



ELSEVIER

***Salmonella* transcriptomics: relating regulons, stimulons and regulatory networks to the process of infection**

Arthur Thompson, Gary Rowley, Mark Alston,
Vittoria Danino and Jay CD Hinton

The advent of *Salmonella* transcriptomics has heralded a new era for gene expression analysis of this formidable intracellular pathogen. Increasing numbers of *Salmonella* transcriptomic datasets will contribute to the comprehensive definition of regulons, stimulons and regulatory networks. This task has highlighted the need for sophisticated computational techniques to describe regulatory interactions.

Addresses

Molecular Microbiology Group, Institute of Food Research, Norwich Research Park, Norwich, NR4 7UA, UK

Corresponding author: Thompson, Arthur
(arthur.thompson@bbsrc.ac.uk)

Current Opinion in Microbiology 2006, 9:109–116

This review comes from a themed issue on
Host–microbe interactions: bacteria
Edited by Chihiro Sasakawa and Jörg Hacker

Available online 18th January 2006

1369-5274/\$ – see front matter
© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.mib.2005.12.010

Introduction

Many serovars of *Salmonella enterica* cause gastrointestinal and systemic diseases in animals. Infections with the serovars Paratyphi and Typhi remain specific to humans and are mainly transmitted through contaminated water; annually, there are approximately 16 million cases of the systemic typhoid disease, resulting in 600 000 deaths [1]. Current estimates suggest that *S. enterica* is responsible for up to 124 million cases of human gastroenteritis each year in industrial and developing nations [2]. Most of these infections are caused by the presence of *S. enterica* in the food chain as a result of the widespread colonisation of farm animals.

Following ingestion, *Salmonella* penetrates gut epithelial cells, resulting in enteritis; in a minority of cases, systemic infections result from the uptake of *Salmonella* by phagocytic cells, followed by transmission to the lymph nodes, spleen and liver. Pathogenic *Salmonella* serovars carry sets of horizontally acquired genes, which are often clustered together into *Salmonella* pathogenicity islands (SPIs). These islands encode proteins involved in several

aspects of disease, including the invasion of epithelial cells, host specificity and the survival and proliferation of *S. enterica* in the very phagocytic cells that are designed to kill it [3,4*]. There are five well-characterised SPIs in *Salmonella typhimurium*, with other virulence-related genes being located more sporadically on the chromosome and virulence plasmid [5]. Over the past 20 years, many studies have described the expression of individual *Salmonella* genes in response to external stimuli or under the control of particular regulatory proteins; during this period, a small number of *Salmonella* genes that are induced during infection were identified.

Here, we summarise the transcriptomic analysis of key aspects of *Salmonella* biology, many of which relate to virulence and infection. We also present approaches for the comparison of regulons and stimulons that could help to describe regulatory networks during the intracellular growth of *Salmonella*. In addition, we discuss potential problems that can arise during comparisons of transcriptomic data and suggest possible solutions.

Pretranscriptomic studies

Prior to the development of functional genomic approaches, promoter trapping techniques, such as *in vivo* expression technology (IVET) and differential fluorescence induction, were used to identify genes expressed during *Salmonella* infection [6,7]. One of the early triumphs of this type of approach was the identification of the *spvABCD* virulence genes, encoded on the pSLT plasmid in *S. typhimurium* and required for systemic infection of mice, as well as various *pags*, *mgtBC*, *pmrB* and SPI2 genes, which are induced inside macrophages [8]. SPI2 encodes 44 proteins required for intracellular proliferation and survival. SPI1 encodes around 50 proteins which are important for the invasion of macrophages and epithelial cells [9]. IVET and differential fluorescence induction have been used to identify 35 *Salmonella* genes involved in the infection of macrophages and/or mice [8]. One limitation of these two approaches is that they are extremely labour intensive. Furthermore, unlike microarray analysis, the reporter-based IVET system can only identify genes that show increased, and not reduced, expression. Signature tagged mutagenesis (STM) and transposon mutagenesis have also been used for virulence gene identification. These genetic methods have been used to dissect the process of *Salmonella* infection, colonisation and host specificity, and represent an important complementary approach to transcriptomics [4*,10].

The transcriptomic era

A breakthrough in the monitoring of gene expression came with the advent of microarray technology [11]. The beginning of the 21st century has seen the development of *Salmonella* transcriptomics, and this was triggered by the publication of the complete genome sequence of *S. typhimurium* strain LT2 [12]. More recently, the sequences of the fully virulent *S. typhimurium* SL1344 strain and several other *Salmonella* serovars have been completed (<http://www.sanger.ac.uk/Projects/Microbes>), or will soon be available. The LT2 microarrays were first constructed in 2001 from PCR products that represented every coding sequence in LT2 [13–15]. Centres using *Salmonella* microarrays include the Institute of Food Research and The Sanger Institute in the UK, and the Sidney Kimmel Cancer Center in San Diego, USA. *Salmonella* microarrays can be obtained from The Institute for Genomic Research Pathogen Functional Genomics Resource Center in the USA (<http://pfgrc.tigr.org/>). Oligonucleotide and PCR primer sets for the construction of *S. typhimurium* LT2 and *S. typhi* microarrays are also available commercially from Qiagen and Sigma-Genosys.

Since 2001, most *Salmonella* transcriptomic experiments have been designed to further our understanding of *Salmonella* virulence (Table 1). The experiments fit into three broad categories: expression profiling of *Salmonella* grown in infection-relevant conditions *in vitro*, transcriptomic analysis of global regulatory mutants attenuated for virulence and profiling of *Salmonella* gene expression within mammalian cells. Here, we present a summary of the currently available transcriptomic data in these areas, and discuss approaches for their comparison and analysis.

Defining *Salmonella* regulons and stimulons

One of the strengths of transcriptomics is the ability to completely define regulons and their regulatory networks, particularly with regard to virulence mechanisms. Transcriptomic analysis of *S. typhimurium* has defined the regulons for polynucleotide phosphorylase (PNPase), CsrA, Fis, histone-like nucleoid-structuring protein (H-NS) and phage shock protein Psp, as well as establishing the effect of various stresses, ranging from microgravity to temperature and bile (Table 1). One of the underlying themes of all of these studies is to understand the mechanisms of gene regulation.

Of the regulators that have been studied, Fis is a nucleoid-associated protein which is expressed upon nutrient upshift and the H-NS nucleoid-associated protein is involved in the integration of environmental signals. CsrA is a post-transcriptional regulator which alters the stability of its target mRNAs and controls metabolic processes, as well as motility and invasion-gene expression. PNPase modulates RNA stability by degrading

certain messages and is a component of the bacterial RNA degradosome [16]. Most importantly, all of these regulators were found to influence the expression of varying numbers of virulence genes. The majority of genes that were most strongly activated by Fis have a role in colonisation of the gastrointestinal tract, and include genes from SPI1, SPI2, SPI3 and SPI5 [17••]. Recently, by demonstrating that an increase in temperature from 25 °C to 37 °C resulted in the activation of over 200 H-NS-dependent *S. typhimurium* genes, we showed that H-NS is an essential component of the thermoregulatory system of *Salmonella* [18].

The regulators CsrA, Fis and H-NS were all shown to induce the expression of many SPI1, motility and chemotaxis-related genes [19]. Of the 87 genes that were upregulated in the *pnp* mutant *in vitro*, 51 genes were located in SPI1 and SPI2 [13]. The differing roles of these four regulatory proteins in virulence gene expression are likely to reflect the complexity of control required for successful adaptation of *Salmonella* to the intracellular host environment. We observed that upregulation of flagellar genes generally accompanies the upregulation of virulence genes, suggesting strong regulatory links between the three type III secretion systems in *S. enterica*.

Before invading the host gut epithelial cell, *Salmonella* has to withstand stresses such as anoxia, acid and exposure to bile during its transit through the stomach and intestinal tract. Bile is a component of the innate immune response and is a complex mixture of several bacteriocidal agents, including bilirubin, cholesterol and bile salts. *S. typhimurium* gene expression profiles were defined after exposure to bile [20]. A total of 101 bile-induced and 129 bile-repressed genes were identified, confirming that bile is an important environmental trigger for transcription in *S. typhimurium*. The bile-repressed genes included several genes belonging to SPI1, such as the major SPI1 regulators *hilAC* and *invF*, as well as motility- and chemotaxis-related genes. SPI1 genes were also downregulated by the addition of spent culture supernatant from the probiotic bacterium *Lactobacillus* [21]; these included HilA-dependent SPI1 genes.

Finally, a review of *Salmonella* transcriptomics would not be complete without mentioning the *S. typhimurium* low-shear modelled microgravity (LSMMG) regulon [14,15]. Here, it was found that virulence gene expression was specifically sensitive to gravitational force, and genes located in SPI1 were downregulated. However, it was found that *Salmonella* virulence was increased by LSMMG, as measured by mouse infection and macrophage survival assays. This intriguing study suggests a novel *Salmonella* virulence mechanism, and underscores the remarkably complex regulatory mechanisms involved in *Salmonella* virulence; such mechanisms could be relevant to space travel.

Table 1

Published *Salmonella* transcriptomic datasets in chronological order

Description	Comments	Date published	Reference
Transcriptional profile of PNPase mutant of <i>S. typhimurium</i>	Increased expression of SPI1 and SPI2 genes in <i>pnp</i> mutant. Role in persistency during murine infection	June 2002	[13]
LSMMG stimulon of <i>S. typhimurium</i>	LSMMG regulates expression of 163 genes in diverse functional groups such as transcriptional regulators, virulence factors, lipopolysaccharide biosynthesis, iron-utilization and FUN genes	October 2002	[15]
LSMMG stress response in <i>S. typhimurium</i> is RpoS independent	Used wild-type and isogenic <i>rpoS</i> mutant strains to show that LSMMG induces acid stress response and enhances virulence in an RpoS-independent manner	November 2002	[14]
Transcriptional response of <i>S. typhimurium</i> during macrophage infection	Identified 919 <i>S. typhimurium</i> genes showing altered expression during infection. Almost half of these were FUN genes. Suggested that intracellular environment was low in phosphate and magnesium, and high in potassium	January 2003	[30]
CsrA regulon of <i>S. typhimurium</i>	Reduced expression of SPI1 genes and <i>pdu</i> , <i>eut</i> , <i>cob</i> and <i>mal</i> operons	June 2003	[19]
Transcriptional response to CAMPs in <i>S. typhimurium</i>	Activation of PhoPQ and RpoS virulence regulons and repression of flagella synthesis and invasion-associated type III secretion system (TTSS) genes	October 2003	[22]
Expression profile during swarming in <i>S. typhimurium</i>	Identified genes specific to surface growth and identification of putative new motility and pathogenicity genes	April 2004	[48]
Response to bile by <i>S. typhimurium</i>	Repression of invasion genes, flagellar synthesis and motility genes	April 2004	[20]
Transcriptional profile of <i>fis</i> mutant of <i>S. typhimurium</i>	Role of Fis in expression of metabolic and TTSS-associated genes	July 2004	[17]
Analysis of <i>S. typhimurium</i> PmrA-regulated genes	Identification of six new loci regulated by PmrAB	September 2004	[26]
Transcriptional response to phage secretin protein IV in <i>E. coli</i> and <i>S. typhimurium</i>	Upregulation of <i>psp</i> operon and <i>pspG</i> (<i>yjbO</i>) gene	December 2004	[49]
Comparison of <i>S. typhimurium</i> and <i>Shigella flexneri</i> transcriptomes during infection of macrophages and epithelial cells	Dataset from [30] re-examined in comparison to exponential growth in LB medium	January 2005	[30,38]
H ₂ O ₂ , mitomycin C and temperate phage induction in <i>S. typhimurium</i> and <i>S. typhi</i>	Measured temporal changes in phage expression and noted differential response of <i>S. typhimurium</i> and <i>S. typhi</i> metabolic genes to H ₂ O ₂ and mitomycin C	February 2005	[50]
Transcriptome analysis of <i>slyA</i> mutant of <i>S. typhimurium</i>	Many SlyA and PhoPQ coregulated genes have functions associated with envelope, virulence and resistance to antimicrobial peptides	April 2005	[25]
Comparison of PhoPQ-dependent gene expression in <i>E. coli</i> and <i>S. typhimurium</i>	Limited overlap of PhoPQ-dependent genes in <i>E. coli</i> and <i>S. typhimurium</i> suggests adaptation to pathogenesis in <i>S. typhimurium</i>	April 2005	[28]
Transcriptomic approach showed that H-NS is an essential component in thermoregulation of <i>S. typhimurium</i>	Temperature shift from 25 °C to 37 °C causes H-NS-dependent upregulation of more than 200 genes. Model proposed	June 2005	[18]
Transcript profiling of <i>S. typhimurium</i> in response to spent culture supernatant of probiotic strain <i>Lactobacillus rhamnosus</i>	Identified cluster of HilA-dependent genes that were downregulated	July 2005	[21]
PNPase controls the expression of only nine genes during macrophage infection	<i>S. typhimurium</i> <i>spvABC</i> genes are upregulated in macrophages, suggesting a basis for a growth advantage of <i>pnp</i> <i>S. typhimurium</i> mutant in mice	2006	[33]
The integration host factor (IHF) integrates stationary phase and virulence gene expression in <i>S. typhimurium</i>	Loss of IHF has a profound effect on expression of major virulence genes and upon epithelial cell invasion.	2006	[51]

Interaction between the CAMP, SlyA, PmrA and PhoPQ regulons

Upon infection of a host, *Salmonella* is challenged with a cocktail of antimicrobial agents. These include cationic antimicrobial peptides (CAMPs), which are produced at several infection sites. Using transcriptomic and proteomic approaches, it was shown that exposure of *S. typhimurium* to CAMPs induced PhoP, an important global regulator required for the survival of *S. typhimurium* in a mouse model [22]. The PhoPQ two-component system responds to Mg^{2+} concentrations; it has been hypothesised that PhoQ directly senses levels of CAMP in host tissues, where Mg^{2+} concentrations are relatively low [23].

Sensitivity to CAMPs is also regulated by SlyA, a transcriptional regulator found in *Salmonella* and *Escherichia coli*; mutation of *slyA* renders *S. typhimurium* highly sensitive to the CAMP polymyxin B [24]. Recently, a transcriptomic analysis found a high degree of overlap between the PhoPQ and SlyA regulons and showed that most of the coregulated genes were involved in virulence and resistance to CAMPs. The level of SlyA protein was shown to be almost unchanged in a *phoP* mutant and a *phoPQ* constitutive mutant compared with the wild-type strain during early stationary and stationary phase growth [25^{**}]. Two intriguing models were presented to explain how PhoPQ and SlyA might interact to regulate virulence gene expression.

The regulatory mechanism for resistance to CAMPs is even more complex, as shown by a transcriptomic study which defined the PmrAB regulon [26]. PmrAB is a two-component system involved in the lipopolysaccharide modification that modulates bacterial resistance to CAMPs. PmrAB activity is dependent on PhoPQ, which acts through PmrD post-transcriptionally to activate the PmrAB system [27]. A comparison of the expression profiles of the PhoPQ–SlyA and PmrAB regulons showed no overlap [25^{**},28^{*}]. Overall, these studies suggest that alternative and complex overlapping regulatory networks have evolved to enable survival of *Salmonella* in the intracellular and intestinal environments.

Intracellular *Salmonella* transcriptomics

One of the most fascinating and potentially therapeutically-important applications of *Salmonella* transcriptomics is to improve our understanding of how *Salmonella* adapts and survives in the intracellular environment of mammalian cells. During its vacuolar sojourn, *Salmonella* is presented with many different stresses and has to evade a variety of host cell defence mechanisms, including an initial toxic oxidative burst followed by the production of antibacterial nitrogen species [29].

The first intracellular *Salmonella* transcriptomic study to be published, involved infection of a cultured murine macrophage cell line with *S. typhimurium* [30]. *Salmonella*

RNA samples were taken for microarray analysis at four, eight and 12 hours postinfection. These time points related to specific stages of infection. At four hours, the *Salmonella* are contained within a protective vacuole (the *Salmonella*-containing vacuole, SCV) and are beginning to divide; toxic nitric oxide production begins at eight hours and is well established by 12 hours. This investigation discovered that expression of 919 out of 4451 *S. typhimurium* genes changed four hours after infection of macrophages, compared with *S. typhimurium* grown in cell culture medium. Only a few genes were differentially regulated between four hours and 12 hours, suggesting that initial bacterial sensing controls most alterations in gene expression required for intracellular growth and survival. Particularly striking was the finding that 408 of the 919 genes are function unknown (FUN) genes, which suggests that there is much more to be discovered concerning the intracellular adaptation of *Salmonella* [31].

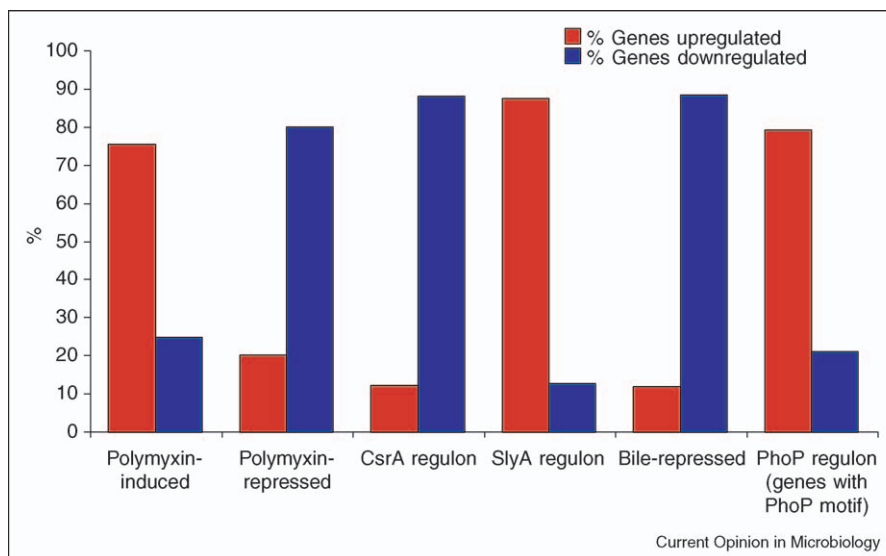
The study also identified several hitherto unknown genes which might have important roles in *S. typhimurium* pathogenesis (STM2839–2841 and STM3117–3120). Other interesting information that emerged from this study was the observation that gene expression patterns can act as sensors of the intracellular environment. For example, genes involved in Mg^{2+} acquisition were highly upregulated, confirming that intracellular *Salmonella* are starved for this ion in the SCV [32]. By analysing expression patterns of genes involved in other uptake systems and biosynthetic pathways, it was deduced that the intracellular environment is aerobic, low in phosphate, high in potassium, rich in amino acids and that gluconate and related carbohydrates might be a principal source of carbon for growth.

The PNPase regulon has uniquely been determined both *in vitro*, in Luria–Bertani (LB) broth, and during infection of macrophages [33]. It was found that the *pnp* mutation changes the expression of nine *Salmonella* genes inside macrophages, whereas expression of 109 genes were altered during *in vitro* growth in LB broth [13]. Notably, the *spvABC* virulence-associated genes were found to be upregulated in the *S. typhimurium pnp* mutant strain in macrophages. This suggests a basis for the persistent infection phenotype of the *S. typhimurium pnp* mutant in mice. The fact that tenfold more genes were regulated by PNPase during *in vitro* growth rather than during macrophage infection will have profound ramifications for the definition of regulons in the future.

New insights from transcriptomic data comparisons

The availability of global transcriptional profiles gives individual published studies a new type of added value for the *Salmonella* research community. For example, once particular regulons have been defined *in vitro*, the relevant genes can be monitored during infection by

Figure 1



The percentage of genes belonging to published stimulons or regulons that are up- or downregulated at four hours postinfection of J774A.1 macrophages with *S. typhimurium*. Comparisons of transcriptomic data were performed using GeneSpring™ (Silicon Genetics). Genes that are induced or repressed by polymyxin or bile were discussed by Bader *et al.* [22] and Prouty *et al.* [20], respectively. The bile-repressed gene list was compiled from genes found to be consistently downregulated at 30 minutes after addition of bile in the three replicate microarray experiments cited in the supplementary data table (<http://genome-www.stanford.edu/microarray>). The CsrA, PhoP and SlyA regulons were from Lawhon *et al.* [19], Monsieurs *et al.* [28*] and Navarre *et al.* [25**], respectively. The PhoP motif is the suggested promoter sequence for the binding of the PhoP protein and is used to define genes that are potentially directly regulated by PhoP [28*]. The probability of all of these percentages of up- and downregulated genes occurring by chance is less than 10^{-3} . The probability was estimated using the binomial distribution and calculated using the `dbinom` command in the statistical package R (<http://finzi.psych.upenn.edu/R/library/stats/html/Binomial.html>).

analysis of the transcriptome defined for mammalian cells. Using this approach, we have now completed the first analysis of the *in vitro*-defined CsrA, SlyA and PhoP regulons, and the bile and polymyxin stimulons during macrophage infection by *S. typhimurium* (Figure 1).

Of 93 CsrA-induced genes identified *in vitro*, we found that 88% were downregulated during infection of J774A.1 macrophages by *S. typhimurium*; consistent with this, *csrA* was also downregulated threefold (Figure 1) [19,30]. We examined the expression of polymyxin-induced and repressed genes in intracellular *S. typhimurium* (data from [22]); 75% of the 381 polymyxin-induced genes were upregulated and 80% of the 373 polymyxin-repressed genes were downregulated during macrophage infection. This comparison underscores the important role of CAMPs in host cell defence mechanisms of phagocytic cells [22]. A similar comparison was performed using the bile stimulon of *S. typhimurium* [20]. We examined the expression of 85 genes that were downregulated after a 30 min exposure to bile to the intracellular *S. typhimurium* transcriptome (data from [20]). Surprisingly, 88% of the 85 genes were downregulated in *Salmonella* from infected J774 macrophages, including several SPI1 and flagella genes.

Of the 41 *S. typhimurium* PhoPQ-activated genes that contained the PhoPQ motif in the promoter region

[28*] we found 79% were induced in intracellular *S. typhimurium*. Because low Mg^{2+} has been shown to activate PhoPQ [34], this probably reflects the low Mg^{2+} environment found in the SCV. In addition, the PhoPQ-activated *mgtBC* genes, which encode a Mg^{2+} transport system, were among the most highly upregulated in the intracellular *S. typhimurium* transcriptomic data.

We also compared the *S. typhimurium* SlyA regulon with the intracellular *S. typhimurium* expression profile; of 24 SlyA-activated genes, we found that 88% were also upregulated in intracellular *S. typhimurium* [25**]. Many of the upregulated genes (78%) were coregulated by PhoPQ [25**]. We found that the expression of the *slyA* gene was unchanged in the intracellular expression data; this is consistent with the suggestion that PhoPQ has no influence on *slyA* expression [25**].

Key points for successful comparisons between different transcriptomic datasets

The comparison of expression levels of genes from different transcriptomic experiments can be problematic but this can be avoided at the experimental design stage. Before any microarray hybridization is performed, it is very important to ensure that a truly representative RNA sample has been purified. The crucial factor is to ensure that bacterial transcription and RNA degradation is

immediately halted at the precise time point that the sample is taken, as a result of the extremely short half-life of bacterial mRNA. In our hands, this is achieved for *Salmonella* with a simple solution of phenol and ethanol [35]. To ensure reproducibility between *in vitro* experiments, bacterial growth conditions should be carefully standardised. For example, we use identical flask sizes, culture volumes and shaking speeds in water baths capable of tight temperature control (± 0.1 °C).

One strategy that aids the comparison of transcriptomic data between experiments is the common reference or indirect approach involving dual dye-labelled microarray hybridizations [36,37]. Here, one of the dyes is used to label a reference nucleic acid, which is used in all experiments. The reference might be cDNA, reverse transcribed from a pool of RNA and used for all experiments, or, more conveniently, genomic DNA from the relevant bacterium. The expression data from different experiments can then be normalised to appropriate controls and subjected to clustering analysis, and compared using a variety of algorithms to form a 'compendium' of related gene expression profiles. The first compendium database of this kind was defined for yeast [38] and a *Salmonella* compendium based on common reference experiments that have defined the major regulons and stimulons, as well as several intracellular growth conditions, is being constructed. This represents a growing resource for examining regulatory mechanisms and transcriptional networks, and defining operon structure.

One of the advantages of comparisons between experiments is that it is possible to choose expression profiles from different growth conditions as a comparator. The expression profiles from the test conditions are normalised to the expression profile of a comparator condition using microarray analysis software such as GeneSpring™ (Silicon Genetics). To avoid erroneous expression levels caused by normalising expression profiles to that of a comparator, an examination of un-normalised transcript levels should always be made to determine whether the data are a true reflection of biologically relevant differential gene expression.

It is also important to consider the precise growth conditions and nutrient availability in the comparator condition because the level of gene expression in the comparator might mask low or high expression in the test condition. For example, the initial analysis of intracellularly regulated genes of *S. typhimurium* during macrophage infection used the expression profile from bacteria grown in RPMI tissue-culture media as a comparator, and identified 919 *S. typhimurium* genes that were up- or downregulated in macrophages [30]. It was subsequently discovered that growth in RPMI media had a dramatic effect upon the *S. typhimurium* transcriptome, including induction of Fe²⁺-regulated genes. Therefore, the intra-

cellular gene expression data were analysed with a new comparator: the transcriptome from bacteria growing exponentially *in vitro* in LB media. This approach identified 1248 *S. typhimurium* genes that were up- or down-regulated during infection of macrophages [39].

Analytical challenges

Transcriptomic data hold promise for developing our understanding of *Salmonella* virulence and regulation; however, we will need to rise to a new set of interesting challenges concerning both experimental approaches and computational biology to make optimal use of the data. It has not yet been possible to determine the transcriptome of *Salmonella* from animal infection models. This is largely because of the problem of extracting bacterial RNA from complex mammalian environments. In addition, the optimisation of cDNA labelling techniques and the development of methods selectively to amplify bacterial RNA in a linear fashion will be important [40].

Computational tools are now being developed to analyse transcriptional regulatory networks. Various clustering techniques based on expression patterns can be used with sequence comparison software to determine upstream regulatory motifs which are useful for defining sets of coregulated genes [41]. Another promising computational approach involves the conceptualisation of complex transcriptional regulatory networks as a combination of a 'parts list' (e.g. transcription factors, binding sites and promoters) and a 'topology model' (a graph of how the parts interrelate); this information can then be used to infer regulatory modules [42]. The parts list can be determined from resources such as BioCyc (<http://biocyc.org>), which makes gene function and operon predictions, and the transcription factor prediction database, DBD (<http://www.transcriptionfactor.org>). DBD detects distant sequence homologies to known DNA-binding domains, and, to date, 151 genomes have been analysed [43]. Obtaining the 'parts list' for *Salmonella* remains a challenge because 30% of the *Salmonella* genome comprises FUN genes [31]; however, DBD has now defined 285, 264 and 420 DNA-binding transcription factors from *S. typhimurium*: LT2, *S. typhi* CT18 and *S. typhi* Ty2, respectively.

Transcriptomic data have been used to identify regulatory modules and module networks from yeast [44], and the Genomica software tool (<http://genie.rockefeller.edu/genomica/index.html>) will enable users to create module maps for *Salmonella*. A further promising computational development is to use gene expression profiles to identify distinct transcriptional subnetworks (called origons); in *E. coli*, it has been shown that origons are responsible for environmental perturbation processing [45••]. Origons might provide further insight into transcriptional control than modulons; all nodes within a modulon must be controlled by a common regulator, whereas in the origon they can be controlled indirectly by diffusion of altered

transcriptional levels through the network [45**]. For transcriptomics to bring insight into the mechanisms that underlie regulatory networks, it is necessary to decipher whether transcriptional effects are direct or indirect, for which the chromatin immunoprecipitation approach is ideal [46*,47].

The next few years will bring a huge increase in the amount of useful information that lies untapped in transcriptomic databases; sophisticated computational approaches such as that described above will prove invaluable in the effort to make the most of these data and to understand *Salmonella* regulatory interactions and the basis for virulence in this complex intracellular pathogen.

Acknowledgements

We thank Dr Rob Foxall at the Institute of Food Research for the probability calculations used in Figure 1; these are available on request. The financial support of the BBSRC Core Strategic Grant is appreciated.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Everest P, Wain J, Roberts M, Rook G, Dougan G: **The molecular mechanisms of severe typhoid fever.** *Trends Microbiol* 2001, **9**:316-320.
2. Herikstad H, Motarjemi Y, Tauxe RV: **Salmonella surveillance: a global survey of public health serotyping.** *Epidemiol Infect* 2002, **129**:1-8.
3. Linehan SA, Holden DW: **The interplay between Salmonella typhimurium and its macrophage host – what can it teach us about innate immunity?** *Immunol Lett* 2003, **85**:183-192.
4. Morgan E, Campbell JD, Rowe SC, Bispham J, Stevens MP, Bowen AJ, Barrow PA, Maskell DJ, Wallis TS: **Identification of host-specific colonization factors of Salmonella enterica serovar Typhimurium.** *Mol Microbiol* 2004, **54**:994-1010.
5. Hensel M: **Evolution of pathogenicity islands of Salmonella enterica.** *Int J Med Microbiol* 2004, **294**:95-102.
6. Valdivia RH, Falkow S: **Fluorescence-based isolation of bacterial genes expressed within host cells.** *Science* 1997, **277**:2007-2011.
7. Mahan MJ, Slauch JM, Mekalanos JJ: **Selection of bacterial virulence genes that are specifically induced in host tissues.** *Science* 1993, **259**:686-688.
8. Rediers H, Rainey PB, Vanderleyden J, De Mot R: **Unraveling the secret lives of bacteria: use of in vivo expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression.** *Microbiol Mol Biol Rev* 2005, **69**:217-261.
9. Hansen-Wester I, Hensel M: **Salmonella pathogenicity islands encoding type III secretion systems.** *Microbes Infect* 2001, **3**:549-559.
10. Lehoux DE, Sanschagrín F, Levesque RC: **Discovering essential and infection-related genes.** *Curr Opin Microbiol* 2001, **4**:515-519.
11. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270**:467-470.
12. McClelland M, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S *et al.*: **Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.** *Nature* 2001, **413**:852-856.
13. Clements MO, Eriksson S, Thompson A, Lucchini S, Hinton JCD, Normark S, Rhen M: **Polynucleotide phosphorylase is a global regulator of virulence and persistency in Salmonella enterica.** *Proc Natl Acad Sci U S A* 2002, **99**:8784-8789.
14. Wilson JW, Ott CM, Ramamurthy R, Porwollik S, McClelland M, Pierson DL, Nickerson CA: **Low-shear modeled microgravity alters the Salmonella enterica serovar Typhimurium stress response in an RpoS-independent manner.** *Appl Environ Microbiol* 2002, **68**:5408-5416.
15. Wilson JW, Ramamurthy R, Porwollik S, McClelland M, Hammond T, Allen P, Ott CM, Pierson DL, Nickerson CA: **Microarray analysis identifies Salmonella genes belonging to the low-shear modeled microgravity regulon.** *Proc Natl Acad Sci U S A* 2002, **99**:13807-13812.
16. Carpousis AJ, Van Houwe G, Ehretsmann C, Krisch HM: **Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation.** *Cell* 1994, **76**:889-900.
17. Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JCD, Dorman CJ: **A global role for Fis in the transcriptional control of metabolism and type III secretion in Salmonella enterica serovar Typhimurium.** *Microbiology* 2004, **150**:2037-2053.
- Using DNA microarrays, combined with β -galactosidase and DNA mobility shift assays, the authors define the Fis regulon of *S. typhimurium*, and identify key roles for this DNA-binding protein in the regulation of central metabolism, virulence and motility gene expression.
18. Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JC, Ladbury JE: **H-NS is a part of a thermally controlled mechanism for bacterial gene regulation.** *Biochem J* 2005, **391**:203-213.
19. Lawhon SD, Frye JG, Suyemoto M, Porwollik S, McClelland M, Altier C: **Global regulation by CsrA in Salmonella typhimurium.** *Mol Microbiol* 2003, **48**:1633-1645.
20. Prouty AM, Brodsky IE, Manos J, Belas R, Falkow S, Gunn JS: **Transcriptional regulation of Salmonella enterica serovar Typhimurium genes by bile.** *FEMS Immunol Med Microbiol* 2004, **41**:177-185.
21. De Keersmaecker SC, Marchal K, Verhoeven TL, Engelen K, Vanderleyden J, Detweiler CS: **Microarray analysis and motif detection reveal new targets of the Salmonella enterica serovar Typhimurium HilA regulatory protein, including hilA itself.** *J Bacteriol* 2005, **187**:4381-4391.
22. Bader MW, Navarre WW, Shiau W, Nikaido H, Frye JG, McClelland M, Fang FC, Miller SI: **Regulation of Salmonella typhimurium virulence gene expression by cationic antimicrobial peptides.** *Mol Microbiol* 2003, **50**:219-230.
23. Vescovi EG, Ayala YM, Di Cera E, Groisman EA: **Characterization of the bacterial sensor protein PhoQ. Evidence for distinct binding sites for Mg²⁺ and Ca²⁺.** *J Biol Chem* 1997, **272**:1440-1443.
24. Shi Y, Latifi T, Cromie MJ, Groisman EA: **Transcriptional control of the antimicrobial peptide resistance ugtL gene by the Salmonella PhoP and SlyA regulatory proteins.** *J Biol Chem* 2004, **279**:38618-38625.
25. Navarre WW, Halsey TA, Walthers D, Frye J, McClelland M, Potter JL, Kenney LJ, Gunn JS, Fang FC, Libby SJ: **Co-regulation of Salmonella enterica genes required for virulence and resistance to antimicrobial peptides by SlyA and PhoP/PhoQ.** *Mol Microbiol* 2005, **56**:492-508.
- Using transcriptomics, these authors observed an overlap between the SlyA and PhoPQ regulons of *S. typhimurium*. Several coregulated genes appear to be specific to *Salmonella*, and the authors suggest that the observed overlap occurred after the evolutionary divergence of *S. typhimurium*, *E. coli* and a common ancestor. The approach demonstrates how microarrays might be used to dissect specific regulons, as well as defining how regulatory networks interact at the transcriptional level.
26. Tamayo R, Prouty AM, Gunn JS: **Identification and functional analysis of Salmonella enterica serovar Typhimurium PmrA-regulated genes.** *FEMS Immunol Med Microbiol* 2005, **43**:249-258.

27. Kato A, Latifi T, Groisman EA: **Closing the loop: the PmrA/PmrB two-component system negatively controls expression of its posttranscriptional activator PmrD.** *Proc Natl Acad Sci U S A* 2003, **100**:4706-4711.
28. Monsieurs P, De Keersmaecker S, Navarre WW, Bader MW, De Smet F, McClelland M, Fang FC, De Moor B, Vanderleyden J, Marchal K: **Comparison of the PhoPQ regulon in *Escherichia coli* and *Salmonella typhimurium*.** *J Mol Evol* 2005, **60**:462-474.
The authors compared the composition of the PhoPQ regulon in *E. coli* and *S. typhimurium* using a combination of expression experiments and motif data. Their results suggest that the PhoPQ system has acquired a specialised function during evolution of these closely related bacteria.
29. Mastroeni P: **Immunity to systemic *Salmonella* infections.** *Curr Mol Med* 2002, **2**:393-406.
30. Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JCD: **Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*.** *Mol Microbiol* 2003, **47**:103-118.
31. Hinton JC: **The *Escherichia coli* genome sequence: the end of an era or the start of the FUN?** *Mol Microbiol* 1997, **26**:417-422.
32. Garcia-del Portillo F, Foster JW, Maguire ME, Finlay BB: **Characterization of the micro-environment of *Salmonella typhimurium*-containing vacuoles within MDCK epithelial cells.** *Mol Microbiol* 1992, **6**:3289-3297.
33. Eriksson S, Clements MO, Rytkönen A, Thompson A, Holden DW, Hinton JCD, Rhen M: **Polynucleotide phosphorylase suppresses virulence gene expression during intracellular replication of *Salmonella enterica*.** *Infect Immun* 2006, in press.
34. Soncini FC, Garcia Vescovi E, Solomon F, Groisman EA: **Molecular basis of the magnesium deprivation response in *Salmonella typhimurium*: identification of PhoP-regulated genes.** *J Bacteriol* 1996, **178**:5092-5099.
35. Hinton JC, Hautefort I, Eriksson S, Thompson A, Mikael R: **Benefits and pitfalls of using microarrays to monitor bacterial gene expression during infection.** *Curr Opin Microbiol* 2004, **7**:277-282.
36. DeRisi JL, Iyer VR, Brown PO: **Exploring the metabolic and genetic control of gene expression on a genomic scale.** *Science* 1997, **278**:680-686.
37. Yang YH, Speed T: **Design issues for cDNA microarray experiments.** *Nat Rev Genet* 2002, **3**:579-588.
38. Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD *et al.*: **Functional discovery via a compendium of expression profiles.** *Cell* 2000, **102**:109-126.
39. Lucchini S, Liu H, Jin Q, Hinton JCD, Yu J: **Transcriptional adaptation of *Shigella flexneri* during infection of macrophages and epithelial cells: insights into the strategies of a cytosolic bacterial pathogen.** *Infect Immun* 2005, **73**:88-102.
40. Mangan JA, Monahan IM, Butcher PD: **Gene expression during host-pathogen interactions: approaches to bacterial mRNA extraction and labelling for microarray analysis.** In *Functional Microbial Genomics*. Edited by Wren BW, Dorrell N: Academic Press; 2002: 137-151. [Bergan T and Norris J (Series Editors): *Methods in Microbiology*, vol 33].
41. Slonim DK: **From patterns to pathways: gene expression data analysis comes of age.** *Nat Genet* 2002, **32**(Suppl.):502-508.
42. Schlitt T, Brazma A: **Modelling gene networks at different organisational levels.** *FEBS Lett* 2005, **579**:1859-1866.
These authors provide a clear discussion of the approaches used, and future challenges in modelling gene networks. The article highlights the gap that exists between network topology and control-logic/dynamic models, due, in part, to the challenge of identifying modules.
43. Madan Babu M, Teichmann SA: **Evolution of transcription factors and the gene regulatory network in *Escherichia coli*.** *Nucleic Acids Res* 2003, **31**:1234-1244.
44. Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N: **Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data.** *Nat Genet* 2003, **34**:166-176.
45. Balázi G, Barabási AL, Oltvai ZN: **Topological units of environmental signal processing in the transcriptional regulatory network of *Escherichia coli*.** *Proc Natl Acad Sci U S A* 2005, **102**:7841-7846.
The authors describe how the transcriptional network of *E. coli*, partly defined with microarray data, is used to define orphans, which represent regulatory subnetworks that originate at distinct classes of sensor transcription factors. The identified features are likely to represent a general framework for environmental signal processing in prokaryotes, which will be important for *Salmonella* biology.
46. Grainger DC, Overton TW, Reppas N, Wade JT, Tamai E, Hobman JL, Constantinidou C, Struhl K, Church G, Busby SJW: **Genomic studies with *Escherichia coli* MeIR protein: applications of chromatin immunoprecipitation and microarrays.** *J Bacteriol* 2004, **186**:6938-6943.
Here, the authors use chromatin immunoprecipitation analysis combined with microarrays to study the global distribution of a transcription factor (MeIR) across a bacterial chromosome. This approach, combined with expression profiles of wild-type and mutant bacteria will be a powerful tool for differentiating direct and indirect regulation by a global regulator.
47. Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJ: **Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome.** *Proc Natl Acad Sci U S A* 2005, **102**:17693-17698.
48. Wang Q, Frye JG, McClelland M, Harshey RM: **Gene expression patterns during swarming in *Salmonella typhimurium*: genes specific to surface growth and putative new motility and pathogenicity genes.** *Mol Microbiol* 2004, **52**:169-187.
49. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, Buck M: **Identification of a new member of the phage shock protein response in *Escherichia coli*, the phage shock protein G (PspG).** *J Biol Chem* 2004, **279**:55707-55714.
50. Frye JG, Porwollik S, Blackmer F, Cheng P, McClelland M: **Host gene expression changes and DNA amplification during temperate phage induction.** *J Bacteriol* 2005, **187**:1485-1492.
51. Mangan MW, Lucchini S, Danino V, Ó Cróinín T, Hinton JCD, Dorman CJ: **The integration host factor (IHF) integrates stationary phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium.** *Mol Microbiol* 2006, in press.