture condition for Typhi, and by comparing the results with Typhimurium, gives further insight into the different strategies employed by these pathogens for adherence to and infection of this host cell type.

## *5.2 Methods*

All reagents were obtained from Sigma-Aldrich, Dorset, UK unless otherwise stated.

### 5.2.1 Strains and cell lines

Assays were carried out using Typhi WT174 and Typhimurium SL3261, and the transposon mutant libraries created in these strains, previously described in Chapter 4. Human monocytic cell line THP-1 was used for cell infections.

### 5.2.2 Growth in RPMI

Typhi WT174 and Typhimurium SL3261 were tested for growth in RPMI-1640 supplemented with 0.3 g/L L-glutamine only, and in RPMI supplemented with 0.3 g/L L-glutamine and buffered with 0.1 M MOPS (final concentration). Cultures of 4.5 mL were grown for 16 h shaking at 37 $^{o}$C before the $OD_{600}$ was determined on a spectrophotometer, using non-inoculated RPMI as a control.

Three independent growth curves per strain were also determined using RPMI supplemented with L-glutamine and buffered with 0.1M MOPS, in a culture volume of 100 mL. The $OD_{600}$ was measured at 18-20 timepoints over an 8 h period and at a final timepoint of 24 h.

### 5.2.3 Preparation of THP-1 cells

THP-1 cells were grown up from frozen stocks in RPMI-1640 supplemented with 10% heat-inactivated foetal bovine serum and 2 mM L-glutamine, and incubated without shaking in vented flasks (VWR, Lutterworth, UK) at 37 $^{o}$C in the presence of 5% $CO_2$. To facilitate growth, culture volumes were split and given fresh media every 3-4 days until the required cell density was reached. Typically, a starter culture of 5 mL was expanded to 30 mL over a week, then split into 3 flasks with each expanded to 60 mL over four days. These were then split into 5 flasks and each finally expanded to 100 mL. Cell density was measured using a 10 μL aliquot of culture mixed with 10 μL Trypan Blue stain pipetted into a chamber on a C-Chip haemocytometer (Labtech International, Ringmer, UK). Visualised under a light microscope, live cells (which exclude Trypan Blue) were manually counted and the density within each flask was estimated. Between one and five 6-well plates were used for each timepoint of the infection assay. The appropriate volume of culture for the entire assay, corresponding to 2 x $10^6$ cells per well, was transferred into 50 mL tubes for centrifugation at 1200 rpm for 8 minutes. The supernatant was poured off and cells were resuspended in 2 mL x number of wells required of warmed, supplemented RPMI. PMA (phorbol myristate acetate) was used to differentiate the THP-1 monocytes; frozen 10 μL aliquots of 50 ng/mL PMA were thawed and mixed with 990 μL warmed, supplemented RPMI. Five hundred μL aliquots were then added to 49.5 mL aliquots of warmed, supplemented RPMI. A volume equivalent to 2 mL x number of wells required was mixed with the cell culture (i.e. a 1 in 2 dilution) and 4 mL cells were seeded into each well required for the infection assay.

The 6-well plates were incubated for six days at 37 $^{o}$C in 5% $CO_2$. Upon the day of infection, the PMA-containing media was removed, cells were washed with dPBS and fresh warmed, supplemented RPMI was added to maintain the cells whilst the bacterial inoculum was prepared.

### 5.2.4 Preparation of transposon libraries

Frozen stocks of the Typhi library were measured by $OD_{600}$ to be at half the concentration of the Typhimurium library. To ensure the overnight cultures started at similar concentrations, a 1 in 5000 dilution of the Typhi library and a 1 in 10,000 dilution of the Typhimurium library was used to inoculate the growth medium. Cultures for each transposon library were grown in 10 mL or 100 mL of RPMI-1640 supplemented with 0.3 g/L L-glutamine and buffered with 1 mL or 10 mL 1 M MOPS, to give a final concentration of 0.1 M, on a shaker at 37 $^{o}$C for 16 h. These cultures were sub-cultured at 1 in 20 into fresh RPMI supplemented and buffered as before, and grown for between 3 and 4 hours to mid-log phase ($OD_{600}$ of 2.4).

### 5.2.5 Optimisation of gentamicin infection assay for TraDIS

#### *5.2.5.1 Preliminary infection*

Silvia Pinero prepared the THP-1 cells for infection.

THP-1 cells were grown up and differentiated as described above, except that the cells were seeded at a density of ~1 x $10^5$ in 1 mL, per well of a 24-well plate. Upon the day of

infection, cells were washed in dPBS and 1 mL fresh supplemented RPMI was added per well.

A culture of Typhimurium strain SL3261 was grown in LB broth (Oxoid, Basingstoke, UK), and incubated shaking at 37 $^{o}$C for 16 h. A control Typhi strain (BRD948) was also cultured. 150 µL of these cultures were then used to inoculate 3 mL fresh LB and incubated shaking at 37 $^{o}$C for 3 h to an $OD_{600}$ of 0.2 for late exponential growth. Per strain, three experimental wells were inoculated for each timepoint (30 minutes, 2 hours and 4 hours), using 5 µL of the bacterial culture per well. The 24-well plate was centrifuged for 5 minutes at 600 x *g* and incubated at 37 $^{o}$C in 5% $CO_2$ for 30 minutes. After 30 minutes, media from all wells was removed and replaced with either 1 mL dPBS + 100 µg/mL gentamicin (30 minute wells) or 1 mL supplemented RPMI + 100 µg/mL gentamicin (2 h and 4 h wells). The 30 minute wells were washed twice in dPBS, and the plate was incubated for a further 90 minutes. At this point, the 2 h wells were washed twice with dPBS, and the plate incubated for a further 2 h. After the full 4 h, the 4 h wells and the blank control were washed twice with dPBS and then 100 µL of 1% Triton-X-100 (VWR) was added to every well and the plate incubated at 37 $^{o}$C for 2 minutes. This 100 µL was serially diluted to $10^{-5}$ and 3 x 10 µL drops were plated out on LB plates (supplemented with 'aro' mix for Typhi) for each dilution. The LB plates were incubated for 16 h at 37 $^{o}$C.

#### *5.2.5.2 Small scale*

Initially, the assay was performed for each transposon library using one 6-well plate per timepoint: 30 minutes, 2 hours and 4 hours. Three wells of the 6-well plate were used as experimental replicates, one as a blank control and the remaining two for plating out viable counts.

At the start of the assay, media was removed from all wells except for the blank control, and a 3 mL bacterial inoculum was added to each experimental well. The plates were centrifuged for 5 minutes at 600 x *g* and incubated at 37 $^{o}$C in 5% $CO_2$ for 30 minutes. A 2-3 mL aliquot of the inoculum was processed for genomic DNA as the input sample for TraDIS. After the 30 minutes, media was removed from all wells, and fresh RPMI additionally supplemented with 100 μg/mL gentamicin was added to wells for the 2 h and 4 h timepoints. The 30 min timepoint wells were washed twice in dPBS supplemented with 100 μg/mL gentamicin and once in plain dPBS. After 2 h and 4 h, the relevant wells were washed 3 times in plain dPBS. Following washing, 500 μL of 1% Triton-X-100 was added to each well to lyse the eukaryotic cells, mixed well by pipetting, and incubated at 37 $^{o}$C in 5% $CO_2$ for 2 minutes. Cell suspensions from the three experimental wells at the same timepoint were pooled for bacterial DNA extraction. Three 10 μL droplets of serial dilutions (from $10^{1}$ to $10^{-6}$) from the inoculum, two wells from each timepoint and from the blank controls were plated out onto L-agar (supplemented with 'aro' mix for Typhi) and incubated for 16 h at 37 $^{o}$C to establish viable counts.

It was unknown whether sufficient DNA could be extracted directly after the infection, or if the recovered bacteria should be grown up overnight. To determine this, in the first infection assay, 200 μL of the pooled cell suspension from each timepoint was used to

inoculate 10 mL RPMI (supplemented with 0.3 g/L L-glutamine and buffered to a final concentration of 0.1 M MOPS) and incubated shaking at 37 $^{o}$C for 16 h before DNA extraction.

#### *5.2.5.3 DNA manipulation (small scale)*

Genomic DNA was extracted from the inocula and pooled cell suspensions from each timepoint using either the Qiagen Generation Capture Columns or the DNeasy Blood and Tissue Kit (Qiagen, Crawley, UK), according to the manufacturer's protocol for Gram negative bacteria, with the optional RNAse step. DNA samples were quantified using the Qubit fluorescence-based dsDNA HS assay (Invitrogen, Paisley, UK).

Quantitative PCR was performed on one set of samples, using genomic DNA from the Typhi transposon mutant pool as DNA standards serially diluted from 30 ng/μL to 1.875 ng/μL. The PCR was run on an ABI cycler (Applied Biosystems, Warrington, UK) using the recommended 25 μL maximum reaction volume. For each sample, 12.5 μL of the QuantiTect SYBR Green PCR mix (Qiagen) was added to 0.125 μL (5μM) of a forward and reverse primer, 11.25 μL distilled $H_20$ and 1 μL of the DNA sample. Each sample was tested using the following primers for two genes, *aroC* and *phsA*:

*aroC*-05: 5'- GTGATCCATCAGTACGATCG and *aroC*-06: 5'- GACAACTCTTTCGCGTAACC

*phsA*-01: 5' – GATTTGAACAGTTAGCACAG and *phsA*-02: 5'- ACAGGTGACGGTCTGGAAAG

### 5.2.6 Large scale infection assay

The assay was carried out as before, except five 6-well plates were utilised for every timepoint assayed. In total, 29 wells were infected with the bacterial inoculum and one served as a blank control for eukaryotic cell contamination. A 4-6 mL aliquot of the inoculum was processed for genomic DNA as the input sample for TraDIS. The cell suspensions from all experimental wells at the same timepoint were pooled for bacterial DNA extraction. A 100 μL aliquot of the pooled cell suspension from each timepoint was serially diluted to $10^{-6}$ and plated onto L-agar (supplemented with 'aro' mix for Typhi) to establish viable counts; three 10 μL droplets of serial dilutions were also plated out from the inoculum and the control from each timepoint, and incubated as before.

#### *5.2.6.1 Single timepoint replicates*

The assay was again carried out as before, except that three independent 100 mL cultures were grown up overnight, used to inoculate 3 independent 270 mL starter cultures and then to infect 3 independent sets of five 6-well plates for 2 hours. For each replicate, a total of 29 wells were infected with the bacterial inoculum and one served as a blank control for eukaryotic cell contamination. All other aspects were performed as before.

#### *5.2.6.2 DNA manipulation (large scale)*

The pooled cell suspension from each timepoint was pelleted and washed twice in RPMI buffered with 0.1 M MOPS (large-scale), or in dPBS (single timepoint replicates). For the large-scale assays, the sample was split into 3 at this point to determine if this increased

DNA yield. For the multiple timepoint replicate assay, the sample was processed as one. Pellets were resuspended in 85 μL $dH_20$, 5 μL DNase I (New England Biolabs, Hitchin, UK) and 10 μL 10 x Buffer (New England Biolabs), incubated at room temperature for 15 minutes and then at 65 $^oC$ for 10 minutes. Samples were centrifuged for 10 minutes at 7500 rpm and resuspended in 180 μL Buffer ATL (Qiagen). Genomic DNA was extracted from the inoculum and from the Buffer ATL samples using the Qiagen DNeasy Blood and Tissue kit, according to the manufacturer's protocol for Gram negative bacteria. DNA samples were quantified using the Qubit fluorescence-based dsDNA HS assay (Invitrogen).

## 5.2.7 TraDIS analysis

### *5.2.7.1 Sample processing and nucleotide sequencing*

Sabine Eckert and Daniel Turner performed the preparation and nucleotide sequencing for the large-scale and single timepoint replicate samples.

All sample clean-ups were performed using Qiagen MinElute columns.

Two μg of DNA per sample was sheared by Covaris AFA (Quail et al. 2008) on an S2 instrument (at settings 5, 20%, 200 bursts per cycle and 90 seconds) to an average length of 230 ± 100 nucleotide bases. The sheared DNA was cleaned up and end-repaired using the NEBNext Sanger Sequencing Sample Preparation kit (New England Biolabs). Samples were cleaned up again and A nucleotides added using the same NEB kit. Following another clean-up, samples were quantified with the Agilent Bioanalyzer 2100

and ligated to a 10-fold excess of annealed adapters (nucleotide sequences below) with the NEB kit.

Adapter 1: PE_AD_T ACACTCTTTCCCTACACGACGCTCTTCCGATC*T

Adapter 2: PE_AD_B GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG

A final clean-up preceded a qPCR to again quantify sample concentration. This was performed using the following primers, against standards of known concentration:

Ad_T_qPCR1: 5' - CTTTCCCTACACGACGCTCTTC

Ad_B_qPCR2: 5' - ATTCCTGCTGAACCGCTCTTC

200 ng of DNA from the bacterial input samples, and 400 ng from samples taken post-infection were PCR-amplified using Jumpstart Taq DNA polymerase (Sigma) with the following primers:

5TMDHp5F2:

5' - CAAGCAGAAGACGGCATACGAGATCGCTGAATTACCCTGTTATCCCTATTTAGGTGAC

PE PCRv3.3:

5' - CAAGCAGAAGACGGCATACGAGATCGGTACACTCTTTCCCTACACGACGCTCTTCCGATCT

at 94 $^{o}$C for 2 minutes, followed by 22 cycles of 94 $^{o}$C for 30 seconds, 65 $^{o}$C for 20 seconds and 72 $^{o}$C for 30 seconds, then 72 $^{o}$C for 10 seconds. The PCR products were quantified by qPCR using the primers below with standards of known concentration.

Syb_FP5: 5' - ATGATACGGCGACCACCGAG

Syb_RP7: 5' - CAAGCAGAAGACGGCATACGAG

Samples were then separated using a 2% TBE agarose gel and fragments in the range of 350 – 500 nucleotide bases were isolated using the Qiagen Gel-Extraction kit, according

to the manufacturer's instructions. These were quantified by qPCR, using primers Syb_FP5 and Syb_RP7 as before. Prior to flowcell preparation, samples were denatured in 0.2 M NaOH for 5 minutes, neutralised with 0.1 M HCl and diluted to a final concentration of 10 pM in hybridization buffer. Flowcells were prepared according to the Illumina protocols and samples were sequenced using a Genome Analyzer II with the custom primer 5TMDH2seq (see Chapter 4).

#### *5.2.7.2 Sequence data processing*

This was undertaken as described in Chapter 4, section 4.2.5.

### 5.2.8 Statistical analysis of required genes

The nature of this analysis was discussed with Leopold Parts who wrote scripts in R to perform the following analyses.

#### *5.2.8.1 Large scale assays*

The processed sequence data from all timepoints post infection were normalised for total sequencing yield and combined. For every gene $g$ present in the input sample and the combined output, with $n_{g,A}$ reads observed in the input and $n_{g,B}$ reads observed in the output, we calculated the $\log_2$ fold change ratio $S_{g,A,B} = log_2 \setminus frac(n_{g,A}+100)(n_{g,B}+100)$. The correction of 100 reads smoothed out the high scores for genes with very low numbers of observed reads. We fitted a normal model to the mode +/- 2 sample standard

deviations of the distribution of $S_{A,B}$, and calculated *P*-values for each gene according to the fit.

For Typhimurium, a cutoff was chosen at a $\log_2$ fold change ratio of at least 4 that gave a false discovery rate (FDR) of 4.5%, corresponding to a *P*-value of 0.001.

For Typhi, the same calculations were applied, but using the number of insertion sites rather than mapped reads since these provided a clearer cutoff, which was chosen at a $\log_2$ fold change ratio of at least 2 that gave an FDR of 3.8%, corresponding to a $10^{-4}$ *P*-value.

#### *5.2.8.2 Single timepoint replicates in Typhi*

The total number of insertion sites per gene for the three input samples and the three 2 h 'output' samples were combined and normalised to the number of insertions found in the first input replicate. In the same manner as above, $\log_2$ fold change ratios in insertion sites were calculated per gene, with corresponding *P*-values. A cutoff was chosen at a $\log_2$ fold change ratio of at least 1, that gave an FDR of 3.6%, corresponding to a $10^{-6}$ *P*-value.

## *5.3 Results*

### 5.3.1 Preliminary Typhimurium assay

Prior to using the Typhimurium SL3261 transposon library itself with the human cell line THP-1, a preliminary infection assay was carried out using the Typhimurium strain in which the library was constructed. This was to ensure that this strain was capable of infecting this human cell line. Differentiated THP-1 cells were infected with the Typhimurium strain (approximately 10 bacteria per macrophage, a multiplicity of infection (MOI) of 10) and with Typhi BRD948 (MOI 60) as a control, and viable counts were performed at 30 minutes, 2 hours and 4 hours post-infection. Figure 5-1 shows the results from each timepoint and demonstrates that Typhimurium is capable of infecting THP-1 cells. In addition, uptake was approximately 20-fold higher for Typhimurium than for the control (Typhi) strain. This was likely due to the expression of the Vi antigen by the latter, which is known to negatively affect uptake (Arricau et al. 1998).

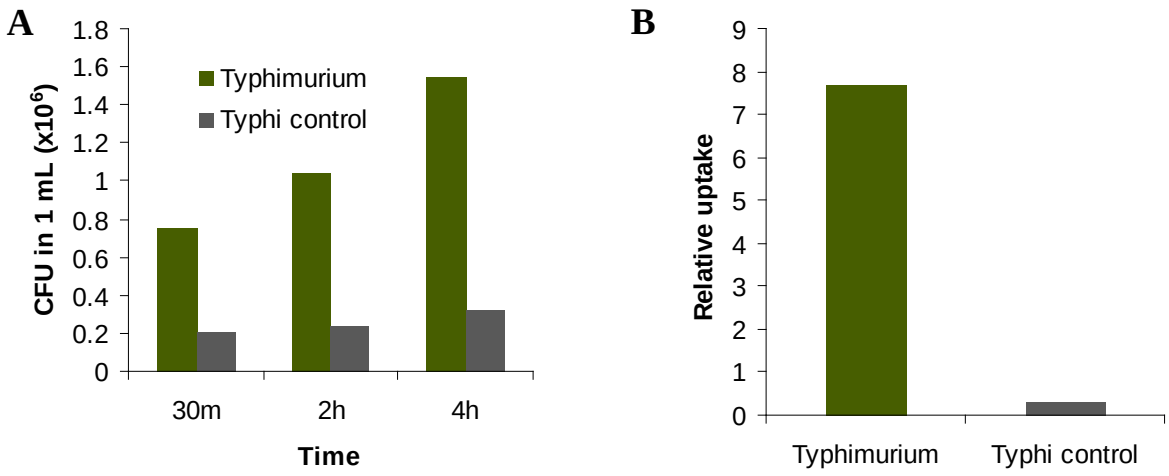


**Figure 5-1 Preliminary infection assay**

A) Total bacteria recovered from THP-1 gentamicin infection assay at three timepoints. CFU, colony forming units given in millions per mL. B). Relative uptake of bacteria, calculated by dividing the CFU at 30 minutes by the inoculum.

## 5.3.2 Optimising the cell infection assay for TraDIS

### *5.3.2.1 Media*

THP-1 cells are cultured in RPMI media for optimum growth. If the bacterial inoculum was grown in LB and then added to the eukaryotic cells in RPMI, there was the possibility that mutants would be negatively affected by the culture media, rather than their infection potential. Hence, both bacterial strains used to construct the transposon libraries were tested for their ability to grow in RPMI. The ability of the Typhi transposon library strain to grow without the addition of 'aro' mix was also tested. Multiple growth curves of each transposon library strain in buffered RPMI, without 'aro' mix indicated that both strains were capable of growth in this culture media, allowing the first infection assay to be carried out using buffered RPMI to grow up the bacterial inocula.

### *5.3.2.2 Infection assays*

Viable counts from the first small scale infection assay using the transposon libraries indicated that the assay had been successful: bacteria were recovered from the eukaryotic cells at all timepoints tested (Figure 5-2). Bacterial numbers did not increase over the timepoints as has been observed in the preliminary infection, but the total number of bacteria recovered at 4 hours with the transposon libraries (~ 4 x $10^7$) was actually in excess of that recovered for the same timepoint in the preliminary Typhimurium infection (1.5 x $10^6$). This is likely due to the high $OD_{600}$ used for infection, which resulted in the multiplicity of infection in this assay being higher than any other. High MOIs were however used in all the infections to ensure every transposon mutant was well represented in the assay (Table 5-1).

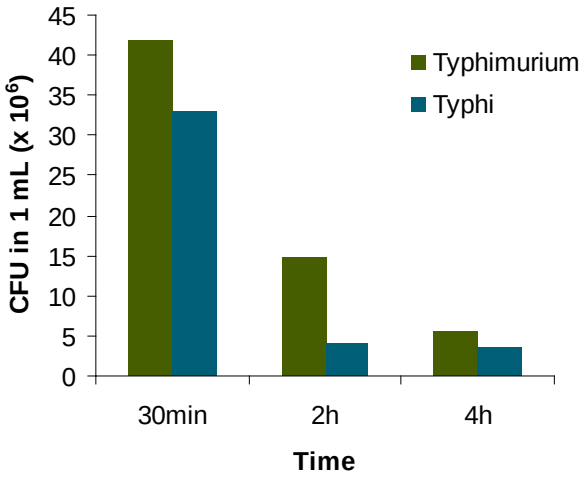


**Figure 5-2 Transposon library cell counts from initial THP-1 infection**

Total number of bacteria recovered after timepoints indicated. CFU, colony forming units given in millions per mL.

**Table 5-1 Small scale infection assays**

| Assay | $OD_{600}$ at infection | MOI | DNA extraction | Comments |
|---|---|---|---|---|
| 1 SL | 0.41 | 750 | direct and overnight | Low DNA yield from both extraction times |
| 1 WT | 0.53 | 960 | direct and overnight | Low DNA yield from both extraction times |
| 2 SL | 0.2 | - | direct only | Eukaryotic cells contaminated |
| 2 WT | 0.2 | - | direct only | Eukaryotic cells contaminated |
| 3 SL | 0.2 | 405 | direct only | - |
| 3 WT | 0.2 | 270 | direct only | Checked by qPCR |
| 4 SL | 0.2 | 315 | direct only | - |
| 4 WT | 0.2 | 345 | direct only | - |
| 5 SL | 0.2 | 225 | direct only | - |
| 5 WT | 0.2 | 285 | direct only | - |

SL, Typhimurium; WT, Typhi; MOI, multiplicity of infection. MOI could not be determined for the second set of infection assays due to contamination on the viable count plates.

The question of when to extract genomic DNA from invading bacteria was addressed in the first infection assay using the transposon libraries. DNA was extracted from transposon mutants harvested directly after each timepoint in the infection, and from aliquots of the same mutants that had been grown overnight in RPMI. Quantification of these samples indicated that yields were extremely low in both cases, although the overnight samples were slightly improved (Table 5-2). However, the small increase in DNA yield was not sufficient to justify the possible bias generated by the transposon mutants from overnight growth.

Further infection assays were performed using a different DNA extraction kit to determine whether increased DNA yields could be obtained directly from the harvested mutants. In total, three independent assays were completed without contamination (Table 5-1), each resulting in greater DNA yields than the original assay (Table 5-2).

**Table 5-2 DNA yields from different extraction kits**

| Time of DNA extraction | Sample | Typhimurium DNA yield (µg) | Typhi DNA yield (µg) | DNA extraction kit |
|---|---|---|---|---|
| Directly following infection | inoculum | 0.034 | 0.25 | Capture columns |
| | 30 minutes | 0.33 | 0.07 | |
| | 2 h | 0.13 | 0.03 | |
| | 4 h | 0.11 | 0.024 | |
| After overnight growth | inoculum | 0.77 | 1.2 | Capture columns |
| | 30 minutes | 0.36 | 1.7 | |
| | 2 h | 0.9 | 2.3 | |
| | 4 h | 0.99 | 1 | |
| Directly following infection* | inoculum | 2.4 | 4.6 | DNeasy Blood and Tissue |
| | 30 minutes | 15 | 7.6 | |
| | 2 h | 12 | 7.7 | |
| | 4 h | 11.1 | 6.9 | |

DNA extraction kits used were all from Qiagen. DNA yield quantified using the Qubit Quant-IT system (Invitrogen). *This infection is one representative assay of three performed under the improved protocol (see text).

With the chance that some eukaryotic DNA might have been carried over in the bacterial genomic DNA extraction, quantitative PCR was performed on the set of Typhi samples from the third infection to determine how much bacterial DNA was present. Two sets of PCR primers known to be reliable were chosen, for the genes *aroC* and *phsA*. The Typhi samples were run with a standard set of dilutions and revealed that, while the inoculum contained the most DNA, very little was present in the samples obtained post infection suggesting that most of the DNA measured by the Qubit system was in fact eukaryotic carry-over.

For this reason, two changes were introduced to the infection assay protocol. Firstly, the infections would be scaled up to improve bacterial DNA yield and enable TraDIS analysis of the results. Secondly, more rigorous wash steps and a DNase I treatment would be used in processing the samples obtained post infection, in order to reduce the amount of eukaryotic DNA being carried through and again improve bacterial DNA yield.

### 5.3.3 Large-scale infection

In the large-scale infection assays, each transposon library was assayed on a separate occasion. The assays themselves were scaled up almost 10-fold, from 3 experimental wells to 29, at each timepoint. The CFU/mL was slightly lower in these assays, but the 10-fold greater total volume generated a much improved yield of DNA.

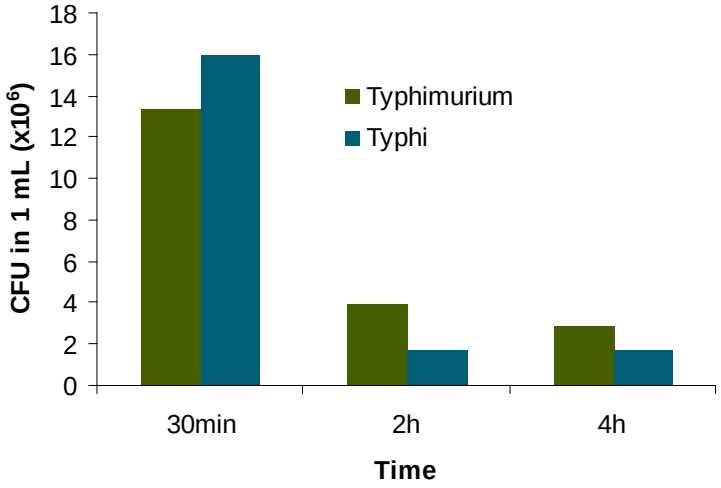


**Figure 5-3 Transposon library cell counts from large-scale THP-1 infection**

Total number of bacteria recovered after timepoints indicated. CFU, colony forming units given in millions per mL.

Despite the wash steps incorporated into the DNA extraction procedure, the amount of DNA recovered from the timepoints post infection suggested significant levels of eukaryotic contamination was still occurring (Table 5-3). Nonetheless, the bacterial yield proved sufficient for TraDIS analysis.

**Table 5-3 DNA yields from large-scale infection**

| Timepoint | Typhi (µg) | Typhimurium (µg) |
|---|---|---|
| Input | 7 | 2.2 |
| 30 minutes* | 124 | 105 |
| 2 hours* | 110 | 45.5 |
| 4 hours* | 78 | 68.5 |

* yield estimated by combining results from one sample split into 3, extracted down 3 columns and quantified as 3 samples (see Methods).

### 5.3.4 TraDIS analysis of large-scale infection assays

The eight samples from the two large-scale assays were each run on a single lane of an Illumina flowcell, which generated between 1.7 and 14.2 million reads (Table 5-4). The lowest sequencing read yields were observed from the Typhi samples obtained post infection. While the sequence reads were not at the same level as the Typhi input sample, there was sufficient information in 1.7 – 3.5 million reads to enable identification of 64,000 – 87,000 transposon mutants. No such issues were observed with the Typhimurium sequencing results.

**Table 5-4 Sequencing output from large-scale invasion assays**

| | Reads | Tagged (%) | Mapped (%) | Inserts |
|---|---|---|---|---|
| **Typhi** | | | | |
| Input | 14,182,028 | 10,557,817 (74) | 7,990,449 (76) | 233,497 |
| 30 minutes* | 3,557,883 | 3,218,768 (90) | 2,771,563 (86) | 87,634 |
| 2 hours* | 2,030,200 | 1,854,321 (91) | 1,657,018 (90) | 97,176 |
| 4 hours* | 1,752,198 | 1,610,693 (92) | 1,459,621 (91) | 64,708 |
| **Typhimurium** | | | | |
| Input | 13,565,707 | 12,457,266 (92) | 7,312,946 (59) | 179,702 |
| 30 minutes* | 14,294,219 | 13,340,388 (93) | 10,537,514 (79) | 46,352 |
| 2 hours* | 13,012,186 | 12,124,834 (93) | 9,633,788 (72) | 43,110 |
| 4 hours* | 10,350,391 | 9,440,397 (91) | 6,411,911 (68) | 25,922 |

Tagged reads are those containing a 100% identical match to the last 10bp of the transposon, indicating a genuine transposon insertion site. The Typhimurium samples were sequenced twice, with very similar sets of results; only one is shown here.

### 5.3.5 Typhimurium

Initial analysis of the TraDIS data comparing the input individually with 30 minutes, 2 hours and 4 hours post infection showed that, in order to quantify changes between timepoints in a rigorous statistical manner, replicates at each timepoint were required. In the absence of these replicates, statistical power was instead increased by treating each of the 3 timepoint samples as independent replicates and combining the data. In this way, we were essentially assaying the genes required for adherence to and invasion of the macrophage, as this was the process each sample had been subjected to. The finer detail of changes across timepoints was lost, but a stronger signal for mutants selected against during infection was observed.

Accordingly, a significant loss of transposon mutants during the infection assay was observed for a total of 316 genes ($\log_2$ FC < -4, *P*-value < 0.001, FDR 6%). The full list can be found in Appendix 8.4.1. These represent candidate genes essential for adherence to and invasion of THP-1 cells.

The candidate on the Typhimurium gene list with the most significant loss of mutants was *surA* ($\log_2$FC = -7.67, $P << 0.001$), which encodes a peptidylprolyl-cis,trans-isomerase involved in energy metabolism. This gene has been previously characterised as a potential vaccine candidate, since it was found to be highly attenuated for virulence in the murine model, and for adherence to and invasion of epithelial cells (Sydenham et al. 2000). Its appearance in this study suggests that *surA* mutants are also defective for adherence/invasion of macrophage-like THP-1 cells. The ligand-binding protein encoded by *sapA* was also identified, as well as the associated ATP-binding protein encoded by *sapD*. The operon *sapABCDF* encodes a peptide transport complex that has been shown to have a role in both antimicrobial peptide resistance and in virulence (Parra-Lopez et al. 1993). Other genes previously associated with a reduced capacity to survive inside macrophages and decreased virulence in the murine include *prc* and *purD* (Baumler et al. 1994), both of which were identified in this study. Together, these provide validation for high-throughput screening as a method to assay genes important for cell adherence, invasion and even survival.

A functional breakdown of the 316 genes is shown in Figure 5-4. When these were compared to the functional breakdown of the whole genome, genes in the categories of pathogenicity/adaptation/chaperones, conserved hypothetical and to a lesser extent information transfer were found to be over-represented (by 3.5%, 3.5% and 1.9%

respectively). Almost one fifth were categorised as conserved hypothetical, indicating that many functions key to cell infection have yet to be elucidated.

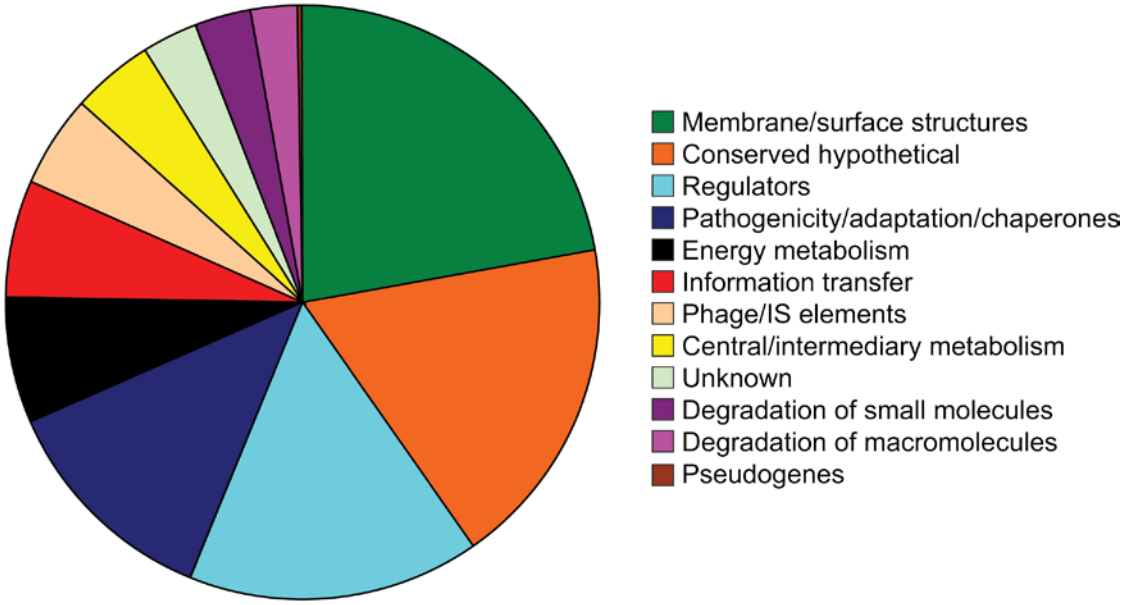


**Figure 5-4 Typhimurium genes essential for macrophage invasion and SCV generation**

Functional categories and colours assigned based on genome annotation.

While the membrane/surface structure genes required for host cell infection were under-represented compared to the whole genome (by 4.3%), fourteen were involved in flagellar regulation, biosynthesis and assembly, including *fliA*, the alternative sigma factor specific for flagellar operons. There were representatives of both SPI-1 and SPI-2 in the pathogenicity/adaptation/chaperone candidate essential genes. From SPI-1, both *iagB* and *spaM*, encoding a cell infection protein and a secreted effector respectively were found to be essential, as well as SL2883, which is part of SPI-1 but has not yet been characterised. Five SPI-2 *sse*/*ssa* genes were also found to be essential, in accordance with this island being known to be required for invasion of macrophages *in vivo*.

In order to determine the metabolic effects these candidate essential genes may have, they were mapped onto StmCyc 4.0 (Figure 5-5). Of note were the multiple genes involved in propanediol utilisation and vitamin B12 biosynthesis and transport, and the number of transporters whose function appeared essential for successful invasion of the macrophage.

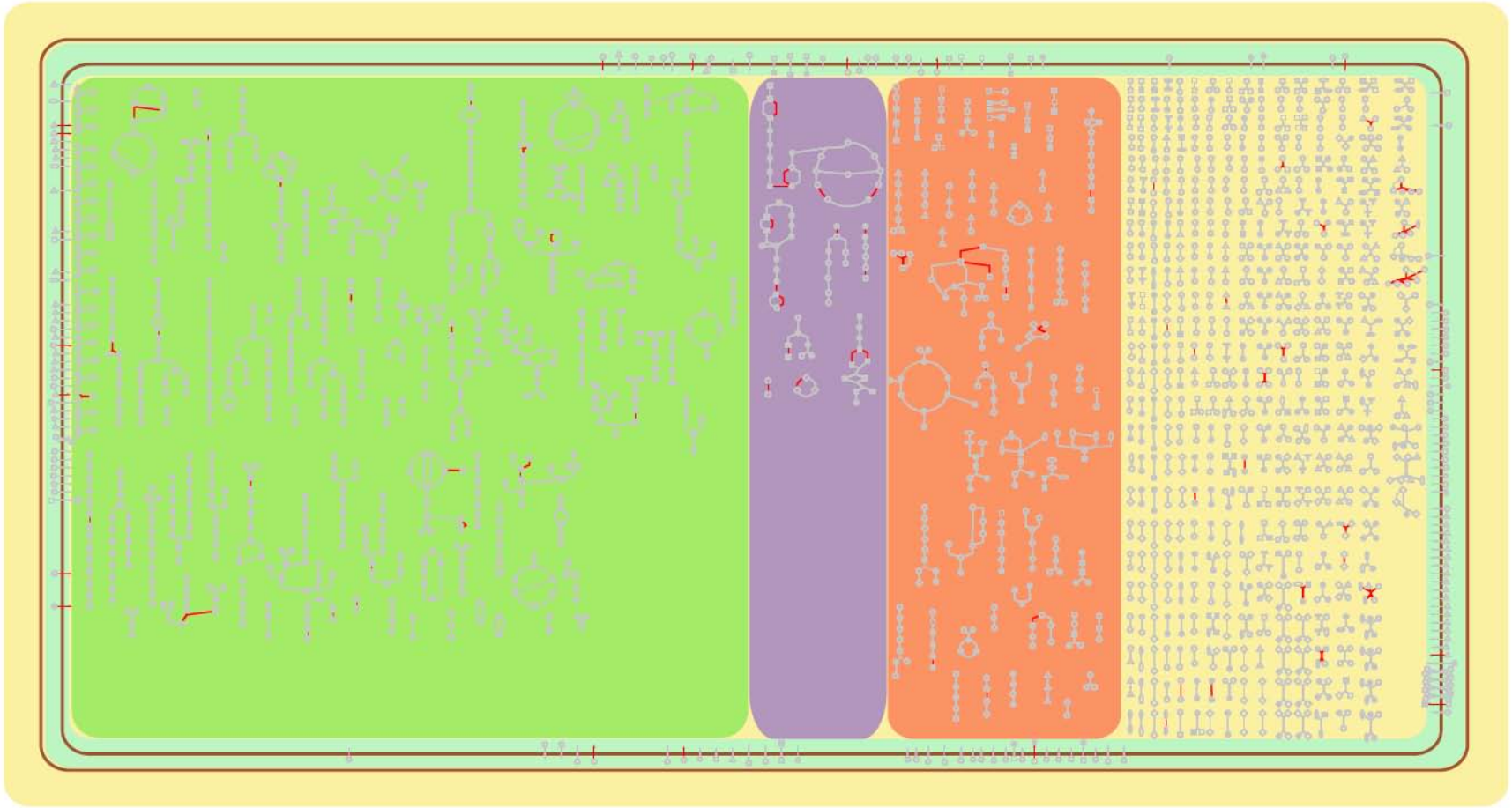

**Figure 5-5 StmCyc view of Typhimurium candidate essential genes for macrophage invasion**

Brown lines indicate bacterial cell membrane; red lines indicate candidate essential genes; greyed out lines represent metabolic reactions. Green background, biosynthetic pathways; dark blue, energy pathways; red, degradation pathways; light blue, transport reactions. Symbols: upward-pointing triangle, amino acids; square, carbohydrates; diamond, proteins; vertical oblong, purines; horizontal oblong, pyrimidines; downward-pointing triangle, cofactors; T, tRNAs; open circle, other; enclosed circle, phosphorylated.

### *5.3.5.1 Propanediol utilisation*

As mentioned in Chapter 3, the utilisation of 1,2-propanediol (PDL), is an anaerobic process dependent upon the synthesis or exogenous provision of vitamin B12. The degradation of PDL was apparently required during the infection assay, since four *pdu* genes were identified on the candidate essential gene list (Figure 5-6).

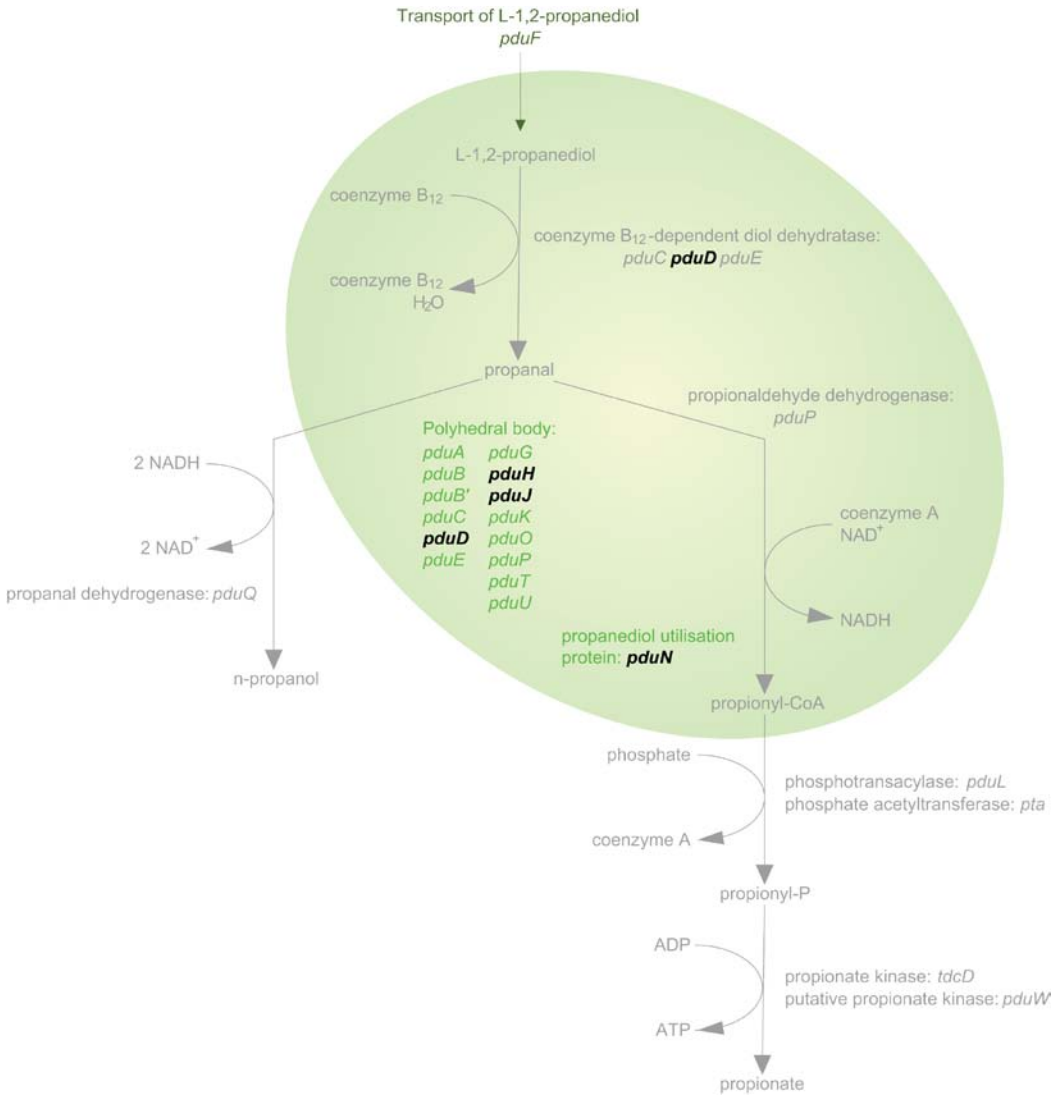


**Figure 5-6 Genes required in propanediol utilisation**

Gene names in black bold type are those identified as candidate essential genes in the Typhimurium large-scale infection assay.

The transport complex for vitamin B12 is encoded by *btuBCDEF*, of which both *btuD* and *btuE* were identified as candidate essential genes for Typhimurium in this assay. In *E. coli*, *btuD* mutants exhibit a defective transport phenotype for vitamin B12, but *btuE* mutants do not (Rioux and Kadner 1989). BtuE is a putative periplasmic protein, and the results of this assay suggest that its role in Typhimurium is required for macrophage invasion, whether this be as part of vitamin B12 transport or some as yet unknown function.

Since vitamin B12 is apparently being taken up from the host cell environment, the requirement for *de novo* biosynthesis could conceivably be reduced. However, the identification of *cbiF*, *cbiO* and *cobS* as candidate essential genes suggests that the metabolic pathways involved in biosynthesis are also required.

#### *5.3.5.2 Transport*

Numerous genes involved in transport reactions were identified as required for successful Typhimurium invasion of THP-1 macrophages. The specific substrates which are transported by these gene products therefore give some insight into the availability of nutrients during and immediately following the infection process. A summary of the genes and substrates transported are given in Table 5-5.

**Table 5-5 Candidate essential transporters**

| Substrate | Type | Genes | Function/family of essential gene |
|---|---|---|---|
| glutamine | ABC | *gln**P**QH* | permease |
| lysine | Secondary | ***lysP*** | APC |
| proline / glycine betaine | ABC | *proVW**X*** | substrate-binding periplasmic protein |
| an amino acid | ABC | SL1563-**6** | unknown |
| spermidine/putrescine | ABC | *pot**A**BCD* | ATP-binding protein |
| PTS carbohydrates | PTS | *pts**H**I*; multiple | HPr protein |
| iron | ABC | *sitAB**C**D* | inner membrane component |
| cobalt | ABC | *cbiMNQ**O*** | ATP-binding protein |
| zinc | ABC | *znuA**C**B* | ATP-binding protein |
| vitamin B12 | ABC | *btuBC**DE**F* | unknown |

Candidate essential genes shown in bold type. ABC, ATP-binding cassette; PTS, phosphotransferase system; APC, amino acid-polyamine-organocation family.

The phosphotransferase system (PTS) histidine protein (HPr), encoded by *ptsH* was found to be essential during the assay. HPr and enzyme I (EI) are both soluble proteins that typically form part of all PTS transporters. The substrate of each PTS is determined by the carbohydrate-specific enzyme II component (EII). Since no EII components were found to be essential, it appears that Typhimurium preferentially utilises carbohydrates during infection, although no particular one is absolutely required.

### 5.3.6 Typhi

The Typhi data was sequenced and analysed after the Typhimurium dataset. Therefore the data from the three output samples were combined immediately for analysis. However, the analysis revealed that only 33 genes were specifically required (Log FC < -2, FDR 3.6%, $P < 0.001$) (Appendix 8.4.2). The assay had been carried out in exactly the

same manner as Typhimurium, so no experimental variation was expected. To be confident of the result, the assay was repeated with 3 independent sample sets of input and the 2 h timepoint. This timepoint was taken to best represent the time after which *Salmonella* should be inside the host cell and have potentially generated an SCV.

All three replicates confirmed that Typhi requires far fewer genes for macrophage infection than Typhimurium, as mutants in only 8 genes were consistently selected against (Table 5-6).

**Table 5-6 Typhi genes required during macrophage assay**

| Gene | Unique ID | Inserts (input) | Inserts (2 h) | $Log_2$ FC | Combined *P*-value | Function |
|---|---|---|---|---|---|---|
| *gidA* | t3465 | 108 | 21 | -1.93 | $2.7 \times 10^{-18}$ | glucose-inhibited division protein |
| *waaP* | t3797 | 43 | 9 | -1.47 | $2.4 \times 10^{-11}$ | phosphorylation of core heptose |
| *thdF* | t3677 | 99 | 30 | -1.46 | $3.9 \times 10^{-11}$ | thiophene/furan oxidation protein |
| *waaG* | t3796 | 41 | 11 | -1.26 | $9.4 \times 10^{-9}$ | LPS glucosyltransferase I |
| *wecB* | t3377 | 49 | 15 | -1.23 | $2.1 \times 10^{-8}$ | UDP-*N*-acetylglucosamine 2-epimerase |
| *dsbD* | t4374 | 105 | 43 | -1.14 | $2.2 \times 10^{-7}$ | thiol:disulphide interchange protein |
| *fkpA* | t4052 | 59 | 21 | -1.14 | $2.1 \times 10^{-7}$ | peptidyl-prolyl isomerase |
| *fimD* | t2317 | 116 | 49 | -1.09 | $6.8 \times 10^{-7}$ | outer membrane usher protein D |
| *rfe* | t3379 | 71 | 29 | -1.05 | $1.6 \times 10^{-6}$ | undecaprenyl-phosphate α-*N*-acetylglucosaminyltransferase |

Total number of insert sites per gene from input and 2 h timepoint were combined from the three replicates and normalised to the values from the first input replicate.

Many of these were in genes relating to outer membrane components; the *waa* genes are both involved in LPS core biosynthesis, and two early steps in enterobacterial common antigen (ECA) biosynthesis are encoded by *wecB* and *rfe*. The role of DsbD is to help fold proteins at the membrane, and correctly folded membrane proteins are needed in

order to adhere to host cells. This is supported by *fkpA*, which encodes a peptidyl-prolyl isomerase (PPIase) that assists protein folding, and is of the same family as *surA* (identified in Typhimurium). Together, these suggest that surface antigens and attachment of Typhi to host cells is important, possibly in triggering phagocytosis rather than invasion.

In *E. coli*, *gidA* and *thdF* have been renamed *mnmF* and *mnmE* respectively. These genes encode proteins that work in concert to maintain the correct reading frame during transcription, by reducing +2 frameshift errors (Bregeon et al. 2001). The authors suggest that *gidA* has further activities, which may be important in Typhi, but since both *gidA* and *thdF* were selected for, it is more likely that the function they perform together is the one required inside the macrophage, for reasons that remain unclear.

This extremely small set of genes apparently required by Typhi for macrophage infection clearly demonstrates the dissimilar manner in which Typhi and Typhimurium enter the cells used in this experiment, while the candidate genes required by Typhimurium suggests utilisation of pathways no longer active in the Typhi genome.

## 5.4 Discussion

The results documented here represent the first genome-wide assay of genes required for infection of human macrophages in serovars Typhimurium and Typhi. Previous studies have concentrated upon understanding which genes are highly expressed or down-regulated during the infection process, which presents its own technical challenges (Daigle et al. 2001; Faucher et al. 2005; Faucher et al. 2006). In this assay, multiple infections were carried out to establish a set of conditions under which each organism in the million mutant pools had an equivalent chance of infecting a macrophage, and enough DNA could be extracted to perform the TraDIS analysis. The nature of an infection assay requires two growth stages of the bacterial inoculum, so to discount the selective effects these might have exerted upon the mutant pools, samples taken post-infection were compared to an 'input' sample taken just prior to infection. The bacteria were also grown in the same cell culture media as the eukaryotic cells, thus reducing the possibility that mutants that were not recovered post infection were unable to respire in the culture media.

The results observed for Typhimurium and Typhi were starkly different, suggesting that these serovars enter THP-1 cells by dissimilar mechanisms. Though it was already known that Typhi is phagocytosed by this eukaryotic cell type (Daigle et al. 2001; Faucher et al. 2006), it appears that Typhi relies entirely upon this mechanism for cell entry. This is unlikely due to the strain used since mutations in the three deleted genes (*aroC*, *aroD* and *htrA*) have been shown not to affect the ability to infect human macrophages (Lowe et al. 1999). Only eight genes were consistently identified as important for cell infection; sufficient mutants in all other genes survived the uptake process and for two hours post-

infection. Of the eight genes, only one has previously been implicated in *Salmonella* virulence; *fkpA* mutants in some Typhimurium strains have shown a reduced ability to survive inside murine macrophages (Horne et al. 1997; Humphreys et al. 2003). The majority of the remaining genes were associated with the bacterial membrane and surface antigens, which suggests that their role lies in adherence to host cells and subsequent phagocytosis. However, the identification of 316 genes important for Typhimurium infection implies that Typhimurium actively invades this cell type.

The genes identified in the Typhimurium assay were from a combination of three timepoints post infection: 30 minutes, 2 hours and 4 hours. Some of these genes had been previously identified as important for survival inside macrophages, rather than invasion. It is possible that either the gene product was required earlier than previously described, or that the signal from some of these genes (e.g. *prc* and *purD*) was in fact strongest from the 2 and 4 hour timepoints. An analysis of single timepoint replicates from 2 hours post infection with Typhimurium is currently underway and will provide further insight.

There is evidence from murine macrophages that with Typhimurium, regardless of the method of entry (invasion/phagocytosis), an SCV is generated that diverges from the typical eukaryotic endocytic degradation pathway (Rathman et al. 1997). It is therefore possible that Typhi, while apparently not invading THP-1s, still generates an SCV after phagocytic uptake. However, the lack of genes identified as required for such a process suggests that the generation of an SCV may occur very slowly, longer than the 2 hour timepoint tested, or is not dependent upon bacterial products. A much later timepoint would need to be tested to confirm this. Whether the genes that are necessary for biogenesis of a Typhi SCV in macrophages would be similar to those in Typhimurium is

unclear, as there is evidence from Typhimurium infection of epithelial cells that SCVs containing non-invasive mutants do not have the same markers as those containing wildtype strains (Gorvel and Méresse 2001). Thus, while the vacuole generated by phagocytosed bacteria may perform the same general function of protection against the host cell environment, structurally it may be somewhat different.

The lack of candidate essential genes for infection of macrophages by Typhi may reflect the level of genome degradation in this serovar. While active invasion is required for infection of epithelial cells, the passive uptake by macrophages may allow certain pseudogenes to be tolerated. This is particularly intriguing when considered in the context of genes required by Typhimurium for macrophage invasion. Two pathways inactivated by pseudogenes in Typhi include vitamin B12 biosynthesis and 1,2-propanediol degradation. As discussed in Chapter 3, these are linked by the absolute requirement for vitamin B12 in propanediol utilisation. However, candidate essential genes involved in both pathways were identified in the infection assay of Typhimurium, an observation supported by experiments performed in murine macrophages (Klumpp and Fuchs 2007). Similarly, genes encoding transporters for proline, cobalt and multiple carbohydrates were required by Typhimurium, but these are also inactivated by pseudogenes in Typhi.

Why all the genes involved in these pathways and transport reactions were not identified as essential in Typhimurium is probably due to a combination of factors. Firstly, there may be redundancy in enzymatic function, which would not be considered essential using this type of assay. Secondly, substrates from elsewhere in the metabolic network may feed into pathways at certain points, alleviating the need for particular enzymes, although their presence may be extremely advantageous. Finally, and most likely, this may be

because the Typhimurium analysis was completed using three datasets taken from different timepoints post-infection. This may have diluted the signal coming from some genes important at the 2 hour and 4 hour stages, such that they did not meet the cutoff criteria. As previously stated, a repeat experiment with multiple Typhimurium replicates taken at 2 hours post-infection should provide greater detail.

It is important to note that a transposon-based assay is one designed to look not for the genes used during a process, but those that are essential for it. There is likely a redundancy in function encoded in the *Salmonella* genome that may account for particular genes found to be important by other methods. Additionally, if mutations in particular genes mean that utilisation of a particular substrate is blocked, other substrates may be used instead. For example, previous studies have indicated that Typhimurium preferentially uses gluconate as a carbon source inside macrophages (Eriksson et al. 2003). None of the necessary *dgo* genes were identified in this assay, but this serves to show that gluconate is not the only substrate upon which Typhimurium can survive inside this cell type. However, the requirement for *ptsH*, a common component of all PTS carbohydrate transporters suggests that the inability to transport at least one of these carbohydrates is detrimental to bacterial survival.

The use of a phagocytic cell line has revealed a distinct difference in the mechanism by which Typhi and Typhimurium infect such cells and suggests that few genes in Typhi are essential to this process. During human infection, Typhi must also infect the epithelial cells of the gut wall, in order to reach the bloodstream. Since these cells are non-phagocytic, Typhi presumably actively invades, likely resulting in a different gene requirement to that found in macrophages. A high-throughput screen of the transposon

library through an epithelial cell line is therefore the next logical step in determining how the metabolic capacity of Typhi affects its ability to infect and survive inside host cells.

## 5.5 Conclusions

Large transposon libraries that comprise over 80 independent mutants per gene provide a useful tool for assaying genes in a high-throughput manner. By making use of the two *Salmonella* libraries, the genes required for infection of a human macrophage-like cell line have been determined. An assay of this nature is difficult to optimise given the variability of the host cells, but a protocol has been established that allows each mutant an equivalent chance of infection, and obtains sufficient DNA post-infection to allow a TraDIS analysis.

The TraDIS analyses from Typhi and Typhimurium infection of THP-1 cells revealed that Typhimurium requires over 300 candidate genes for successful invasion and early survival. A subset of these genes were mapped to metabolic pathways and transporters, and indicates that Typhimurium utilises metabolic capabilities no longer present in Typhi due to inactivation by pseudogenes. Conversely, Typhi requires less than 10 genes for infection, apparently relying on phagocytosis for cell entry, with almost no genes clearly required for early survival inside the host cell.