5 Global transcriptome analysis of an *S.* Typhi *ompR*-null Mutant.

5.1 Aims of this chapter

The aim of the work described in this chapter was to characterise the *in vitro* regulon of the virulence mediating global transcriptional regulator, OmpR. This approach was facilitated by exploiting spotted PCR-based *pan-Salmonella* DNA microarrays and by developing a method for high-throughput chIP-seq.

5.2 Introduction

Bacterial species have evolved mechanisms through which to sense their environment and tightly regulate gene expression in response to specific signals. This dynamic response to changing stimulus has allowed bacteria to survive in their respective niches. Adaptive response studies have revealed many different strategies employed by bacteria to regulate gene transcription and translation. Transcriptional repressors and activators are the most widely studied regulatory systems and are essential for effective cellular function. However, more recently discovered *cis*-acting RNA elements and small RNA are believed to also play an important role. Approximately 227 genes in the *S. enterica* genome are thought to encode regulatory proteins [76,77,136]. Interestingly, many of these conserved regulators have been shown to tightly regulate recently horizontally acquired *Salmonella* DNA islands, not present in *E. coli*. The *S.* Typhi genome encodes 15 putative two-component transcriptional regulators. Each of these regulatory systems harbours both sensing and transcriptional activator domains either as a single, or as two individual, proteins. Transfer of the external signal is normally mediated via exchange of a phosphate group. The autophosphorylating sensing domains is an inner-membrane spanning sensor associated with a histidine kinase domain [217]. Changes in the environment, such as

reduced magnesium or low osmolarity, induce autophosphorylation and transfer of the phosphate group to the receiver domain of the transcriptional activator. Binding of the phosphate group alters the DNA-binding domain affinity to specific DNA motifs. Consensus sequences for most two-component regulators in *E. coli* have been described. The OmpR/EnvZ (*ompB* locus) two-component system is orthologous to the eponymous system in *E. coli* and senses changes in osmolarity [139]. The *S.* Typhi EnvZ protein has diverged from *E. coli* with substitutions in 13 amino acids, however, OmpR is completely conserved between these two species. While maintaining the key porin *ompC* within the regulon, the *S.* Typhi OmpR system has captured the Vi antigen locus [116] and co-regulates expression of SPI-1 [227] and SPI-2 [228]. *ompR*-null mutants are highly attenuated in the murine model [70].

DNA microarray technology has facilitated genome wide analysis of RNA expression in sequenced biological systems. Comparative genomic hybridisation (CGH) and comparative transcriptome analysis have been widely used to identify and quantify nucleic acid content of these bacterial cells. To further characterise the regulon of the *ompB* locus total RNA was isolated from *S.* Typhi under different growth and media conditions and directly compared to an isogenic strain with a defined deletion in the *ompR* gene on the WTSI generation III spotted DNA microarray. These data were confirmed using real time PCR and enzymatic assays to investigate differential protein abundance of the transcribed genes. To confirm OmpR directly regulates the differentially transcribed genes, a method for chIP-seq, to identify protein-binding sites, was developed using high-throughput Illumina sequencing technology and the data mapped back to the entire genome.

5.3 Results

5.3.1 Analysis of the *S.* Typhi OmpR regulon, preliminary work

Since this work was initiated at the WTSI before the CL3 containment facility was established, an attenuated and safe derivative of *S.* Typhi was selected for these experiments. BRD948 is a candidate vaccine strain derived from *S.* Typhi Ty2 that harbours mutations in *aroC*, *aroD*, and *htrA*. These mutations serve as attenuating lesions with *htrA* preventing the detection of bacteria in the blood of orally immunised volunteers. An *ompR* mutant of BRD948 was generated and both derivatives had similar and indistinguishable growth characteristics on LB agar and in LB broth. The DNA microarray used in this study was designed at the WTSI and was generated using specific unique PCR products (200-500bp) representative of the 4097 predicted coding sequences of CT18. Appropriate positive and negative controls were printed onto the arrays. The PCRs were carried out using specific primers (Sigma-Genosys) in a two-step protocol.

5.3.2 Microarray analysis

5.3.2.1 Biological conditions, quality control and analysis

Total RNA was prepared from BRD948 and BRD948 Δ*ompR::kan* cells grown under three different conditions, mid-log (OD_{600}=0.6), early stationary phase (OD_{600}=1.1) and SPI-2 inducing conditions [157]. For each biological replicate, differentially

labelled cDNA was hybridised to four slides and the dye used to label each RNA was inverted for two of these. This was performed for three biological replicates in each condition. To identify systematic error and to assess if the microarray experiments were reproducible the samples were clustered (figure 5.1 (a)). Further analysis was carried out on these samples as the technical replicates generally clustered together. The relationship between the RNA populations isolated from each of the strains was calculated by average-linkage hierarchical clustering using the Pearson Correlation as part of the GeneSpring microarray analysis software V5.0 (Silicon Genetics). Lowess normalised intensities were plotted to determine possible RNA degradation or gDNA contamination (figure 5.1(b)). Differences of less than two-fold were discarded (p-values < 0.05).

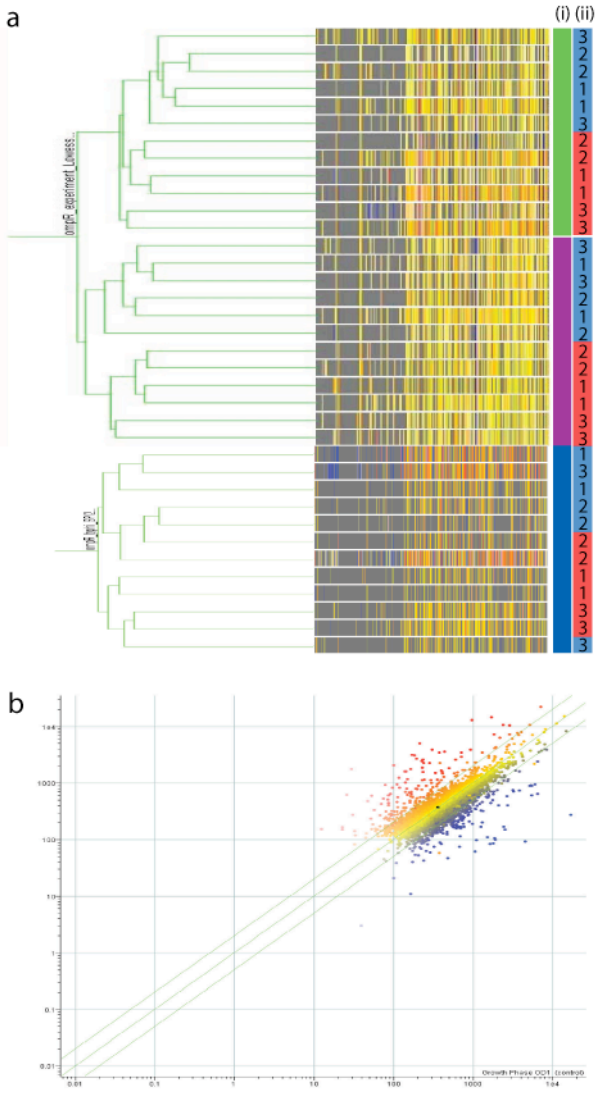

Figure 5.1 Quality control for expression microarrays.

(a) Clustering of microarray intensity data. The vertical colour bars (i) align with each experiment, green is exponential growth, purple is early stationary phase and dark blue represents SPI-2 inducing conditions. The (ii) colour bars represent dye labelling with red, Cy5 conjugated to BRD948 cDNA and Cy3, BRD948 Δ*ompR*. Blue was the inverse and the digits represent each biological replicate. (b) Fluorescence intensity scatter plot. Spots were assigned present or absent during the data extraction. Present spots were normalised using the LOWESS method and the intensity values for control (BRD948) were plotted against BRD948 Δ*ompR*. Data points outside the 2-fold range are clearly visible on this plot.

5.3.2.2 The OmpR regulon during exponential growth (OD_{600}=0.6)

A total of 219 genes had significantly different levels of fluorescence intensity directly related to labelled nucleic acid abundance and differential levels of transcription. Of the differentially transcribed genes, 153 were increased in BRD948 Δ*ompR::kan* (appendix 9.8). Differentially expressed genes were sorted (figure 5.2) according to functional classification as predicted in the *S.* Typhi CT18 sequence annotation. Briefly, for all differentially expressed genes, thirty-five were classified as structural, 82 respiratory, 12 regulatory, 25 transport associated, 29 enzymatic with the remaining 36 genes annotated as hypothetical proteins. The location of each of the differentially transcribed genes is illustrated in figure 5.3 using DNAplot [229]. Genes previously shown to be under OmpR regulation, such as *ompC*, *ompS1* [143], *viaB* locus, *ompB* locus and *sprAB* were all significantly different.

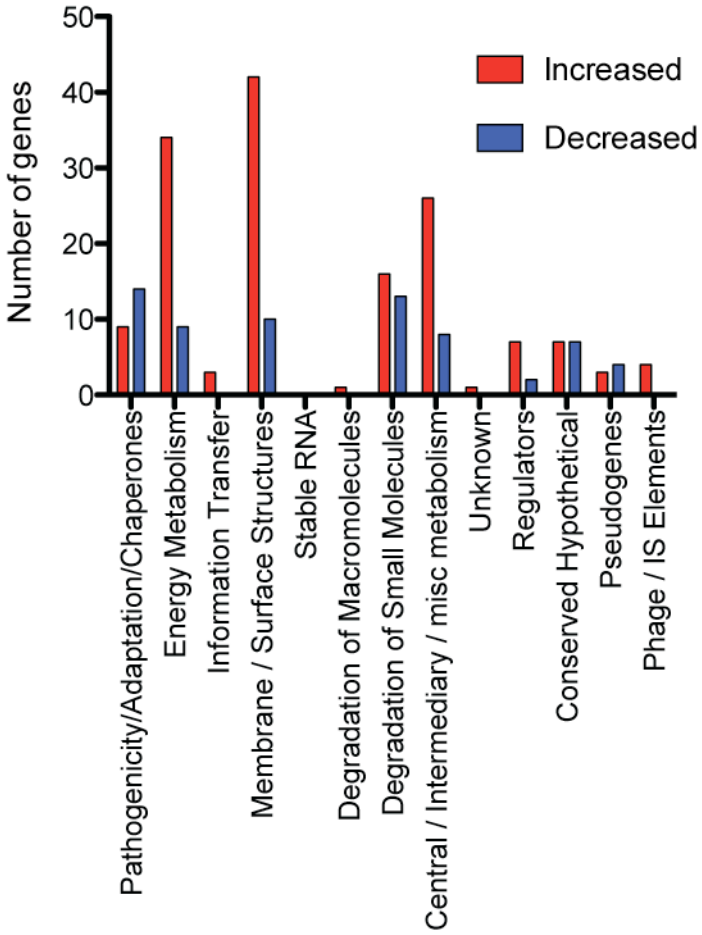

Figure 5.2 Numbers of genes differentially transcribed by functional class.

Genes were grouped according to functional classification. The number of genes in each class with increased transcription is represented by red bars and decreased, blue bars.

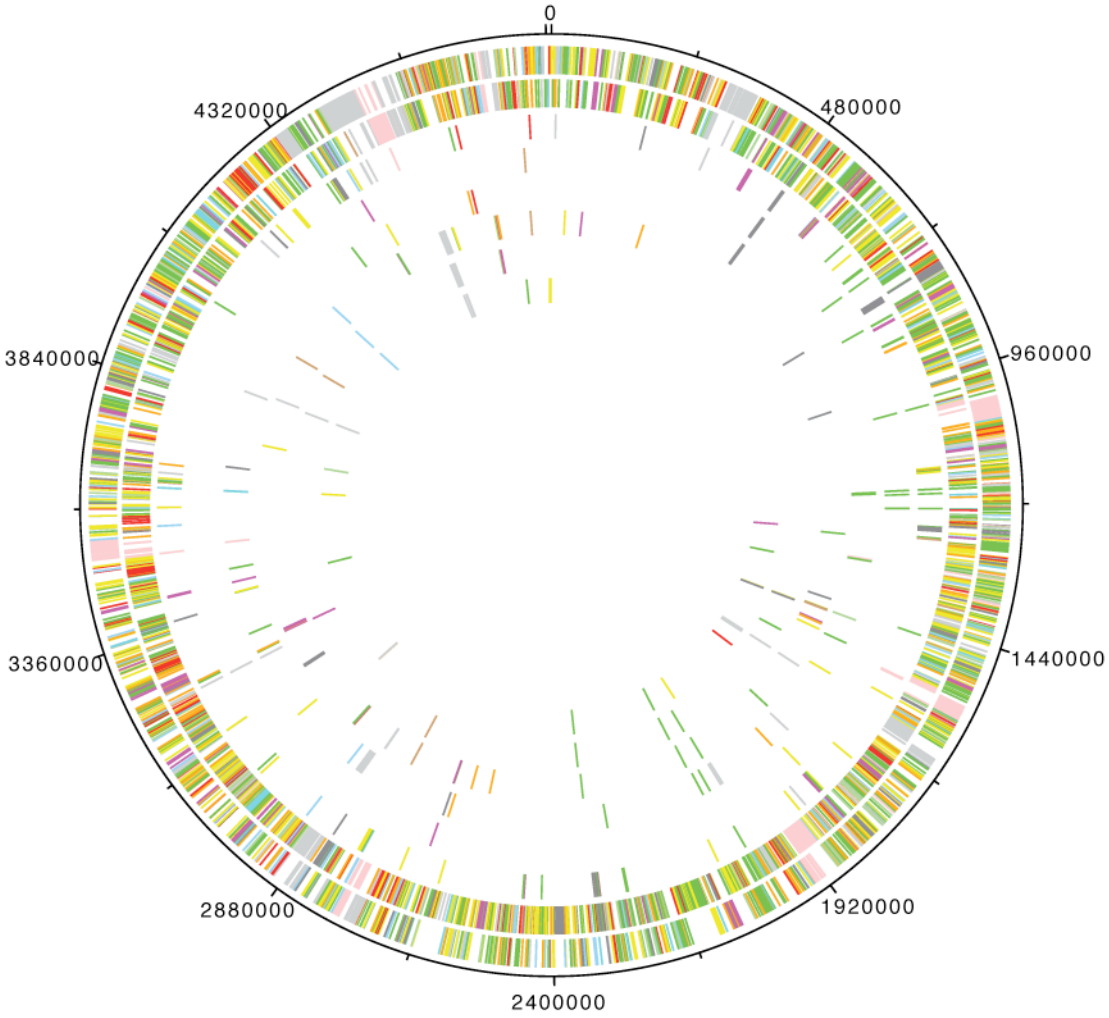

Figure 5.3 Circular plot of the S. Typhi genome and genes differentially transcribed in the *ompR* mutant.

Circles are numbered from 1(outermost) to 8(innermost). 1, forward strand annotated CDS; 2, reverse strand CDS; 3, midlog genes increased transcription; 4, stationary phase increased transcription; 5, SPI-2 genes increased transcription; 6, midlog genes decreased transcription; 7, stationary phase decreased transcription; 8 SPI-2 decreased transcription.

5.3.2.3 The OmpR regulon during early stationary phase (OD_{600}=1.1)

The global fluorescence intensity and quantity of fluorescent spots was reduced in these data, despite using the same mass of total RNA (16µg). With fewer global mRNA transcripts, forty-three genes were identified as being differentially expressed

during early stationary phase in the BRD948 Δ*ompR::kan*, 16 of which were up regulated. Structural genes were the major functional class affected with 19 genes significantly different, including the *viaB* locus, *ompC* and *ompS1*. Other classes were as follows: 9 respiratory, 1 regulatory (*ompR*), 3 transport-associated, 5 enzymatic and 6 hypothetical proteins. Forty-one of the 43 genes were differentially expressed in both the exponential growth phase and early stationary phase (see appendix 9.8 for full gene list and comparison).

5.3.2.4 The OmpR regulon grown under SPI-2 inducing conditions

SPI-2 expression is essential for the formation of and survival within the SCV. Expression of SPI-2 is influenced by OmpR and requires low pH and low Mg^{2+} conditions. Briefly, *S.* Typhi cells were grown to exponential phase, harvested, washed twice and resuspended in SPI-2 inducing media and incubated at 37°C for 20 hours. This was performed in parallel with BRD948 harbouring the plasmid p/c/1, a *lacZ* reporter construct with the promoter region for *ssaG* (from SPI-2) upstream of the *lacZ* gene. Expression of LacZ was confirmed by Miller β-galactosidase assays [181]. The *ssaG*-LacZ reporter was used to confirm that the growth conditions were optimal for SPI-2 activation.

The global fluorescence intensity and quantity of fluorescent spots was reduced in these data in a similar manner to the early stationary phase data. Of the 78 differentially transcribed genes, 54, were increased in expression in the BRD948 Δ*ompR*::*kan*. Forty-two of the genes were classified as structural, 3 transport-associated, 12 respiratory, 12 hypothetical, 8 enzymatic and 1 regulatory (*ompR*).

Interestingly, of the 80 genes differentially expressed in this experiment and the 218 identified during exponential growth, only 20 were identical in both data. Similarly, when compared with the early stationary phase set of 43 differentially expressed genes, only 12 were the same (see appendix 9.8 for full gene list and comparison).

5.3.3 Real time PCR confirmation of expression profile

To validate the expression profile, an experiment to test cDNA abundance was designed using real-time PCR [230]. Briefly, the technique requires quenched fluorescent probes (Applied Biosystems) which bind specifically to the target amplicon. During polymerase activity, the quencher is cleaved from the hybridised probe (by the polymerase) and fluoresces when excited by a LASER. An excess of probes ensures fluorescence increases every cycle and, if the primers are 100% efficient, this intensity should double every cycle. Requisite controls should ensure no contaminating gDNA is amplified (template mRNA prior to reverse transcription), inclusion of a no template control, and a gene, or set of genes, for normalisation. Eukaryote gene expression experiments, utilising real-time PCR, generally normalise to 18S, β-actin or GAPDH which are stably expressed under all conditions. The prokaryotic transcriptome is not thought to contain such stably expressed "housekeeping" genes, therefore, 10 different genes of varied, but essential biological function, were used to normalise to. 16S was not chosen due to the predicted difference in relative abundance to most mRNA transcripts.

A subset of genes that were of biological interest was chosen to interrogate further and potentially validate the array dataset (figure 5.4). From the real time data *sucA, sdhC, ompF, narK, hyaA, slsA, aceA* and *tviB* agree with the microarray data,

however, *fliC, cadA* do not. The *ompC* gene was too variable to determine any difference.

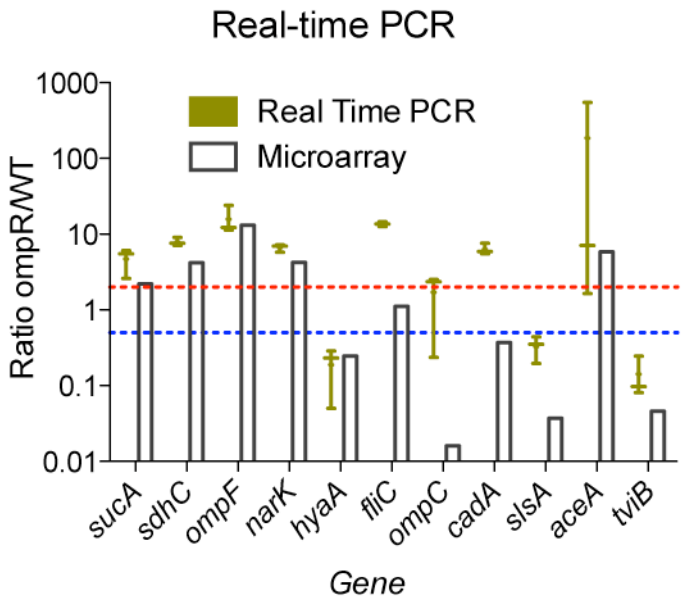

Figure 5.4 Comparison of expression values determined by real-time PCR and microarray.

Real time PCR intensity for each target gene in BRD948Δ*ompR::kan* was compared to the average intensity determined for 10 control genes and then compared to intensity in BRD948.

5.3.4 Succinate dehydrogenase enzyme assay

The succinate dehydrogenase (*sdhCDABsucABCD*) operon, which encodes essential enzymatic components of the TCA cycle [204], has not previously been described as part of the OmpR regulon. Confirmation of the microarray data by real time PCR further confirmed a transcriptional difference between the levels of expression in BRD948 compared to BRD948 Δ*ompR::kan*. In order to determine if this increase in mRNA abundance is translated into functional proteins we performed a succinate dehydrogenase assay [231]. Briefly, BRD948 and BRD948 Δ*ompR::kan* were grown to the same density (OD_{600}= 0.6) as the exponential growth microarray experiment

and harvested. The cells were lysed using a constant cell disrupter and spun (10 000 x g) to remove whole cell debris then centrifuged at 100 000 x g to collect the inner membrane. The pellet was then resuspended, the protein quantified and bound oxaloacetate removed. A colorimetric assay was used to determine the succinate dehydrogenase activity and subsequent kinetics.

The rate of change of absorbance during the colorimetric reaction reveals the enzyme activity, hence, the abundance of the target protein. Direct comparison using the same quantity of inner membrane proteins illustrates a significant difference (p=0.0094) between BRD948 and BRD948 Δ*ompR::kan* (figure 5.5).

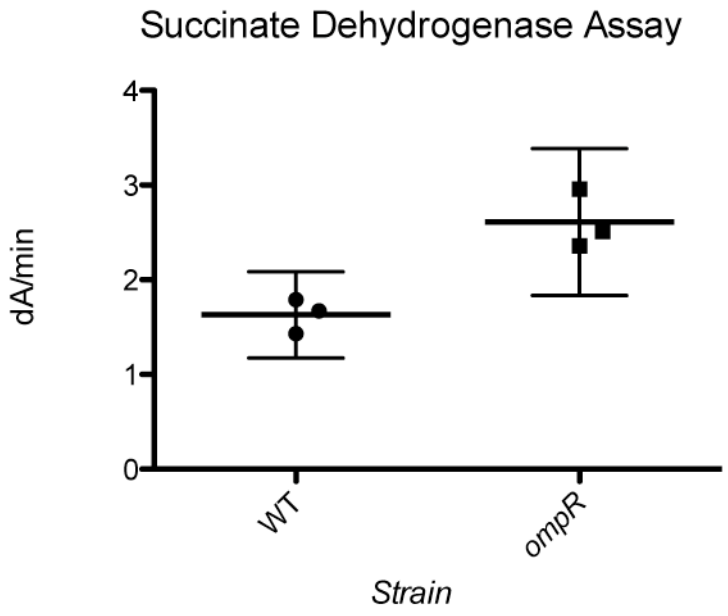

Figure 5.5 Succinate dehydrogenase assay.

Succinate dehydrogenase activity, measured by light absorption, of equal quantites of protein derived from both BRD948 and BRD948Δ*ompR* grow under the same conditions as the microarray experiment.

5.3.5 chIP-seq to determine ompR-binding domains

5.3.5.1 Rabbit anti-OmpR

chIP-seq techniques are dependent on the ability to specifically precipitate the protein/DNA complexes under study. One method by which to achieve this is to use specific antibody to a particular DNA binding protein. Consequently, attempts were made to raise antibody to the *S.* Typhi OmpR protein. In order to express an OmpR antigen suitable for generation of anti-OmpR antibodies the commercially available vector, pBAD202 was used to clone and his-tag *ompR.* DNA encoding the *ompR* open reading frame was generated by PCR from *S.* Typhi BRD948 template DNA. The OmpR protein was expressed in *E. coli* K12 and the purified OmpR antigen was sent to CovalAb (UK) and an antibody raised in and purified from rabbits. Western blotting analysis after pre-absorption of sera in an *ompR*-null mutant fixed culture revealed polyclonal anti-OmpR was not specific enough to distinguish between OmpR and other proteins by western blot.

5.3.5.2 Construction of an OmpR:3xFLAG

An alternative approach to raising antibody is to tag a protein with a specific peptide sequence recognised by a commercially available generic antibody. Consequently, a BRD948 *ompR:3xFLAG* strain was constructed by exploiting a suicide vector to deliver a FLAG-tagged *ompR* open reading frame back *in situ* to the chromosome. Briefly, we cloned a PCR amplicon of the *ompR* open reading frame, designed to incorporate the 3xFLAG tag DNA sequence, by overlap extension PCR (figure 5.6). This sequence maintained translation of OmpR linked to the 3xFLAG peptide and

encoded a terminal stop codon. The sequence after the exogenous stop codon included part of the C-terminal sequence of *ompR* that was predicted to include the Shine-Dalgarno sequence requisite for *envZ* translation in the +1 frame. Homologous recombination was used to replace the wild type *ompR* gene with the FLAG-tagged version in situ on the BRD948 chromosome. To determine the functionality of the OmpR:3xFLAG protein *in vivo* Vi polysaccharide agglutinations were performed and the resulting strain TT53.8 was found to be still Vi positive. Vi expression is lost if the *ompR* gene of BRD948 is inactivated.

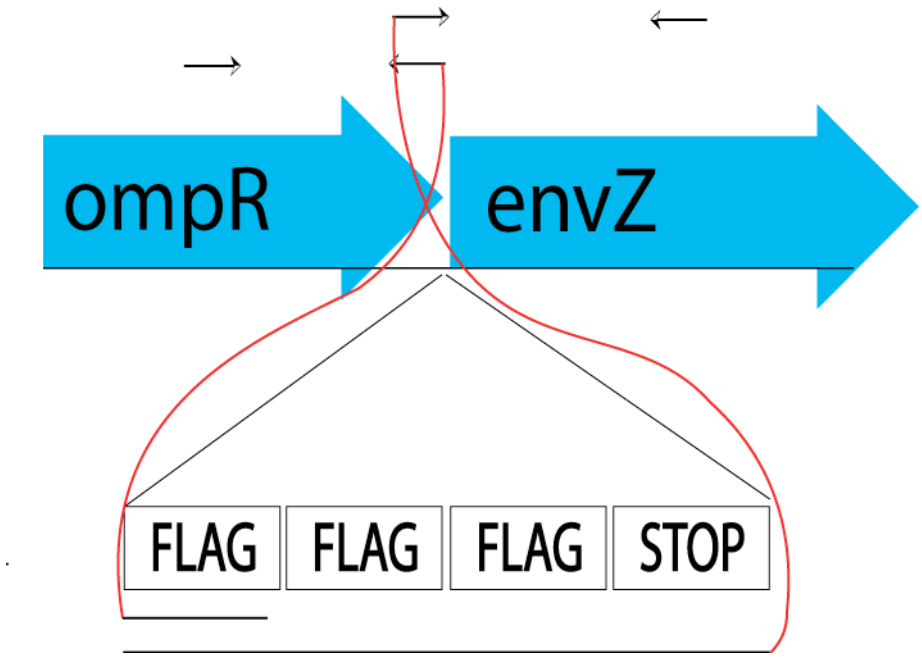

Figure 5.6 Schematic of overlap extension PCR to incorporate 3XFLAG peptide into the C-terminal region of OmpR.

Arrows represent primers homologous to Ty2 regions. Arrows are appended to sequences including 3XFLAG and 20bp overlapping regions. Two rounds of PCR are required to generate the cloning product.

5.3.5.3 chIP-Seq Quality Control

To determine if the monoclonal anti-FLAG antibody was efficacious in binding to OmpR:3xFLAG and able to enrich for the heterologous protein, a western blotting

and immunoprecipitation experiment was performed (figure 5.7). Further quality controls were performed to determine if it was possible to enrich for known OmpR-bound DNA sequences. This was undertaken by performing real time PCR comparison of enriched DNA sequences after immunoprecipitation of WT and the OmpR:3xFLAG. These data show enrichment for the heterologous protein and Illumina sequencing of co-purified DNA identified DNA binding sequences upstream of *tviA* and *ompC*.

Figure 5.7 Western blot of immunoprecipitated OmpR:3XFLAG.

Lanes number from left to right. Lane 1, Seeblue protein ladder (Invitrogen), Lane 2, immunoprecipitated culture using anti-FLAG. No anti-FLAG antibodies bound to BRD948 cell extracts (data not shown). MW for the marker proteins shown in kDa.

5.3.5.4 chIP-Seq using the Illumina sequencing platform

Paired-ended sequence data generated by sequencing DNA co-purified with OmpR:3xFLAG was mapped using the same parameters as previously performed for the transcriptome data without assigning the pileup of reads to each strand. Plots were z-score normalised (R, affy) and the differences between wild-type and ompR:3FLAG generated sequences determined. Plots were then read into Artemis and peakfinder

was used to determine enrichment for OmpR:3xFLAG bound sequences. The peakfinding programme identified 161 peaks. These were then filtered manually, based on the proximity to or location of the peak. Peaks in the middle of a CDS were removed unless there were multiple peaks nearby. 5 predicted DNA binding motifs (z-score > 10) were predicted and mapped back to 43 separate loci within the enriched sequences. Only 5 predicted motifs had a z-score greater than 10. Artemis assigns a score to each peak and this is represented in appendix 9.9 alongside information on the binding sites, number of sites and sequence motif.

Of the 5 putative OmpR-bound motifs, the TGTWACAW was the most common and it is found in all of the known OmpR regulated genes or operons. This motif was identified in 22 enriched sequences, 12 of which were found as a couplet. The couplet binding sites were upstream of *ompR, ompS, csgD, sdhC, mreB, dppA,* a probable phage integrase t4357 that is close to *tviA,* the conserved hypothetical t3244, putative outer membrane protein t3144, *galP, rfbB, fadL* and the conserved hypothetical *yabB*. Notably, upstream of *tviA, ompX* and *ompC*, only one of these sites TGTWACAW is present. The second commonest was AATATATA. This is found in 8 of these sequences and as a couplet upstream of *csgD* and as a singlet upstream of *ompS, stdA, pckA, glpF* and hypothetical proteins t4222 and t1022. The site GTGTSTAY is represented 7 times in these data and only as a singlet. These sites are upstream of *ssrA, cspE, ompR, glpF,* the probable phage integrase t4357 and genes t1320 and *yafK*. A further two sites were identified and these appear in 5 and 1 of the upstream sequences.

Peakfinder identified a series of three peaks upstream of the sigma factors *rpoS* and *rpoH*, fimbrial gene *stdA, hilC* and *glpF*. However, no motif was identified in these

sequences. Sequences upstream of subunits L10, L11, L20, L34 and L36 of the 50S ribosomal protein were also enriched in these data and again no motifs were identified by YMF [184,232]. Interestingly, there were a number of genes identified in the data with a common substrate or function such as L-arginine (*argG, argT, artJ*) and two cold shock proteins *cspE* and *cspA*.

5.3.5.6 Comparison with genes differentially transcribed in microarray data

Of the 127 upstream sequences that were enriched by immunoprecipitation 16 were associated with genes differentially transcribed in the microarray experiment (table 5.1) with 7 encoding the commonest motif (TGTWACAW).

Table 5.1 Genes differentially expressed in microarray experiments and with chIP-seq enrichment sites mapped upstream

Gene ID	Gene Name	Annotation	Expression Ratio	t-test P-value
STY0773	gltA	citrate synthase	3.24E+00	7.68E-11
STY0775	sdhC	succinate dehydrogenase cytochrome b-556 subunit	4.19E+00	1.85E-14
STY1167	-	hypothetical protein	2.03E+02	1.40E-03
STY1925	-	putative membrane protein	2.19E+00	2.64E-10
STY2203	ompS1	outer membrane protein S1	4.06E-01	4.38E-04
STY2424	mglB	D-galactose-binding periplasmic protein precursor	2.11E+00	2.60E-07
STY2493	ompC	outer membrane protein C	1.60E-02	4.20E-06
STY2585	argT	lysine-arginine-ornithine-binding periplasmic protein precursor	3.51E+00	1.24E-06
STY2623	fadL	long-chain fatty acid transport protein precursor	2.11E+00	1.27E-03
STY3049	rpoS	RNA polymerase sigma subunit RpoS (sigma-38)	2.65E+00	1.44E-09
STY3176	stdB	probable outer membrane fimbrial usher protein	2.47E+00	2.92E-11
STY3395	air	aerotaxis receptor protein	2.68E+00	1.59E-08
STY3405	-	probable membrane transport protein	1.12E+01	2.92E-05
STY4168	dppA	periplasmic dipeptide transport protein precursor	2.17E+00	2.50E-07
STY4294	ompR	two-component response regulator OmpR	2.58E-01	1.26E-04
STY4662	tviA	Vi polysaccharide biosynthesis protein	2.32E-02	5.92E-06

5.4 Discussion

5.4.1 Microarray Data

DNA microarrays are a high-throughput platform designed to represent a global snapshot of significant differences in a transcriptional profile. This experimental design offers only to highlight the differences between two conditions and precludes any information on the global transcriptome of the individual strains. Information is also limited to the double stranded DNA sequences present on the array, thus limiting the data to the sequenced strain global annotation, rather than the actual experimental organism transcriptome. It is not expected the genomic DNA would differ significantly in terms of genetic drift between the sequenced Ty2 and BRD948 Ty2. Also, a spotted array does not distinguish between the coding or non-coding strands, potentially increasing the noise, by hybridising any overlapping transcripts. Recently identified small non-coding RNA were not present on this array.

Previous literature on the OmpR regulon is summarised in table 5.2 and a comparison is made with our findings. These data generally agree with the literature. Individual genes such as the outer membrane porins, C and S1, and the *viaB* locus, which are strongly regulated by OmpR, and transcribed under these conditions, are important validations for these data. However, the *ssrA* gene and regulator of SPI-2 is co-regulated by OmpR. Under SPI-2 inducing conditions, there are 4 SPI-2 genes with reduced transcription, suggesting there is dysfunctional transcription of this region further validating the dataset.

Table 5.2 Experimentally verified and published genes in the OmpR regulon of *S.* Typhi or *S.* Typhimurium identified as being differentially regulated in the microarray experiments (2-fold, p<0.05)

			Spotted Arrays		
Gene	OmpR-regulated	Reference	OD_{600} = 0.6	OD_{600} = 1.1	SPI-2 Induced
ompC	Yes	[233]	Yes	Yes	Yes
ompF	Yes	""	Yes	Yes	No
tviA	Yes	[116]	Yes	Yes	Yes
tviB	Yes	""	Yes	Yes	Yes
tviC	Yes	""	Yes	Yes	No
tviD	Yes	""	Yes	Yes	No
tviE	Yes	""	Yes	Yes	Yes
vexA	Yes	""	Yes	Yes	No
vexB	Yes	""	Yes	Yes	No
vexC	Yes	""	Yes	Yes	No
vexD	Yes	""	Yes	Yes	No
vexE	Yes	""	Yes	Yes	No
ssrA	Yes	[234]	No	Yes	No
ompS1	Yes	[144]	Yes	Yes	Yes
ompS2 STY1649	Yes	[143]	No	Yes	No
tppB STY1670	Yes	[154]	No	Yes	No
flhD	No	[146]	No	No	No
csgD	Yes	[235]	No	No	No
osmZ/*hns* STY1299	No	[236]	No	No	No
hilD/*sprB*	Yes	[237]	Yes	No	No
cadC/*cadB*	Yes	[238]	Yes	No	No
aas	Yes	[155,156]	No	No	No

5.4.1.1 Regulators

There are 12 known regulators with differential expression patterns in the *S.* Typhi mid-log experiment. It appears the loss of OmpR could be impacting globally on these operons during exponential growth. Of these twelve genes, most are regulators of stress responses and of anaerobic metabolism. Anaerobic carnitine metabolism, regulated *caiF*, was decreased in expression (0.4, p=2.7E-04), allantoin regulator (*allR*) and the allantoin operon were significantly increased (range 2.0-4.0, all p<1.1E-4) indicating the OmpR deficient organism is exploiting allantoin as a anaerobic nitrogen source [239]. The expression of the two-component regulator for tricarboxylate transport is increased by approximately 4-fold (p=2.6E-11).

Tricarboxylates can be utilised as an anaerobic or aerobic carbon source [240]. Increased transcription of the *rpoS* gene has the potential to complicate this analysis as the RpoS regulon contains more than 50 genes [241]. Expression of the propionate catabolism regulator, *prpR,* is increased as is the propionate operon. This pathway is known as the 2-methylcitrate pathway and allows growth on propionate as the sole carbon and energy source [242]. Transcription of the two-component fumarate response regulator (*dcuR*) is also increased by approximately 4-fold. This regulator activates the anaerobic fumarate respiratory system, transport of fumarate and is activated by C4-dicarboxylates [243].

There are two putative regulatory genes also induced in the *S.* Typhi *ompR* mutant derivative, *rsD* and *ydeW*. In the *S.* Typhi exponential phase growth the SPI-1 AraC-like regulator transcripts, *sprA* (*hilC*) and *sprB* were repressed. These genes have been associated with the regulation of *hilA*, a regulator of SPI-1 mediated cell invasion. Over-expression of these regulatory genes also results in increased ability to enter host cells [94].

In these data, it is unclear whether OmpR is a direct activator or repressor of these regulators, or those indirect effects such as dysfunctional porin expression and changes in inner membrane transport induce anaerobiosis. Such major disruptions to the outer membrane, affecting homeostasis of the cell, may increase demand for less efficient systems such as anaerobic respiration thus indirectly increasing expression of anaerobic related transcripts.

5.4.1.2 Membrane transport

Gram-negative bacteria require transport of nitrogen, carbon, water, electrons and ions across both the inner and outer membranes. The outer membrane acts as a general barrier to osmotic shock and maintains the ability to increase (OmpF) or decrease (OmpC) the porin size depending on the immediate environment [138]. OmpR is known to regulate this system [139]. There are also porins OmpS1 and OmpS2 that have been characterised in the *S.* Typhi OmpR regulon [143,144]. The genes encoding these proteins exhibit some form of deregulation in these data.

With many regulatory genes associated with stress response and anaerobic respiration it is no surprise to see an increase in inner membrane transport activity. During exponential growth of the *S.* Typhi *ompR* derivative there are 11 different transport systems significantly altered (figure 5.8). Allantoin permease increases the cellular availability of allantoin for anaerobic respiration [239]. Molybdenum uptake is induced to increase the amount of cofactor needed for reductases that catabolise nitrogen compounds [244,245]. Small peptide scavenging is also induced, suggesting limited availability of free peptides and recycling of peptides from the cell's outer wall [246,247]. Further evidence to suggest the cell is undergoing anaerobic respiration is the induction of the nitrite extrusion transport system (*narK*). Nitrite is toxic to the cell and a bi-product from anaerobic nitrate metabolism [248]. Glutamate and aspartate transport is also important for controlling cellular nitrogen concentration and is increased in the *S.* Typhi *ompR* derivative [249].

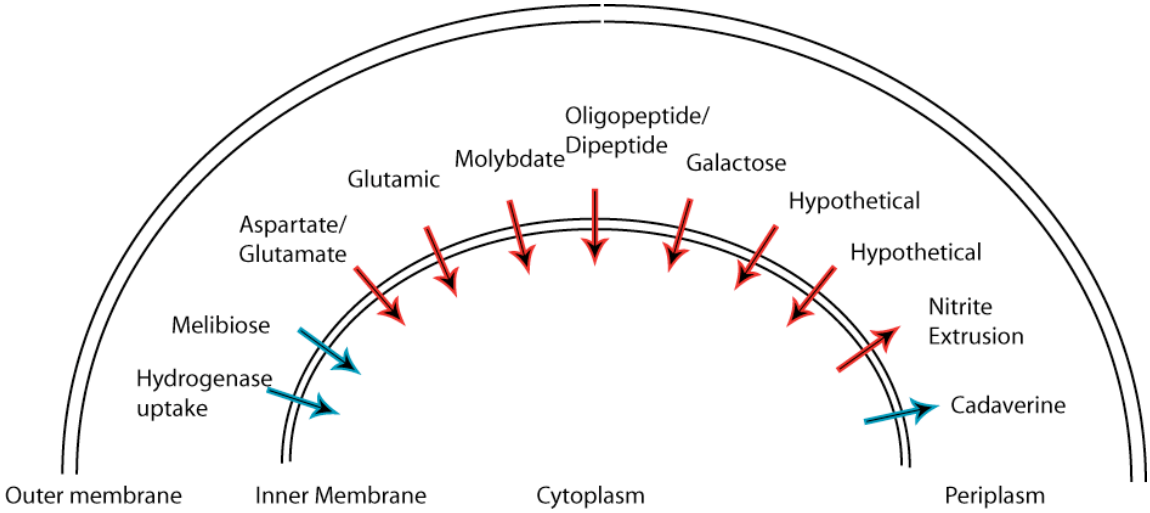

Figure 5.8 Transporter transcript changes.

Schematic diagram representing the shift in inner membrane transport in the *S*. Typhi *ompR*-null mutant. Arrows in red indicate significant increase (2-fold, $p<0.05$) and blue, decrease (2-fold,, $p<0.05$). Arrow represents the predicted direction of substrate.

Cadaverine transport has been associated with blocking of outer membrane porins as well as playing a role in cellular response to low pH by neutralising the external media [250,251]. This may have important implications for the ability of *S.* Typhi to survive within the hostile environment of the stomach and host. The Vi polysaccharide is thought to inhibit the bacterium's ability to induce TTSS mediated endocytosis encoded within SPI-1, however, with cadaverine transport in the same regulon, this may confer resistance to low pH if Vi expression is reduced to invade the host. Unsurprisingly during early stationary phase, very little transport across the membrane is affected and this may be due in part to reduced mRNA transcripts during this period of attenuated growth.

Under SPI-2 inducing conditions there is an increase in transcript of genes required for the scavenging of Mg^{2+} ions (*mgtB, corA*) and an increase in amino acid permease (t3197 (STY3460)). The bacterium is starved of magnesium and essential amino acids in the host vacuole [252]. The two-component sensor, PhoP/PhoQ, which senses the

change from extracellular host to intracellular host environment, is activated by low magnesium, thus evidence of increase in the uptake of this integral cation is not unexpected.

5.4.1.3 Respiration

Respiration is the transfer, or donation of electrons by low redox potential donors, via a range of redox co-factors to the terminal electron acceptor, oxygen. Anaerobic respiration involves reduction of nitrogen oxy anions or nitrogen oxides and is less energy efficient. The energy released in this step-wise process is used to generate a *trans*-membrane electrochemical gradient to generate activation energy for the synthesis of ATP. The ability of different bacterial species to colonise particular hostile niches lies in part in the evolution of alternative strategies for this electron transport mechanism [253].

In the *S.* Typhi mid-log experiment, over half of the genes with significantly different transcription are annotated as genes associated with respiratory processes or are genes encoding enzymes. Figure 5.9 [253] represents the significantly different loci, both increased and decreased, associated with respiration. There is a significant increase in genes that respond to oxygen limitation, such as nitrogen metabolism (*narGHJdsbEccmF1E1C1I, nirDB* and *napCBHGAyojFnapF*) [254,255,256]. Glyoxylate induced genes, such as succinate dehydrogenase (*sdhCDAsucABC*) [204] and isocitrate lyase (*aceABK*) [257] are up-regulated as well as alternative cytochrome genes in the *nrf* locus (*nrfABCD,STY4480*) [258], which is known to be induced under glucose limitation. Interestingly, both aerobic (*fadAB*) [259] and anaerobic (*fadL*) systems for β-oxidation of acyl-CoA to acetyl-coA synthesis [260], a substrate that can be fed into the glyoxylate cycle, are increased in the *ompR* mutant.

This, combined with increases in propionate catabolism (*prpBCDE*) [261], suggests the organism's carbon metabolic programme has been rewired to catabolise poor carbon sources, possibly found in the mammalian gut. The significantly increased allantoin utilisation locus (*allARgipglxRallPybbWallB*) [239] expression appears to remain under regulatory control even though both *allA* and *ybbW* are pseudogenes.

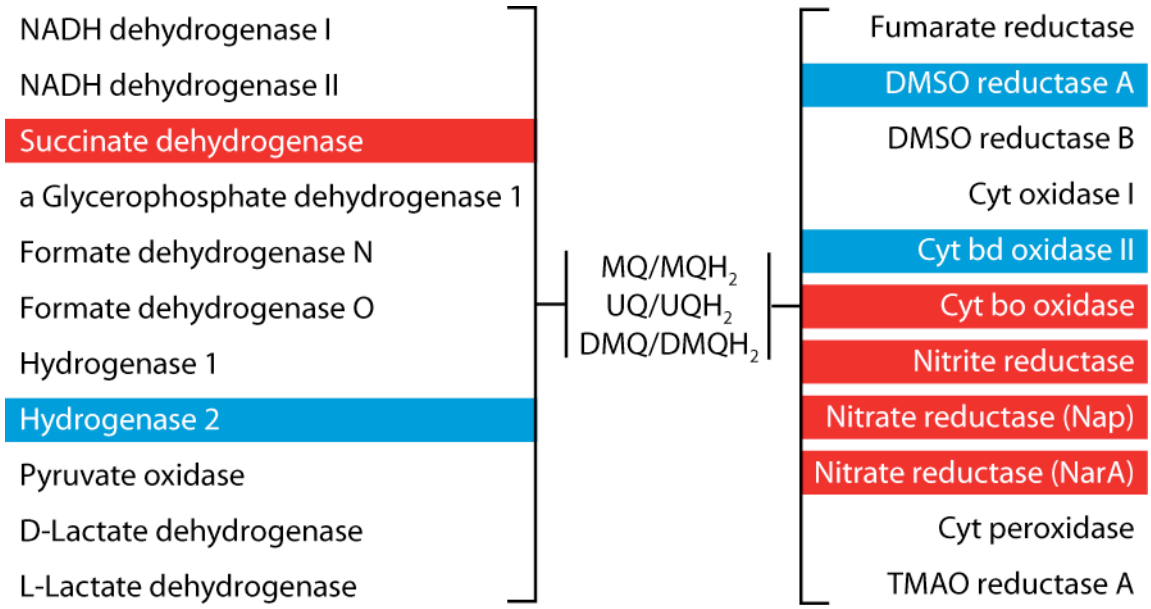

Figure 5.9 Respiration.

Significantly different genes involved in respiration. Red, increased (2-fold, $p<0.05$) and blue decreased (2-fold, $p<0.05$).

Respiratory operons with decreased expression also confer information about the cell's physiological state. Anaerobic dimethyl sulfoxide (DMSO) reduction (*dmsA, dmsA2CSTY1596*) and hydrogenase activity (*hybDCBAO*) also suggest the anaerobic metabolism switch is tightly regulated by OmpR.

The *S.* Typhi stationary phase experiment respiratory genes were not differentially expressed and these may not be expressed at high levels in either strain. However, the data indicate significant increase in transcription of the cytochrome-o biogenesis

pathway genes (*cyoABCDE*) [262]. This is also observed in the SPI-2 induced experiment.

5.4.1.4 Chemotaxis and motility

S. Typhi contains 10 chemotaxis genes involved in regulating motility two of which are putative pseudogenes. In this "stressed" exponentially growing *S.* Typhi *ompR* derivative, transcriptional increase of chemotaxis-associated genes is evident. Significantly different levels of expression in the aerotaxis gene, *air*, represent a natural response in order to translocate to a niche where, the more energy efficient aerobic respiration can be utilised [263]. For cells in early stationary phase, there is evidence for an increase in the flagellar regulation (*flgM*, sigma factor), structural (*flgK*, hook-associated) and biosynthesis genes (*flgN*) [264,265].

5.4.1.5 Genes of unknown function

In these data there are two putative operons encoding genes of unknown function that are differentially expressed (figure 5.10(a) and (b)) in the *ompR* mutant. The genes orthologous to STY1164-1170 are significantly increased in expression and the orthologous genes STY4801-4806 are significantly reduced. These genes may fall within the OmpR "virulence" or "respiratory" regulon or may be regulated due to indirect effects on the cell as discussed previously.

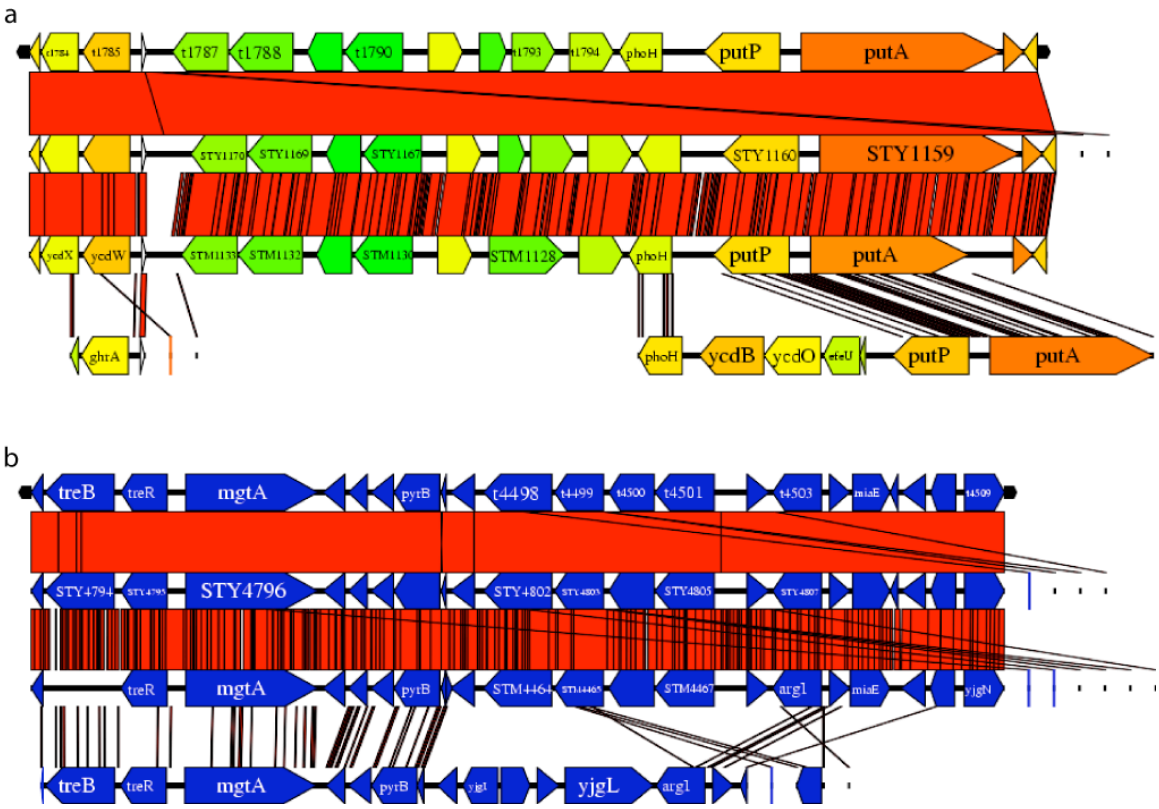

Figure 5.10 Genes of unknown function.

(a) STY1164-STY1170 genes compared with Ty2 (top), *E coli* (bottom), and LT2 (second from bottom). (b) STY4801-4805 genes compared with Ty2 (top), *E coli* (bottom), and LT2 (second from bottom). Both operons were consisted of hypothetical genes and were found to be differentially transcribed in the *ompR* mutant Neither operon is present in *E. coli* as the above Colibase [266] comparison illustrates.

5.4.1.6 Indirect comparison with RNA-seq data

The transcriptome sequencing data (chapter 4) identified 305 genes that were differentially transcribed (2-fold, $p<0.05$) and 39 of these genes were also differentially regulated in the microarray experiment. One of the genes, *yaeG*, was increased in the Illumina comparison and decreased in the microarray comparison and is considered an outlier (figure 5.11).

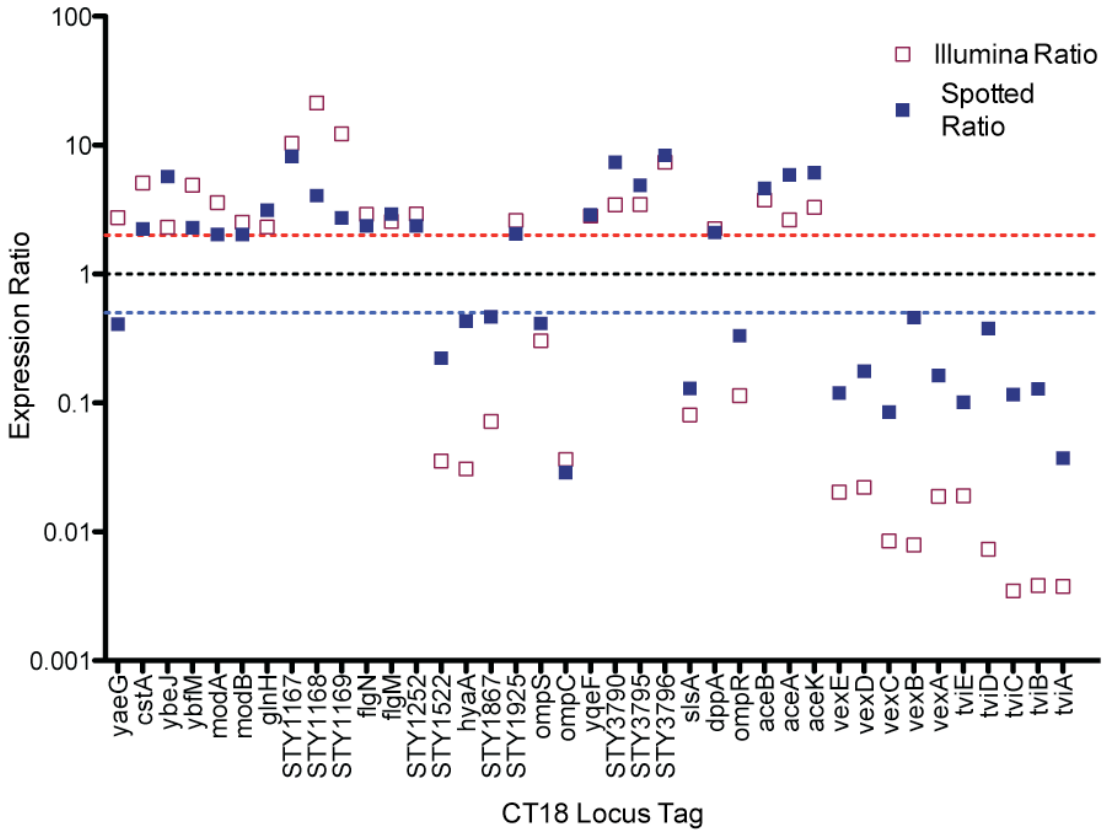

Figure 5.11 Expression ratio for genes found to be differentially expressed in microarray and deep-sequencing experiments.

Spotted array data are represented by filled blue square, Illumina data by red outline square. The red dotted line illustrates a two-fold increase and the blue dotted line, a two-fold decrease in nucleic acid content.

Of the 17 genes that were decreased in expression in both experiments, the Illumina data exhibited a consistently lower ratio of *ompR*/WT transcripts, suggesting background autofluorescence induced during microarray scanning may directly affect the results. Genes previously characterised as OmpR-regulated with decreased levels of expression in the mutant were *tviABCDE, vexABCDE, ompC, ompS*. Genes not previously described in the *ompR* regulon identified in these data include the *slsA* (t3757), *hyaA2* (t1458) and a putative secreted choloylglycine hydrolase (t1459). The function of these genes is discussed in section 4.3.4.3.

Intriguingly, 21 genes were increased in expression in response to the loss of *ompR*, which has previously described as a transcriptional activator. Two contiguous flagellin regulatory genes, *flgN* (t1749) and *flgM* (t1748) were increased in transcription in the *ompR* mutant and are both required for flagellin assembly. FlgM is a negative regulator of flagellar biosynthesis and a mutation is attenuating in *S.* Typhimurium [265]. FlgN is required for the efficient initiation of filament assembly [264]. The glyoxylate shunt genes (*aceBAK)* are also increased in expression; fatty acid catabolism by isocitrate lyase is crucial for macrophage persistence in *Mycobacterium tuberculosis* [267] and *Rhodoccocus equi* [268]. Three genes *t3544*, *t3543* and *t3538* that are predicted components of a ribose/arabinose transport operon were also increased in expression. Furthermore, hypothetical genes *t1788-90* were greatly increased in expression in the *ompR* mutant. These genes are contiguous and encode proteins with sequence identity to sialic acid transport, a secreted protein and a sialic acid lyase respectively and are not present in *E. coli.* Molybdate transport is a crucial co-factor for anaerobic metabolism and the expression of two genes, *modAB,* required for transport are increased in the *ompR* mutant. The genes *dppA, cstA, ybeJ, ybfM, glnH, t1709* were also increased and these encode for proteins annotated as periplasmic dipeptide transporter, carbon starvation response, glutamate transport, putative outer membrane, glutamine transport and a hypothetical protein, respectively.

5.4.2 chIP-seq

The list of genes identified as being downstream of enriched peaks contains previously identified OmpR-regulated genes such as *ompS, ompC, tviA, csgD* and *ssrA*. Many of these genes are preceeded by a preferred DNA-binding motif that is

commonly present as a doublet. This has increased the confidence in predicting genes not previously described as being OmpR regulated.

Combined with the microarray data, this technique has corroborated 16 genes, 7 of which encode the TGTWACAW motif. Three hypothetical genes are included in this dataset, t1790 (hypothetical protein), t1080 (putative membrane protein) and t3145 (probable membrane transport protein).

5.5 Conclusion

The OmpR protein of *S. enterica*, although not essential for *in vitro* cellular survival, appears to be an important regulator of the adaptive response in preparation for invasion of and survival within the host cells. It remains to be fully elucidated whether all the pleiotropic transcriptional effects observed with BRD948 harbouring the *ompR* null mutation are caused directly or indirectly by lack of OmpR binding to consensus sequences, directly activating or repressing transcription. Regardless of this, an interesting transcriptional signature has been identified. Genes of unknown function that are maintained within this regulon and are not present in *E. coli* are interesting inclusions to these regulon data, which are strongly bifurcated into control of anaerobic respiration or virulence determinants.