The Primary Transcriptome of Salmonella enterica Serovar Typhimurium and Its Dependence on ppGpp during Late Stationary Phase

Vinoy K. Ramachandran¹, Neil Shearer², Arthur Thompson²*

1 Department of Plant Sciences, University of Oxford, Oxford, United Kingdom, 2 Institute of Food Research, Norwich, United Kingdom

Abstract

We have used differential RNA-seq (dRNA-seq) to characterise the transcriptomic architecture of S. Typhimurium SL1344, and its dependence on the bacterial alarmone, guanosine tetraphosphate (ppGpp) during late stationary phase, (LSP). Under LSP conditions we were able to identify the transcriptional start sites (TSSs) for 53% of the S. Typhimurium open reading frames (ORFs) and discovered 282 candidate non-coding RNAs (ncRNAs). The mapping of LSP TSSs enabled a detailed comparison with a previous dRNA-seq study of the early stationary phase (ESP) transcriptional architecture of S. Typhimurium SL1344 and its dependence on ppGpp. For the purposes of this study, LSP was defined as an aerobic LB culture grown to a later optical density reading (OD\textsubscript{600} = 3.6) compared to ESP (OD\textsubscript{600} = 2.3). The precise nucleotide positions of the majority of S. Typhimurium TSSs at LSP agreed closely with those identified at ESP. However, the identification of TSSs at different positions, or where additional or fewer TSSs were found at LSP compared to ESP enabled the genome-wide categorisation of growth phase dependent changes in promoter structure, the first time such an analysis has been done on this scale. Comparison of the ppGpp-dependency ESP and LSP TSSs for mRNAs and ncRNAs revealed a similar breadth of ppGpp-activation and repression. However, we note several ncRNAs previously shown to be involved in virulence were highly ppGpp-dependent at LSP. Finally, although SPI1 was expressed at ESP, we found SPI1 was not as highly expressed at LSP, instead we observed elevated expression of SPI2 encoded genes. We therefore also report an analysis of SPI2 transcriptional architecture at LSP resulting in localisation of SsrB binding sites and identification of a previously unreported SPI2 TSS. We also show that ppGpp is required for nearly all of SPI2 expression at LSP as well as for expression of SPI1 at ESP.

Background

Infections caused by non-typhoidal Salmonella are one of the most frequent causes of food-borne illness worldwide. Although it is difficult to estimate precisely due to unreported incidents, a recent report suggests that there are 93.6 million cases of gastroenteritis due to Salmonella species globally each year with 135,000 deaths [1].

In the present study we focus on S. Typhimurium, which mainly causes a self-limiting enterocolitis but can cause a more serious disease in immunosuppressed patients [2,3]. Once ingested S. Typhimurium is able to penetrate intestinal epithelium cells, a process that is dependent on the expression of a type 3 secretion system (T3SS) encoded by a horizontally acquired set of virulence genes within Salmonella Pathogenicity Island 1 (SPI1). In the case of immunocompromised humans, S. Typhimurium can become systemic resulting in a typhoid-like fever due to its ability to replicate and survive within macrophages; this is facilitated by the intracellular expression of a second T3SS encoded within SPI2 [4]. In the present study we focus on defining the transcriptomic architecture of S. Typhimurium during stationary phase. Stationary phase occurs when nutrients become scarce and bacteria develop a multiple-stress resistant state. Morphological changes are observed, including rounded shape, loss of flagella and thickening of the cell wall. General metabolism is redirected, macromolecular degradation is increased, and storage and osmoprotection compounds are synthesized [5]. The reorganization of the nucleoid is accompanied by an overall reprogramming of gene expression, much of it mediated by the bacterial alarmone ppGpp and the stationary phase sigma factor RpoS [6,7]. The adaptation to stationary phase starvation and stress also involves the expression of virulence factors including genes encoded within SPI1 and SPI2, several of which have previously been shown to be ppGpp-dependently activated [23,24,26].

Guanosine tetraphosphate is produced by the RelA and SpoT enzymes in all beta- and gammaproteobacteria and whereas RelA only has ppGpp synthetic function, SpoT is able to both synthesise and hydrolyse ppGpp [for reviews see [7–9]]. RelA mediates the stringent response during amino acid starvation and SpoT modulates ppGpp levels in response to a variety of stresses including phosphorus, nitrogen carbon, iron, fatty acid limitation. The effect of ppGpp occurs via direct binding to RNA polymerase...
(RNAP) resulting in the modulation of transcription [10]. Interestingly, it has also been shown that ppGpp plays a key role in coupling virulence to metabolic status in several pathogenic bacteria including Mycobacterium tuberculosis [11,12], Listeria monocytogenes [13], Legionella pneumophila [14,15], Vibrio cholera [16] and Pseudomonas aeruginosa [17]. In S. Typhimurium, virulence in Balb/C mice was shown to be completely dependent on ppGpp due to the presence of SpoT rather than RelA [18]. The transcriptomic architecture of S. Typhimurium at ESP during expression SPI1 and its dependence on ppGpp, has previously been defined using dRNA-seq resulting in the definition of TSSs for the majority of S. Typhimurium genes and the discovery of many new ncRNAs [19,20]. In the present study we used dRNA-seq to define the transcriptomic architecture of S. Typhimurium and its dependence on ppGpp during later stationary phase when the expression of SPI2 encoded genes are elevated, in addition to changes in the expression patterns of many other genes and ncRNAs involved in the transition to a nutrient deprived and microaerophilic environment.

Our previous analysis of the transcriptomic architecture of S. Typhimurium during early stationary phase (ESP) identified the TSSs for 78% of the S. Typhimurium genome [19]. In the current study we define the TSSs for 53% of the S. Typhimurium genome during LSP and show that the majority of LSP TSS positions agree to within 0–3 nt of our previously defined ESP TSSs. However, we found that 13.5% of the LSP TSSs were present at LSP and not ESP or had differently positioned TSSs compared to previously defined ESP TSSs [19]. We report the discovery of 282 new candidate ncRNAs of which 159 were antisense RNAs (asRNAs); five of the asRNAs were located opposite SPI2 encoded genes. Finally, the transcriptomic architecture of S. Typhimurium at stationary phase was found to be highly ppGpp-dependent with 33% of the TSSs for protein coding genes and 29.4% of the candidate LSP ncRNAs being directly or indirectly controlled by ppGpp.

Results
Identification of stationary phase transcriptional start site positions and operons
A dRNA-seq approach was used to identify the precise nucleotide positions of TSSs from RNA samples isolated from an S. Typhimurium parental strain, SL1344, and an isogenic ΔrelAΔspoT strain grown aerobically to LSP in LB. The respective growth curves for the parental and ΔrelAΔspoT strains were shown to be almost identical (Fig. S1). The growth point (OD₆₀₀ = 3.6) and time at which samples were taken for dRNA-seq analysis are indicated in Fig. S1. A description of how dRNA-seq (differential RNA-seq) was used to define TSSs position is described in [19,21]. Briefly, the procedure involves preparing two cDNA libraries from the same RNA sample, one of which is enriched for primary transcripts by treating with terminator exonuclease which specifically degrades processed transcripts. Comparison with the untreated library reveals the location of TSSs due to the elevated read numbers of transcripts from this library compared to the treated library. The mapped reads from this study and also from the ESP dataset [19] can be viewed on JBrowse (http://jbrowse. molmicro.org/) by following the link provided at www.ifr.ac.uk/Safety/MolMicro/.

We mapped a total of 2186 LSP TSS’s on to the SL1344 chromosome (including all annotated ORF’s, stable RNAs and ncRNAs and a total of 78 for the SLP1–3 endogenous plasmids (Tables S1, S2, S3, S4, S5). The TSSs were categorised as primary, secondary, internal or present only in the ΔrelAΔspoT strain; however many had multiple associations and these are defined and summarised in Fig. 1. Primary LSP TSSs were identified for a total of 2530 mRNAs (including intra-operonic genes) representing 53% of the annotated SL1344 genome (Genebank ID FQ312003.1; Table S5). The proportion of genes for which LSP TSSs were identified was lower than was previously defined for ESP (78%) [19], most likely due to restricted growth and metabolic activities at LSP; however a comparison of the TSS positions for mRNAs found at ESP and LSP revealed that 86.5% (1576) of LSP TSSs were either identical or within 1 to 3 nt of the previously defined TSSs positions for ESP (Table S1) [19]. Of the remaining 13.5% of the LSP TSSs, 6.0% (111) were specific to LSP (i.e. found within genes for which no ESP TSSs had previously been defined), and 7.5% (134) were repositioned by 10 nt or greater (median = 63 nt) at LSP compared to ESP (Table S2) [19]. The accuracy of dRNA-seq for the identification of TSSs (92% of ESP TSSs were located within 10 nt of experimentally determined TSSs [19]) suggests that the repositioning of 7.5% of the LSP TSSs relative to their ESP TSS positions was not due to experimental inaccuracies. A large number (924) of the chromosomal TSSs and 23 SLP TSSs that had previously been found at ESP were not present at LSP, nor was there any alternative LSP TSS, suggesting the corresponding genes were not transcribed at LSP or were subject to other regulatory mechanisms (Table S7). A previous DOOR-based prediction of operon structure inferred from ESP dRNA-seq data revealed 625 predicted operons [19]. Here we confirmed the structure of the predicted ESP operons and in addition defined 46 operon structures specific to LSP (Table S1); Finally, it has long been known that the majority of TSSs start with a purine residue, and analysis of the TSS located upstream of annotated LSP ORF’s (Table S1) revealed that 71% of the transcripts started with a purine residue (A - 46%, G - 25%), in agreement with the known preference for a purine residue at the + 1 position [22] (data not shown).

Transcriptomic architecture of SPI2 during late stationary phase
SPI2 is expressed during intracellular replication of S. Typhimurium within host cells [4]. Whist stationary phase culture in LB is unlikely to represent a physiologically apt model of the intracellular environment, we nevertheless observed that the expression of the majority of SPI2 but not SPI1 genes were elevated under LSP conditions compared to ESP; this expression pattern has previously been noted and was also verified from a microarray based transcriptomic analysis shown in Fig. S6 [19,20,23–26]. The consecutive expression pattern of SPI1 followed by SPI2 expression during transit through stationary phase (Fig. 2 and Fig. S6) mimics the infection situation whereby SPI1 largely precedes SPI2 expression and they rarely occur simultaneously in the same niche [27]. The elevated expression of SPI1 (but not SPI2) at ESP was defined by the increased reads mapping to the SPI1 locus (Fig. 2AC; data from [19]). Conversely, elevated SPI2 expression at LSP (but not SPI1) is shown in Fig. 2CD, confirming previous findings [23–26], and verified in Fig. S6. During both ESP and LSP, Fig. 2 and Fig. S6 also clearly show that SPI1 and SPI2 expression were dependent on ppGpp for activation. The expression of SPI2 during LSP enabled us to undertake an analysis of the transcriptomic architecture of SPI2 under these conditions.

Due to its importance in Salmonella pathogenicity, SPI2 has been widely studied in terms of its organisation and regulation and in the role of individual effector proteins. However we were able to gain new insights into features of SPI2 regulation from a comparison of our dRNA-seq data under SPI2 inducing conditions with previously
published work. Analysis of the dRNA-seq data confirmed a TSS 167 nt upstream of the ssrA translational start at genomic position 1436618 (only 1 nt distant from a previously reported ssrA TSS [28], and in addition, confirmed the position of a second previously reported TSS at genomic position 1436769 (Fig. 3CD, [19]). No TSSs could be identified immediately upstream of ssrB or orf242 (SL1324), although sequencing reads extended into both ORFs implying that under the growth conditions used here, they were co-transcribed with ssrA (Fig. 3C). Further examination of the dRNA-seq data identified TSSs upstream of ssaB, ssaG, ssaM and ssaR (Fig. 3ABD, Table S1). The TSSs for ssaB, ssaG and ssaM agree closely with those determined experimentally (Fig. 3D) [28,29]. The ssaR TSS has not previously been identified although promoter activity and SsrB binding in the ssaR upstream region has been reported and a consensus SsrB binding site was identified upstream of the ssaR TSS [30] (Table S6).

The sseA TSS has previously been mapped at two adjacent sites at genomic positions 1440403-4, [29], (Fig. 3D). We were unable to clearly identify a TSS for sseA at this position in the LSP dRNA-seq data, however, according to data from ESP, an additional TSS for sseA was identified within the 3’ end of the ssaE coding region at genomic position 1440133, which lies immediately upstream of sseI [19], (Fig. 3AD). We also note an SsrB binding site 36 nt from this TSS, shown in Table S6. In the LSP dRNA-seq data a sudden increase in transcript levels was observed within a few nt’s of the TSS internal to ssaE (Fig. 3AD), making it highly likely that this is the TSS for ssaE; however, the precise position of the TSS may be obscured in the LSP sequencing data, possibly due to interaction with other regulatory mechanisms. Alternately, or in addition, the transcript levels of the ssaE operon may be being affected by a ~150 nt transcript which was observed anti-sense to the ssaE-sseA intergenic region (Fig. 3A); we are currently investigating this possibility.

The SPI2 operons are all activated by SsrB and yet many SPI2 encoded genes are also controlled either directly or indirectly by multiple ancestral proteins [31]. Knowledge of the precise location at which SsrB binds to its dependent promoters in relation to RNA polymerase will facilitate research on the interaction between the different regulatory components. Where they have been mapped, SsrB dependent promoters have been found to be diversely organised with SsrB binding to sites upstream, overlapping or downstream of the TSSs of different genes [29]. Although for some genes it has been noted that SsrB can directly activate transcription by binding immediately upstream of their promoters [29], a full analysis of SsrB dependent promoters has been limited by the number of targets for which both SsrB binding site and TSS information was available. The dRNA-seq analysis presented here and published previously [19] has extended the number of SsrB dependent promoters for which TSSs have been mapped. We combined this information with previously published TSSs and SsrB binding sites [30] to map the promoters of all the SPI2 operons and 14 additional SsrB regulated operons encoded outside of the SPI2 locus (Table S6). This analysis revealed that the majority of SsrB binding sites were located immediately upstream to, or overlapping a region 35 nt upstream of the TSS (Table S6). At promoters where SsrB does not appear to bind close to the −35 region it is also possible that SsrB activates transcription to initiate at alternative TSSs used only under certain growth conditions; for example, we identified a weak TSS signal 34 nt downstream of a consensus SsrB binding site at the sseI promoter, although the

![Figure 1. Annotation of TSSs.](image-url)
previously reported TSS is located 100 nt upstream of the SsrB binding site (Table S6).

Comparison of ESP and LSP TSS positions

The LSP mRNA TSSs were compared to our previously identified ESP TSSs ([19], Tables S1, S2). Although a close correlation was found between the majority of ESP and LSP TSS positions, we found 134 TSSs which had different positions at LSP compared to ESP (10 nt); these will be discussed in the next section of the results. In addition we found 111 LSP-specific TSSs for which no previous ESP TSSs had been identified, [19], Fig. 4, Fig. S2 and Table S2). The 111 LSP-specific TSSs were found within the promoters for 105 genes. A survey of the functions of the 105 genes revealed that several were anaerobically induced and involved in fermentative metabolism (adhC, citA, fucR, napB, fdoG, pflD), as might be expected to occur in a stationary phase microaerophilic environment. At LSP, lack of nutrients is also a limiting growth factor and this was suggested by the presence of LSP TSSs for genes encoding gluconate, glucose and iron transporters (glnR, dhaT, ptsG, entC, sitA). A low phosphate and magnesium environment is suggested by the presence of LSP TSSs for pstB, which encodes a phosphate transport ATP-binding protein [32], and rstB which is the sensor kinase in a two component system (with RstA) [33]. Expression of rstB is induced under low magnesium growth conditions through the PhoPQ two-component system [34]. The RstA/B regulon is induced at stationary phase and includes genes involved in acid tolerance, curli fimbria formation and anaerobic respiration [33]. RstAB and PhoPQ are thought to form a signal relay cascade involved in the adaptive response to acid conditions [35]. RstA expression is also able to promote RpoS degradation in S. Typhimurium and has been shown to be involved in modulating Salmonella biofilm formation [36]. The low pH, phosphate and magnesium environment implied at LSP by the identification of TSSs for pstB and rstB has also been shown to trigger activation of SPI2 under minimal media growth conditions [25,37–39].

Promoters with differing TSSs at LSP and ESP

A comparison of the TSSs at LSP and ESP found within the same promoters was performed and revealed that 134 TSSs either differed in position (>10 nt) and/or were found to have fewer or more TSSs at LSP compared to ESP (Fig. 4). For ease of reference the different types of rearrangements of TSSs between LSP and ESP, along with the numbers of genes associated with each type of rearrangement were categorised alphabetically (Fig. 4; Fig. S2 also shows rearrangements that occurred in one or two genes only). The genomic positions of the TSSs at LSP, relative to ESP and their corresponding genes and function, are shown in Table S2. In summary, we found the majority of differences between ESP and LSP TSS positions occurred where one or more of the ESP TSSs were not present at LSP (274 TSSs) and these constituted types A, C, E and H shown in Fig. 4 (and M, N in Fig. S2). We identified 84 LSP TSSs where the ESP-defined TSS was absent but was instead replaced by a differently positioned TSS at LSP (shown by red arrows in Fig. 4); these constituted types B and G in Fig. 4 (and J, K, L in Fig. S2). We also identified 37 genes which had LSP...
TSSs in addition to one or more ESP TSSs; these constituted types D and F and are also shown by red arrows in Fig. 4. Finally we found that in the case of 17 genes, TSS position was ppGpp-dependent (categorised as ‘R’ in Fig. S2 and Table S2).

Interestingly several genes in which one or more ESP TSSs were absent at LSP (and therefore presumably not used at LSP) had virulence related functions; these included the SPI1 regulators hilC and hilD and also the SPI1 encoded genes prgH, invH (categorised as type ‘A’ in Fig. 4 and Table S2). Other virulence related regulators in which one or more TSSs were absent at LSP compared to ESP included csrA, dksA, hfq, hns, mig-14, pagK, pipA (SPI5), slyA and virK. Similarly we could also identify 70 genes at LSP where the ESP TSS was absent but where an LSP TSS at a differing position was present (designated as type ‘B’ in Fig. 4 and Table S2). Other virulence related regulators in which one or more TSSs were absent at LSP compared to ESP included csaA, dksA, hfq, hns, mig-14, pagK, pipA (SPI5), sfA and virK. Similarly we could also identify 70 genes at LSP where the ESP TSS was absent but where an LSP TSS at a differing position was present (designated as type ‘B’ in Fig. 4 and Table S2). A survey of the functions of the latter genes revealed several with metabolic, virulence and regulatory functions. These included the anti-sigma factor encoding genes rseE and rseA, which are involved in sigma E activation in response to envelope and acid stress, and involved in virulence in S. Typhimurium [40,41]. Other virulence related genes of type B where the ESP TSS was absent and the LSP TSS was in a different position included lrp, pipD (SPI5), pagC, ygdP and vssH. Lrp has been shown to affect many stationary phase induced genes including those involved in the response to nutrient limitation, high concentrations of organic acids, and osmotic stress [42]. Lrp has also been shown to act as a virulence repres sor in S. Typhimurium [43]. YgdP is a dinucleoside polyphosphate hydrolase, mutant alleles of ygdP have been shown to reduce invasion of Salmonella by repressing SPI1 expression [44]. RfaH encodes a stationary phase-expressed transcriptional anti-terminator and affects operons encoding extracytoplasmic cell components involved in the virulence of E. coli pathogens and has also been shown to down-regulate key virulence factors in S. Typhimurium [45,46].

Non-coding RNAs at stationary phase

The new candidate LSP ncRNAs are listed in Table S3. Candidate ncRNAs were defined as either intergenic (>250 nt from a TSS) or antisense (strictly defined as being opposite the CDS of a gene). We found a total of 282 candidate ncRNAs of which 159 were asRNAs and 123 which were located in intergenic regions. Candidate asRNAs were found to 5 SPI2 genes, 4 of which encode secretion system effectors (ssaB, ssaC, ssaV and sseG), and orf2. Other virulence related genes with candidate asRNAs included virK, srb, pagC and two genes related to iron transport: woc and fhuE (Table S3). A MEME analysis of the 282 candidate ncRNAs revealed conservation of a consensus sequence closely matching the E. coli σ70 (~10) binding site (TATAAT) (Fig. S3). An analysis of the first nucleotide of the candidate LSP ncRNA transcripts also revealed a tendency for a purine residue (A – 41%, G – 29%), as has been previously noted for the +1 position of TSSs.
Finally, we determined the LSP read counts for the 114 sRNAs which have previously been validated by Northern blot analysis (Table S8, [20]). We were able to identify TSSs for all but 21 of the validated sRNAs (for read counts).

Several of the validated sRNAs with the highest read numbers at LSP (SdsR, CsrB, CsrC, RprA, GlmY and MgrR) have been shown to be highly expressed during stationary phase; CsrB and CsrC have both been shown to play a role in virulence in *Salmonella* [47–49].

Defining ppGpp-dependent gene expression at stationary phase using dRNA-seq

The dRNA-seq analysis revealed that the positions of 17 mRNA TSSs were ppGpp-dependent (categorised as type ‘R’ in Fig. S2 and listed in Table S2); an example of a ppGpp-dependent alteration in TSS position is shown for *gltA* which encodes citrate synthase (Fig. S4). Whether the ppGpp-dependent alterations in TSS position have any regulatory significance remains to be determined. We also observed large ppGpp-dependent changes in the read counts mapping to the TSSs identified at LSP (Table S1). We interpreted this as ppGpp-dependent transcriptional regulation; RNA-seq data has been used previously to identify differences in the transcriptional status of bacterial genes and validated by comparison with microarray experiments [50–52]. In order to estimate the expression level of a promoter from dRNA-seq data we calculated the number of non-enriched reads mapping between the primary TSS and 50 nt downstream of the TSS. ppGpp-activated expression was defined as 4-fold or higher transcript levels in the parent strain compared to the *ΔrelA ΔspoT* strain and ppGpp-repressed expression was defined as 4-fold or higher transcript levels in the *ΔrelA ΔspoT* strain compared to the parent strain (Table 1). For LSP mRNAs, we identified a total of 511 TSSs that were ppGpp-repressed by 4-fold or greater; these included 88 TSSs for which no mapped reads could be discerned in the SL1344 parent compared to the *ΔrelA ΔspoT* datasets. Only 96 (5.2%) TSSs were found to be ppGpp-activated by 4-fold or greater. In total 33% of the TSSs for protein coding genes were

![Figure 4. Use of alternate TSSs at LSP compared to ESP.](https://example.com/fig4.png)

Letters A to H annotate the different categories of arrangements of TSSs found at LSP compared to ESP that we were able to identify. Letters X, Y, Z and W represent TSS positions at ESP and LSP. The number of genes for which each type of rearrangement occurs is shown. For nucleotide positions of TSSs and genes to which the different types of rearrangements apply see Table S2. Red arrows indicate TSS positions that were identified at LSP but not ESP. Rearrangements which apply to one or two genes are shown in Fig. S2.

doi:10.1371/journal.pone.0092690.g004
ppGpp-dependent at LSP compared to 34% at ESP; together these are indicative of the major role ppGpp plays in modulating the stationary phase gene expression programmes (Table 1 and Table S1). The genes for the ppGpp-repressed TSSs were categorised according to function (Fig. 5), and the majority (184 genes) were of putative or unknown function. The largest functional classes of ppGpp-repressed genes were involved in purine/pyrimidine metabolism (36%), protein folding, degradation and stabilisation (34%) and nucleotide and nucleoside interconversions (29%) (Fig. 5). Other highly represented classes of ppGpp-repressed genes were involved in translation and protein modification (24%) and fatty acid and lipid metabolism (24%). Of the 92 genes that were ppGpp-activated greater than 4-fold, the majority were of putative or unknown function however the next largest category (13 genes; 22%) had virulence functions and included 5 SPI2 genes (ssaB, ssaG, sseA, sseI, sseJ, sseL) (Table 1 and Table S1).

For the stable RNA genes, the TSSs for all but 2 of the operons containing rRNA genes were ppGpp-repressed (greater than 4-fold), consistent with the ppGpp-repression of TSSs for genes involved in translation/protein modification described above, and with the general down-regulation of protein synthesis occurring at LSP (Table S4). However, in contrast to the general trend of ppGpp-repression of stable RNA genes, we found 7 tRNA genes, cognate to leucine, arginine, and tyrosine which were ppGpp-activated greater than 4-fold (Table S4).

It is known that ppGpp-repressed promoters tend to contain GC rich discriminator regions between the TSS and −10 regions that play a role in destabilising the Rnap-promoter complex resulting in transcriptional repression (e.g. tRNA promoters [53,54]). We therefore compared the GC content of the discriminator regions for all of the LSP TSSs to the level of repression by ppGpp. The median fold-repression of all of the ppGpp-repressed TSSs compared to GC content revealed a bias towards ppGpp-dependency with higher GC content (Fig. 6). A MEME analysis of all of the discriminator regions upstream of all of the ppGpp-repressed TSSs revealed no conserved region, however, 25 TSSs (out of 41) that were ppGpp repressed greater than 20-fold revealed a GC rich conserved region (Fig. 6, inset).

**Discussion**

The definition of the transcriptomic architecture for *S. Typhimurium* at later stationary phase resulted in the prediction of TSSs for 53% of ORFs which were found to closely correlate with the TSS positions for these ORFs previously defined at ESP [19]. The fewer mRNA TSSs that we were able to identify at LSP compared to ESP (53% and 78% respectively) are likely to be representative of the regulatory changes accompanying the decreased transcription occurring at LSP. However, we were able to define 245 LSP specific TSSs and found that 111 of these were within genes which had no identifiable TSS at ESP; the remaining

---

**Table 1.** The relative proportions of ppGpp-repressed and ppGpp-activated TSSs within the *S. Typhimurium* genome at LSP.

| TSSs type       | % ppGpp-repressed | % ppGpp activated | Total TSSs |
|-----------------|-------------------|-------------------|------------|
| ORF P           | 34                | 6                 | 1453       |
| ORF P, I        | 34                | 8                 | 64         |
| ORF S           | 34                | 1                 | 170        |
| ORF S, I        | 33                | 0                 | 12         |
| ORF I           | 16                | 8                 | 122        |
| sRNAs           | 28                | 6                 | 159        |
| Other ncRNAs    | 15                | 10                | 123        |
| Validated ncRNAs| 12                | 11                | 105        |
| tRNAs           | 22                | 10                | 40         |
| rRNAs           | 86                | 0                 | 14         |

Compiled from Tables S1, S3, S4 & S8.

doi:10.1371/journal.pone.0092690.t001
134 TSSs differed in position between LSP and ESP suggestive of the regulatory changes occurring between these growth phases. Functionally, the genes with TSSs that were either specific to LSP (i.e. not present at ESP), or in a different position compared to ESP TSSs were mostly involved in metabolic and regulatory changes which could be related to nutritional and other stresses encountered at stationary phase, and also within several virulence-related genes.

We were also able to map the transcriptomic architecture of SPI2 at LSP and identify a previously unreported TSS for **ssaR**. We were unable to identify a TSS upstream of **ssrB** or **orf242**, suggesting they were co-transcribed with **ssrA** (Fig. 3C). Feng and colleagues [28], found **ssaA** and **ssrB** expression to be uncoupled, with **ssrB** having its own promoter which was dependent on both OmpR and SsrB. However, it appears that this promoter is only active under specific growth conditions since other workers found no evidence of SsrB binding or functional promoter activity directly upstream of **ssrB** [30,63]. The finding that **orf242** is apparently co-transcribed with **ssrAB** is potentially interesting since a previous study found no evidence of it being required for virulence [64]. The mapping of TSSs for SsrB-dependent genes enabled us to locate the relative positions for predicted SsrB binding sites within their promoters, revealing that most sites were located close to or overlapping the -35 regions suggesting that SsrB activates transcription through direct interactions with RNA polymerase using a CRP-like Class II activation mechanism (reviewed in [65]). Alternative architecture at other promoters suggests SsrB can employ additional activation mechanisms, for example at the **slrP** promoter where the SsrB binding site overlaps the TSS (Table S6). Based on a MEME and FIMO analysis.

**Figure 5. Functional category analysis of ppGpp-dependent genes.** Functional categories were compiled from the Kyoto Encyclopedia of Genes and Genomes (KEGG; www.genome.jp/kegg) and The Comprehensive Microbial Resource (CMR) at the J. Craig Venter Institute (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) and a manual inspection based on the published literature. The total number of ORFs present in each category is indicated in parentheses.

doi:10.1371/journal.pone.0092690.g005
We report the discovery of 282 new candidate ncRNAs at LSP and 159 of these were asRNAs. The remaining candidate LSP ncRNAs (123) were located in the intergenic regions of genes and we note that 16 and 3 candidate ncRNAs were opposite the 5' and 3' regions of genes respectively and hence may play a cis-acting regulatory role in the expression of these genes. Indeed, several of the candidate ncRNAs were located opposite the 5' regions or CDS's of genes involved in virulence or combating stress; for example we found two ncRNAs (LSPncRNA41 and LSPncRNA39) which were opposite the 5' UTRs of abpC and ompC respectively, the former encodes a subunit of alkyl hydroperoxide reductase involved in combating oxidative stress and the latter encodes the bacterial siderophore enterochelin, involved in iron uptake within the host [66]. Oxidative stress has been noted to be present at stationary phase [67]. We were able to identify TSSs for the majority of sRNAs previously validated by Northern blots and found that some of the highest read numbers mapped to sRNAs that have been shown to play a role in adaptation to stationary phase or in S. Typhimurium virulence (Table S8). SelR sRNA was the most abundant later stationary phase sRNA and has been shown to down-regulate the major outer membrane porin, OmpD in S. Typhimurium [68]. We also note that CsrB and CsrC were amongst the most highly expressed sRNAs at LSP (Table S8). The CsrB and CsrC sRNAs antagonise CsrA which has many regulatory targets including biofilm formation, motility, virulence and metabolism, (for reviews see [69,70]). Both CsrB and CsrC were found to be strongly induced during growth in nutrient-poor medium, and it has been suggested that together with the BarA-SirA system, may form a mechanism by which bacteria sense the energy/growth status of the cell and switch from growth on glycolytic to gluconeogenic substrates which occurs at stationary phase [71,72]. Interestingly we found the TSS for CsrB but not CsrC to be ppGpp-repressed by a factor of 3.65-fold at LSP, (and 3.56-fold at ESP [19]), suggesting an indirect mechanism by which ppGpp could post-transcriptionally affect the function of many cellular processes in response to changes in the amino acid availability [71]. Several sRNAs have been shown to be encoded within Salmonella pathogenicity islands [73] and we were able to identify TSSs to four of these [IsrB-1, IsrJ, IsrK and IsrH-2; Table S8]. IsrJ has been implicated in the invasion of Salmonella into nonphagocytic cells and IsrH was found to be up-regulated during infection of J774 macrophages; IsrB was upregulated when cells were grown in minimal media and IsrH was up-regulated under a variety of stress conditions and during stationary phase [73]. Finally, we report two LSP ncRNAs that were located on the opposite strand to other ncRNAs: LSPncRNA125 and LSPncRNA312 were opposite the previously validated sRNAs Stnc710 and Stnc1060 respectively [19,20]. Several other examples of opposing ncRNAs have previously been noted potentially adding to the complexity of post-transcriptional regulatory mechanisms [19].

An analysis of the ppGpp-dependence of LSP mRNA TSSs showed very similar levels of ppGpp-repression compared to ESP TSSs (28% and 27% respectively, [19]). This is likely to be indicative of the major role ppGpp plays in modulating stationary phase gene expression, however not all of the ppGpp-repressed TSSs found at ESP were repressed at LSP and vice versa. These observations support evidence suggesting that although ppGpp plays a major role in transcriptional regulation, its action is essentially passive and modified by other regulatory elements such as sigma factors [74]. Fig. 2 reveals the reduced read numbers mapping to SPI1 TSSs at LSP compared to SPI1 at ESP [19]. At LSP there were no identifiable TSSs for hilD or rtsA and neither could we identify a TSS for hilD in the parental strain at LSP (Table S7). RtsA is a transcriptional activator of hilA and HldA is a major activator of SPI1 genes; HilD is the coordination point for several regulators that operate at the translational and transcriptional levels to control SPI1 and SPI2 expression [75]. Interestingly we observed a TSS with a perfectly conserved -10 region (TATAAT) in the parental strain at genomic coordinate 3039861 which was 552 nt upstream of the translational start site of hilD and opposite to the CDS of pgdH (Fig. S5). This TSS may

Figure 6. Correlation of ppGpp-repression of LSP TSSs of annotated genes with %GC content of discriminator regions. The analysis of GC content within the discriminator region was calculated using a custom Perl script. The inset shows the GC rich conserved region identified using MEME within the discriminator regions of promoters that were ppGpp-repressed by 20-fold or greater. doi:10.1371/journal.pone.0092690.g006
represent a potential ppGpp-dependent start site for \textit{hilD}, or spurious transcription, or promote an antisense transcript to \textit{pyrH} (TSS position 304120). We note that this TSS was also identified as ppGpp-repressed at ESP with very similar read numbers compared to LSP [19]. In addition we also identified a TSS mapping to genomic coordinate 3047036 within SPI1 at LSP which had higher read counts in the \textit{\Delta relA\Delta spoT} strain compared to the parent strain (Fig. S5). The latter TSS was internal to the \\textit{sitF} coding region and upstream of \\textit{sitT} and \\textit{sitF}. We also note that this TSS was also identified at ESP, but was ppGpp-activated by a factor of 2-fold [19]. In relation to virulence we also note ppGpp-dependent activation of the sRNA, \\textit{IsrH}-2 (16-fold); this RNA was found to be located within \\textit{IsrH}-1 and in accordance with our results, the former but not the latter sRNA was found to be expressed at stationary phase and to be highly induced within infected macrophages [73] (Table S8). Finally we note that the majority of stable RNA TSSs were ppGpp-repressed at LSP, apart from 7 tRNA TSSs which were ppGpp-dependently elevated greater than 4-fold. This was in contrast to ESP where the majority of tRNA TSSs (all but one) were ppGpp-dependently elevated greater than 4-fold. This is perhaps indicative of changes in the ppGpp-dependent differential processing or stability of tRNAs between ESP and LSP rather than any direct ppGpp-dependent regulation [19].

\textbf{Methods}

\textbf{Bacterial Strains and RNA extraction}

The virulent \textit{S. Typhimurium} parent strain, SL1344 and the isogenic \textit{\Delta relA\Delta spoT} strain were obtained as described in [19]. Bacterial cultures were grown overnight in Luria-Bertani broth (LB) at 37°C, 250 rpm, from -70°C glycerol stocks and used to inoculate into 50 ml of fresh LB in 250 ml conical flasks. The cultures were grown aerobically at 37°C with shaking at 250 rpm to an OD\textsubscript{600} of 0.6 (defined in this study as later stationary phase; LSP), conditions previously shown to induce SPI2 gene expression [23,24]. At this point, bacterial cultures were harvested and total RNA was extracted as described in [19]. The growth curves for SL1344 parent strain and the \textit{\Delta relA\Delta spoT} strain, and time of harvesting are shown in Fig. S1.

\textbf{Library preparation and sequencing}

In order to differentiate primary from processed transcripts, total RNA from biological replicates of each strain were divided into equal quantities and one half treated with Terminator 5'-phosphate-dependent exonuclease (TEX; Epicentre Biotechnologies) as previously described [19]. TEX specifically degrades RNAs with a 5’ phosphate end but does not degrade transcripts with a 5’ triphosphate end (primary transcripts) and thereby enriches for primary transcripts. The TEX treatment and cDNA library construction was performed by Vertis Biotechnology, AG, Germany (http://www.vertis-biotech.com) as described in [19]. For the Illumina-Solexa sequencing, four strand specific cDNA libraries were prepared: SL1344 parent non-enriched (SL1344-NE), SL1344 parent enriched (SL1344-EN), \textit{\Delta relA\Delta spoT} non-enriched (DM-NE) and \textit{\Delta relA\Delta spoT} enriched (DM-EN). Each library was tagged with a different barcode at the 5’ end to enable multiplexing during sequencing. Four of the Illumina-Solexa libraries were pooled and sequenced in a single lane using 50 single-read cycles on an Illumina Genome Analyzer II sequencing machine (GATC Biotech, Germany), yielding a total of \( \sim 2 \times 10^8 \) reads.

\textbf{Analysis of sequences and statistics}

The mapping statistics for the Illumina-Solexa cDNA libraries are shown in Table S3. Following sequencing, custom PERL scripts were used to separate the cDNA libraries based on their barcodes and to remove 5’ linker regions. Illumina-Solexa generated reads were mapped onto the \textit{S. Typhimurium} SL1344 genome (Genbank ID: FQ312003.1) including the three native virulence plasmids (SLP1, SLP2 and SLP3) using the \textit{segemehl} program [76]. Mapped reads were converted to a graph file and visualized on an Integrated Genome Browser (IGB) [77]. For comparison of expression levels between strains, mapped read numbers were normalised to their respective gDNA library sizes.

\textbf{Supporting Information}

\textbf{Figure S1} Growth curves for \textit{S. Typhimurium} SL1344 parental and \textit{\Delta relA\Delta spoT} strains. (DOCX)

\textbf{Figure S2} Single promoter rearrangements at LSP compared to ESP. Red arrows indicate LSP-specific TSSs. (DOCX)

\textbf{Figure S3} MEME analysis of 282 candidate LSP ncRNAs showing conserved -10 region (TATAAT). (DOCX)

\textbf{Figure S4} ppGpp-dependent location of \textit{gldA} transcriptional start sites. (DOCX)

\textbf{Figure S5} Promoter architecture and ppGpp-dependency of SPI1. (DOCX)

\textbf{Figure S6} Heat map showing clustered expression levels of genes encoded within SPI1 and SPI2 from early to late stationary phase in \textit{S. Typhimurium} SL1344 parent and \textit{\Delta relA\Delta spoT} strains from microarray transcriptomic data. (DOCX)

\textbf{Table S1} Master table of TSSs and ppGpp-dependent expression levels for \textit{S. Typhimurium} SL1344 ORFs at LSP. (XLSX)

\textbf{Table S2} TSSs with absent or altered positions at LSP compared to ESP. (XLSX)

\textbf{Table S3} TSSs and ppGpp-dependent expression of candidate LSP ncRNAs. (XLSX)

\textbf{Table S4} Expression levels and ppGpp-dependent expression of tRNAs and rRNAs at LSP and ESP. (XLSX)

\textbf{Table S5} Mapping statistics for SL1344 parental and \textit{\Delta relA\Delta spoT} libraries. (XLSX)

\textbf{Table S6} Location of SsrB binding sites relative to transcription start sites at SsrB dependent promoters. (DOCX)

\textbf{Table S7} ESP TSSs with no corresponding LSP TSS or alternative TSS. (XLSX)
Table S8  TSSs and ppGpp-dependent expression of known and predicted sRNAs that have been validated by Northern blot analysis. (XLSX)

Author Contributions

We conceived and designed the experiments: VKR AT. Performing the experiments: VKR. Analyzed the data: VKR AT. Contributed reagents/materials/analysis tools: VKR. Wrote the paper: AT VKR NS.
52. Oliver HF, Orsi RH, Ponnala L, Keich U, Wang W, et al. (2009) Deep RNA sequencing of L. monocytogenes reveals overlapping and extensive stationary phase and sigma E-dependent transcriptions, including multiple highly transcribed noncoding RNAs. BMC Genomics 10: 641.

53. Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, et al. (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. GCB 125: 1069–1082.

54. Zhou YN, Jin DJ (1998) The rpoB mutants destabilizing initiation complexes at stringently controlled promoters behave like “stringent” RNA polymerases in Escherichia coli. Proc. Natl Acad Sci U S A 95: 2908–2913.

55. Papenfort K, Pfeiffer V, Lucchini S, Sonawane A, Hinton JC, et al. (2008) SigmaE-dependent small non-coding RNAs. BMC Genomics 10: 641.

56. Guillier M, Gottesman S (2006) Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. Mol Microbiol 59: 231–247.

57. Guillier M, Gottesman S (2008) The 5’ end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. Nucleic Acids Res 36: 6781–6794.

58. Jiang H, Cao M, Cao X, Gu H, Zeng K (2010) sRNA (sraB) regulate the resistant ability of Salmonella enterica serovar enteritidis to egg albumen. Wei Sheng Wu Xue Bao 50: 1537–1544.

59. Sharma CM, Papenfort K, Permitzsch SR, Mollenkopf HJ, Hinton JC, et al. (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Mol Microbiol 81: 1144–1153.

60. Sharma CM, Darfeuille F, Planinga TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev 21: 2804–2817.

61. Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, et al. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the Salmonella core genome. Mol Microbiol 66: 1174–1191.

62. Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, et al. (2006) SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. Mol Microbiol 62: 1674–1688.

63. Bustamante VH, Martinez LC, Santana EJ, Kroullier LA, Seele-Morrin O, et al. (2008) HiD-mediated transcriptional cross-talk between SPI-1 and SPI-2. Proc Natl Acad Sci U S A 105: 14591–14596.

64. Hensel M, Nikolaus T, Egelseer C (1999) Molecular and functional analysis indicates a mosaic structure of Salmonella pathogenicity island 2. Mol Microbiol 34: 489–498.

65. Lee DJ, Minchuk SD, Busby SJ (2012) Activating transcription in bacteria. Annu Rev Microbiol 66: 125–152.

66. Naiz M, Theurl I, Schroll A, Theurl M, Fritsche G, et al. (2009) Absence of functional Hfe protects mice from invasive Salmonella enterica serovar Typhimurium infection via induction of lipocalin-2. Blood 114: 3642–3651.

67. Dukan S, Nyström T (1999) Oxidative stress defense and deterioration of growth-arrested Escherichia coli cells. J Biol Chem 274: 26027–26032.

68. Frohlich KS, Papenfort K, Berger AA, Vogel J (2012) A conserved RpoS-dependent small RNA controls the synthesis of major pools of OmpD. Nucleic Acids Res 40: 3623–3640.

69. Lucchetti-Miganet C, Burrowes E, Bayse C, Ermel G (2008) The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts. Microbiology 154: 16–29.

70. Timmermans J, Van Melder L (2010) Post-transcriptional global regulation by CsrA in bacteria. Cell Mol Life Sci 67: 2897–2908.

71. Jonas K, Melefejs O (2009) The Escherichia coli CsrB and CsrC small RNAs are strongly induced during growth in nutrient-poor medium. FEBS Lett 587: 80–86.

72. Wolfe AJ (2003) The acetate switch. Microbiol Mol Biol Rev 69: 12–50.

73. Padalou-Beauch G, Heuvelberg R, Elghabli-Weiss M, Baruch K, Roseumshine I, et al. (2006) Small RNAs encoded within genetic islands of Salmonella typhimurium shows host-induced expression and role in virulence. Nucleic Acids Res 36: 1913–1927.

74. Jishage M, Kvint K, Shingler V, Nystrom T (2002) Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev 16: 1260–1270.

75. Golubeva YA, Sadik AY, Ellermeier JR, Slauch JM (2012) Integrating global regulatory input into the Salmonella pathogenicity island type III secretion system. Genetics 190: 79–90.

76. Hoffmann S, Otto C, Kurtz S, Sharma CM, Kratovich P, et al. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol 5: e1000502.

77. Nicol JW, Helt GA, Blanchard SG Jr, Raja A, Loraine AE (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics 25: 2730–2731.