

In Vivo Expression of *Salmonella enterica* Serotype Typhi Genes in the Blood of Patients with Typhoid Fever in Bangladesh

Alaullah Sheikh^{1,2,9}, Richelle C. Charles^{2,3,*}, Nusrat Sharmeen¹, Sean M. Rollins^{2,3}, Jason B. Harris^{2,3}, Md. Saruar Bhuiyan^{1,2}, Mohammad Arifuzzaman¹, Farhana Khanam¹, Archana Bukka⁴, Anuj Kalsy², Steffen Porwollik⁵, Daniel T. Leung^{2,3}, W. Abdullah Brooks¹, Regina C. LaRocque^{2,3}, Elizabeth L. Hohmann², Alejandro Cravioto¹, Tanya Logvinenko⁶, Stephen B. Calderwood^{2,7}, Michael McClelland⁵, James E. Graham⁴, Firdausi Qadri^{1,†}, Edward T. Ryan^{2,3,8,†}

1 International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), Dhaka, Bangladesh, **2** Division of Infectious Diseases, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **3** Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States of America, **4** Department of Microbiology and Immunology, University of Louisville School of Medicine, Louisville, Kentucky, United States of America, **5** The Vaccine Research Institute of San Diego, San Diego, California, United States of America, **6** Division of Biostatistics, Institute for Clinical Research and Health Policy Studies (ICRHPS), Tufts Medical Center, Boston, Massachusetts, United States of America, **7** Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts, United States of America, **8** Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, United States of America

Abstract

Background: *Salmonella enterica* serotype Typhi is the cause of typhoid fever. It is a human-restricted pathogen, and few data exist on *S. Typhi* gene expression in humans.

Methodology/Principal Findings: We applied an RNA capture and amplification technique, Selective Capture of Transcribed Sequences (SCOTS), and microarray hybridization to identify *S. Typhi* transcripts expressed in the blood of five humans infected with *S. Typhi* in Bangladesh. In total, we detected the expression of mRNAs for 2,046 *S. Typhi* genes (44% of the *S. Typhi* genome) in human blood; expression of 912 genes was detected in all 5 patients, and expression of 1,100 genes was detected in 4 or more patients. Identified transcripts were associated with the virulence-associated PhoP regulon, *Salmonella* pathogenicity islands, the use of alternative carbon and energy sources, synthesis and transport of iron, thiamine, and biotin, and resistance to antimicrobial peptides and oxidative stress. The most highly represented group were genes currently annotated as encoding proteins designated as hypothetical, unknown, or unclassified. Of the 2,046 detected transcripts, 1,320 (29% of the *S. Typhi* genome) had significantly different levels of detection in human blood compared to *in vitro* cultures; detection of 141 transcripts was significantly different in all 5 patients, and detection of 331 transcripts varied in at least 4 patients. These mRNAs encode proteins of unknown function, those involved in energy metabolism, transport and binding, cell envelope, cellular processes, and pathogenesis. We confirmed increased expression of a subset of identified mRNAs by quantitative-PCR.

Conclusions/Significance: We report the first characterization of bacterial transcriptional profiles in the blood of patients with typhoid fever. *S. Typhi* is an important global pathogen whose restricted host range has greatly inhibited laboratory studies. Our results suggest that *S. Typhi* uses a largely uncharacterized genetic repertoire to survive within cells and utilize alternate energy sources during infection.

Citation: Sheikh A, Charles RC, Sharmeen N, Rollins SM, Harris JB, et al. (2011) *In Vivo* Expression of *Salmonella enterica* Serotype Typhi Genes in the Blood of Patients with Typhoid Fever in Bangladesh. PLoS Negl Trop Dis 5(12): e1419. doi:10.1371/journal.pntd.0001419

Editor: Christian Johnson, Fondation Raoul Follereau, France

Received: July 2, 2011; **Accepted:** October 21, 2011; **Published:** December 13, 2011

Copyright: © 2011 Sheikh et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the ICDDR,B and grants from the National Institutes of Health, including the National Institute of Allergy & Infectious Diseases (A1072599, A1077883 [ETR]; A1058935 [SBC]; A1052237, A1073971, A1075093, A1083964, A1039557, A1077645, A1083646, and BARD IS-4267-09 [MMCC]; A148635 [JEG]; ICDDR,B core funds (FQ, WAB); the PneumoADIP Project of the Johns Hopkins Bloomberg School of Public Health (WAB); a ICDDR,B-US Centers for Disease Control and Prevention Cooperative Agreement (WAB); a Training Grant in Vaccine Development from the Fogarty International Center (TW05572 [AS, MSB, FQ]); an American Recovery and Reinvestment Act (ARRA) FIC Post-doctoral Fellowship in Global Infectious Diseases (TW005572-S1 [RCC]); Career Development Awards from the Fogarty International Center (K01 TW007409 [JBH], K01 TW07144 [RCL]), and NIAID (K08 A1089721 [RCC]); and a Physician Scientist Early Career Award from the Howard Hughes Medical Institute (RCL). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rccharles@partners.org

⁹ These authors contributed equally to this work.

[†] These authors also contributed equally to this work.

Author Summary

Salmonella enterica serotype Typhi is the cause of typhoid fever and infects over 21 million cases and causes 200,000 deaths each year. *S. Typhi* only infects humans and this has greatly limited studies of *S. Typhi* pathogenesis. To study bacterial gene expression in human hosts, we used Selective Capture of Transcribed Sequences (SCOTS) and array hybridization to identify *S. Typhi* mRNAs expressed in the blood of 5 patients with *S. Typhi* infection. In total, we detected the expression of 2,046 *S. Typhi* genes (44% of the *S. Typhi* genome) in human blood; of these, 1,320 (29% of the *S. Typhi* genome) had significantly different levels of detection in human blood compared to *in vitro* cultures. Our results provide insight into *S. Typhi* pathogenesis, identifying both previously described and novel interactions occurring between host and microbe during the natural course of human infection. Further study of these genes, especially those of unknown function, may further our understanding of *S. Typhi* pathogenesis and aid in vaccine, diagnostic, and/or drug target development.

Introduction

Salmonella enterica serotype Typhi is a Gram-negative bacterium and the cause of typhoid fever. Typhoid fever affects over 21 million people each year, killing 200,000 [1]. *S. Typhi* is a human-restricted pathogen and this has greatly limited studies of *S. Typhi* pathogenesis. Our current understanding of *S. Typhi* responses during infection is largely based on the study of murine models with the related bacterium *S. Typhimurium* (i.e., a bacteria that causes a typhoid-like illness in mice) [2], a separate mouse model of *S. Typhi* infection [3], and *ex vivo* macrophage and epithelial cell models of *S. Typhi* and *S. Typhimurium* [4]. However, these studies have limitations, and do not fully replicate human disease. For instance, despite high sequence similarity, 13% of the genes in the *S. Typhi* genome are absent from *S. Typhimurium*, and the *S. Typhi* chromosome contains over 200 pseudogenes that *S. Typhimurium* does not [5,6].

Here we report the application of an mRNA/cDNA capture and amplification technology, Selective Capture of Transcribed Sequences (SCOTS), combined with cDNA hybridization technology [7–12], to directly assess the gene expression profile of *S. Typhi* in the blood of humans with typhoid fever in Bangladesh. We previously applied this technology to *S. Paratyphi A*, the 2nd leading cause of enteric fever, and detected expression of over 1700 bacterial genes during human infection [12]. Here we report the extension of this analysis to *S. Typhi*.

Methods

Ethics statement

This study was approved by the Ethical and Research Review Committees of the International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh (ICDDR,B) and the Human Research Committee of Massachusetts General Hospital; the study was conducted according to the principles expressed in the Declaration of Helsinki/Belmont Report. Written informed consent was obtained from all individuals or their guardians prior to study participation.

Study subject selection and sample collection

Individuals presenting to the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) Hospital or the Kamalapur field site of ICDDR,B were eligible for enrollment if

they met the following criteria at presentation: age of 1–59 years, fever duration of 3–7 days ($\geq 39^{\circ}\text{C}$), no obvious focus of infection, and no alternate diagnosis. We collected 2 ml of venous blood from participants, immediately placed these specimens in TRIzol (Invitrogen Life Technologies, Carlsbad, CA) at a 1 (blood):2 (TRIzol) volume ratio, and stored the samples at -70°C for later analysis. We simultaneously obtained 3–5 ml of blood for microbiologic analysis using a BacT/Alert automated system. We sub-cultured positive bottles on MacConkey agar, and identified *S. Typhi* isolates using standard biochemical tests and reaction with *Salmonella*-specific antisera [13]. After we collected blood, we treated patients with parenteral ceftriaxone, oral ciprofloxacin, or oral cefixime for up to 14 days at the discretion of the attending physician.

cDNA synthesis

To generate *S. Typhi* cDNA from blood samples, we used TRIzol-preserved blood of patients whose initial cultures were subsequently confirmed to grow *S. Typhi*. To create a corresponding *in vitro* *S. Typhi* cDNA sample for comparison, we grew each patient's bacterial isolate to mid-log growth phase (OD_{600} 0.45–0.6) in Luria Bertani (LB) broth, and preserved the samples in TRIzol at a 1 (mid-log culture):2 (TRIzol) volume ratio. We extracted total RNA from TRIzol preserved samples per the manufacturer's instructions (Invitrogen) and treated recovered RNA with DNase I on RNeasy columns (Qiagen Inc., Valencia, CA). We then converted 5 μg of total RNA into cDNA for each sample, as previously described with a few modifications [12]. Briefly, we used random priming (T-PCR) to obtain a representative amplifiable double-stranded cDNA population by using Superscript III (Invitrogen) with a conserved primer with a defined 5' end terminal sequence and a random nonamer at the 3' end [14]. We then synthesized second strands using the same primers and Klenow fragment (Invitrogen) according to the manufacturer's instructions, and then equilibrated samples based on 16S *S. Typhi* rRNA.

Selective Capture of Transcribed Sequences (SCOTS)

We separated bacterial cDNA from host DNA using SCOTS, as previously described [12]. Briefly, we mixed denatured biotinylated *S. Typhi* gDNA with blocking ribosomal *S. Typhi* DNA, and added this denatured mixture to both *in vivo* and *in vitro* cDNA samples. After hybridizing samples overnight at 67°C , we captured biotinylated *S. Typhi* gDNA-cDNA hybrids using streptavidin-coated magnetic beads (Dynabeads M-280 streptavidin, Invitrogen), eluted captured cDNA with NaOH, PCR-amplified cDNA samples with conserved primers, and purified products using Qiagen PCR column purification kits. We performed three rounds of capture and amplification to separate *S. Typhi* cDNA from host DNA and to generate the cDNA mixture used for microarray hybridization.

Salmonella microarray analysis

We labelled *in vivo* and *in vitro* cDNA recovered from SCOTS with Cy3 and Cy5, respectively, and hybridized these preparations to *Salmonella* ORF microarrays (version STv7S; McClelland Laboratory, Vaccine Research Institute of San Diego, CA, <http://www.sdibr.org/Faculty/mcclelland/mcclelland-lab>) in duplicate and with two dye reversals as previously described [12]. These microarrays contained gene-specific PCR-products of 4,600 ORFs from *Salmonella enterica* serotype Typhi CT18 (98.6% genome coverage) and 4,318 ORFs of strain Ty2 (98.0% genome coverage). The arrays also contained 1049 *S. enterica* ORFs absent from the *S. Typhi* genome. We used an equal amount of *in vivo* and

in vitro Cy dye-labeled product on all slides for a given patient. We used ScanArray software (ScanArray express, version 3.0.1) to quantify signal intensities.

For each individual patient, we considered a gene to be detected *in vivo* if at least 2 of the 3 replicate gene spots on each of the four slides for that infected human was at least ten median absolute deviations greater than the median of spots on the microarray corresponding to genes absent from the *S. Typhi* CT18 or Ty2 genomes. For those genes we detected *in vivo*, we evaluated whether there was a difference in expression when compared to detection levels for *in vitro* grown organisms. For this latter statistical analysis, we included genes with a coefficient of variation in signal intensity less than 50% within an array, and employed repeated measures ANOVA (to within slide replicate spots) with type (*in vivo* versus *in vitro*) and dye effects to LOESS-normalized, log-transformed data. Those genes with a False Discovery Rate of less than 0.05 computed using Benjamini-Hochberg multiple testing adjustment and a 2-fold variation in signal intensity were considered differentially expressed *in vivo* versus *in vitro*. We deposited data in the NCBI Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo), accessible through GEO accession number GSE30565. We based functional classification of genes on J. Craig Venter Institute annotations (<http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi>).

Quantitative PCR analysis

We used quantitative real time PCR (RT-qPCR) to confirm microarray results for a subset of genes. We compared mRNA levels in the peripheral blood (*in vivo* sample) of infected patients (i.e. the 5 patients included in our SCOTS array analysis and 5 additional patients) to three *in vitro* culture replicates of a *S. Typhi* isolate (from Patient 1) grown to mid-logarithmic phase in LB (*in vitro* sample), as previously described [12]. To maximize the likelihood of detecting differences in gene expression in comparative samples, we selected eight representative genes from operons involved in intra-cellular invasion or survival (STY4609, *sopE*, invasion-associated secreted protein; STY3639, *trxA*, thioredoxin); alternate energy usage (STY2244, *pduB*, putative propanediol utilization protein; STY0417, *psiF*, phosphate starvation-inducible

protein; STY2701, *eutN*, a putative ethanolamine utilization protein; STY0634, *jepC*, a ferric enterobactin transport ATP-binding protein); and bacterial adhesion (STY0207, *staA*, putative fimbrial protein and STY4543, *pilO*, putative pilus assembly protein), focusing on genes with high baseline signals and fold-increases by SCOTS-cDNA hybridization analysis comparing *in vivo* (high signal) to *in vitro* (low signal) samples. We also quantified by RT-qPCR the expression levels of two house-keeping genes that were predicted by SCOTS-cDNA hybridization to be equally expressed in *in vivo* and *in vitro* samples (STY0724, encoding a glutaminyl-tRNA synthetase, *glmS*; and STY3081, encoding an enolase, *eno*). We were unable to reproducibly assess expression levels of genes predicted by SCOTS-cDNA hybridization to be down-regulated in blood samples compared to *in vitro* grown organisms. To generate cDNA for quantitative RT-PCR from TRIzol-preserved samples, we used SuperScript II (Invitrogen) with random hexamers (Sigma, St. Louis, MO) according to the manufacturer's instructions, and performed RT-qPCR analysis using iQ SYBR Green Supermix reagent (Bio-Rad; Hercules, CA) and a CFX96 Real-time PCR detection system (Bio-Rad; Hercules, CA) as previously described [12]. Primers are listed in the Supplemental Table S1. We used no-template controls and samples lacking reverse transcriptase as baseline reactions for each sample. After calculating the threshold cycle (C_T) in the low/linear portion of product curves, we quantified gene copy numbers using pGEM-T Easy-based plasmids (Promega, Madison, WI) containing the gene of interest. To calculate the control gene copy number, we used plasmid size and A260 readings, and normalized gene copy numbers based on cDNA copies of 16S rRNA. We assessed singularity of product species and size by melting curve analysis, as previously described [15].

Results

Patient samples

Of the 89 patients enrolled for blood sample collection, we identified 10 patients with confirmed *S. Typhi* bacteremia at the time of TRIzol-preserved blood collection. We performed SCOTS-cDNA hybridization screening analysis using samples

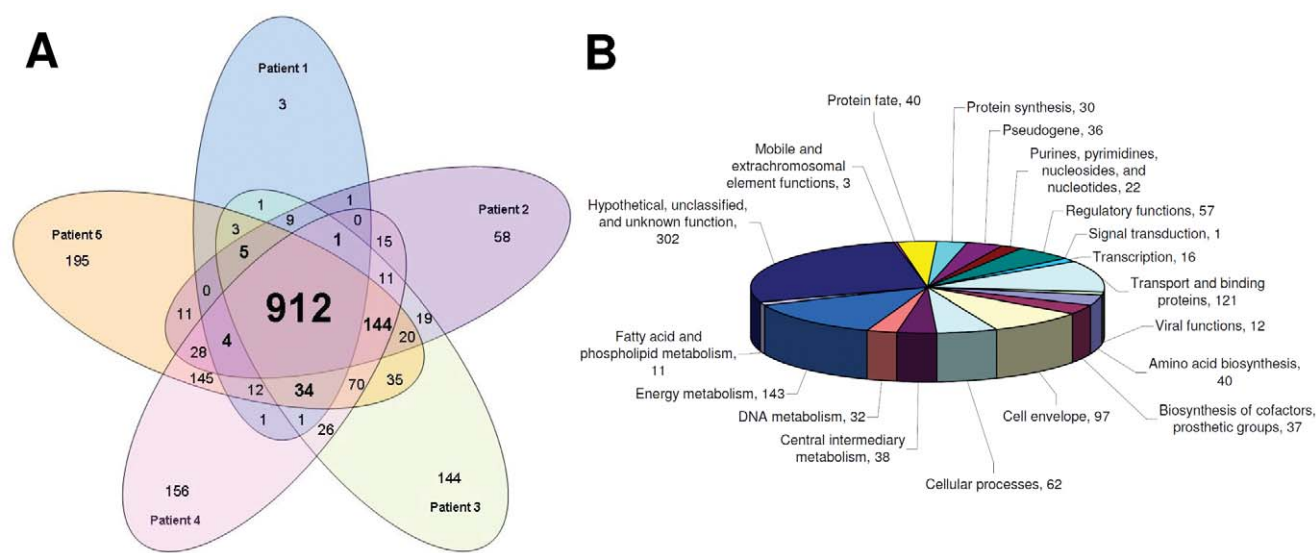


Figure 1. *S. Typhi* mRNA detected *in vivo* by SCOTS-microarray analysis. (A) Venn diagram of the 2046 *S. Typhi* transcripts detected *in vivo* by patient, and (B) functional classification of proteins encoded by the 1100 *S. Typhi* transcripts detected in 4 or more patients. doi:10.1371/journal.pntd.0001419.g001

from patients 1–5, and performed RT-qPCR on samples from patients 1–10, as sample quantity permitted.

S. Typhi transcripts detected in the blood of infected humans

Using SCOTS-cDNA hybridization technology, we detected expression of 2046 *S. Typhi* genes in the blood of bacteremic patients. This represents approximately 44% of the *S. Typhi* ORFeome (Figure 1A, Supplemental Table S2). Of these, we detected expression of 912 genes in all 5 patients (45% of detected transcripts), and 1100 in at least 4 of the 5 patients (54% of detected transcripts).

The products encoded by the 1100 genes identified in 4 or more patients fell into a number of functional categories (Figure 1B). The most highly represented group were genes currently annotated to

encode hypothetical proteins or proteins designated as unknown or unclassified. The next most highly represented groups were genes that encode products involved with energy metabolism, transport and binding, followed by genes encoding products of the cell envelope or associated with cellular processes and pathogenesis. Ninety-five of the 1100 genes were located within known *Salmonella* pathogenicity islands (SPI 1–7, 9, 10, 13, and 16), and 29 are known components of the PhoP regulon, a major virulence regulon in *Salmonella*, involved in intra-macrophage survival.

S. Typhi transcripts with a different level of detection in *in vivo* versus *in vitro* bacterial samples

A total of 31 genes were detected in 4 or more patients *in vivo*, but not detected in any *in vitro* sample (Table 1). The majority of

Table 1. *S. Typhi* transcripts detected only in *in vivo* samples.

CT18 Locus	Ty2 Locus	Gene	Function
Biosynthesis of cofactors, prosthetic groups, and carriers			
STY0828	t2092	<i>bioF</i>	8-amino-7-oxononanoate synthase
STY3725	t3471	<i>thiG</i>	thiamine biosynthesis protein
Cell envelope			
STY1609	t1379		hypothetical protein
STY4620	t4314	<i>nucD2</i>	putative lysozyme
Central intermediary metabolism			
STY0417	t2480	<i>psfF</i>	phosphate starvation-inducible protein PsfF
STY2702	t0393	<i>eutD</i>	putative phosphate acyltransferase
STY4773	t4468	<i>ppa</i>	inorganic pyrophosphatase
Energy metabolism			
STY1917	t1086	<i>hyaE</i>	hydrogenase-1 operon protein HyaE
STY2316	t0768	<i>manB</i>	Phosphomannomutase
Fatty acid and phospholipid metabolism			
STY2700	t0395	<i>eutE</i>	putative aldehyde dehydrogenase
Protein synthesis			
STY4360	t4067	<i>rplW</i>	50S ribosomal subunit protein L23
Regulatory functions			
STY3707	t3448	<i>yifE</i>	conserved hypothetical protein
Transport and binding proteins			
STY0065	t0058	<i>oadG</i>	oxaloacetate decarboxylase gamma chain
STY0123	t0110	<i>yabJ</i>	hypothetical ABC transporter
STY0636	t2276	<i>fepD</i>	ferric enterobactin transport protein FepD
STY2341	t0744	<i>mdtC</i>	putative RND-family transporter protein
Hypothetical, Unclassified and Unknown proteins			
	t3166		hypothetical protein
STY0321	t2569		Rhs-family protein
STY1058-60,1063-64,1069			putative prophage proteins
STY1323	t1640		conserved hypothetical protein
STY1548	t1434		conserved hypothetical protein
STY1732	t1256	<i>ydhZ</i>	conserved hypothetical protein
STY1916	t1087	<i>hyaD</i>	hydrogenase-1 operon protein HyaD
STY2244	t0835	<i>pduB</i>	putative propanediol utilization protein PduB
STY2608	t0487		conserved hypothetical protein
STY3448	t3185	<i>yraN</i>	conserved hypothetical protein

S. Typhi transcripts detected in the blood of 4 or more of 5 patients, but not in bacterial samples grown *in vitro*.

doi:10.1371/journal.pntd.0001419.t001

these genes are involved with survival in nutrient-limited conditions including *psiF*, a phosphate starvation-inducible protein; *bioF* and *thiG* involved in vitamin biosynthesis; *eutD*, *oadG*, and *pduB* involved in use of alternative carbon sources; and *fepD* involved in iron acquisition.

Of the 2046 transcripts detected in human blood samples, 1320 (representing 29% of *S. Typhi* ORFeome) had significantly different levels of detection in *in vivo* samples compared to bacterial samples grown *in vitro* (Figure 2A, Table S2). Detection levels for 141 transcripts were significantly different between *in vivo* and *in vitro* samples in all 5 patients, and 331 in at least 4 patients. These 331 encode products that fall into a number of functional categories (Figure 2B). The most highly represented group included proteins annotated as hypothetical, unknown, or unclassified. Other highly represented groups included energy metabolism, transport and binding, the cell envelope, and cellular processes and pathogenesis.

Quantitative Real Time-PCR analysis

To confirm *S. Typhi* mRNA expression levels in human blood compared to *in vitro* grown bacteria, we used RT-qPCR to assess the copy number of the following eight genes that had high *in vivo* baseline reactivity as well as fold-change between *in vivo* and *in vitro* samples by SCOTS array analysis: thioredoxin, *trxA* (STY3639); a putative fimbrial protein, *staA* (STY0207); an invasion-associated secreted protein, *sopE* (STY4609); a putative propanediol utilization protein, *pduB* (STY2244); a putative pilus assembly protein, *pilO* (STY4543); an phosphate-inducible starvation protein, *psiF* (STY0417); a putative ethanolamine utilization protein, *eutN* (STY2701); and a ferric enterobactin transport ATP-binding protein, *fepC* (STY0634). Compared to expression levels in *in vitro* grown bacteria, we found increased expression of all 8 genes in the blood of infected humans, including in humans not analyzed by the SCOTS-cDNA hybridization screening protocol (Figure 3, A–H). As predicted by our SCOTS screening, we found no differences by RT-qPCR in the expression of housekeeping genes *glnS* (STY0724) and *eno* (STY3081) in blood versus *in vitro* bacterial samples (Figure 3, I–J).

Discussion

S. Typhi is a human-restricted pathogen, the cause of typhoid fever, and a significant cause of global morbidity and mortality. Despite this, there are limited data on bacterial events within humans infected with *S. Typhi*. Here we describe the application of a cDNA capture-amplification approach combined with microarray hybridization technology to assess *S. Typhi* gene expression directly in the blood of infected humans. In total, we detected 2046 *S. Typhi* transcripts in human blood (45% of *S. Typhi* transcriptome); we detected 1100 in at least 4 of 5 patients. Two major virulence determinants of *Salmonella* are the ability to invade host cells and the ability to survive and replicate within host cells. The PhoPQ-two component regulatory system is involved in intra-macrophage survival and antimicrobial resistance [16], and *Salmonella* pathogenicity island-1 (SPI-1) and SPI-2 encode type three secretion systems (T3SSs) involved in invasion of host cells and intracellular survival and replication, respectively [17,18]. In our analysis, we identified 29 genes involved in the PhoP regulon as more highly expressed in human samples, including the two component regulator itself, *phoPQ*; *virK*, a virulence protein; *mgfBC*, involved in magnesium transport; *pmrF*, an antimicrobial resistance protein; and *slyB*, an outer membrane lipoprotein [19,20]. We also identified 95 genes located within previously described SPIs, including SPI-1 and 2, as well as genes within SPI-3–7, 9, 10, 13, and 16.

The role of SPI-1 in invasion of epithelial cells has been well established [21]. We detected a number of transcripts associated with SPI-1 genes, including a number that encode effector proteins injected into eukaryotic cells via the SPI-1 T3SS, such as SipB. We also detected a number of transcripts encoding SPI-1 T3SS effector proteins expressed from other SPIs, including *sopE* (expressed from SPI-7) and *sopB/sigD* (expressed from SPI-5); SopB/sigD is involved in creation and maintenance of the *Salmonella* Containing Vacuole (SCV), crucial to intra-cellular survival of *Salmonella* in eukaryotic cells [22]. Of note, we similarly identified SPI-1 transcripts in our recent analysis of *S. Paratyphi* A cDNA in the blood of infected humans in Bangladesh [12]. Our

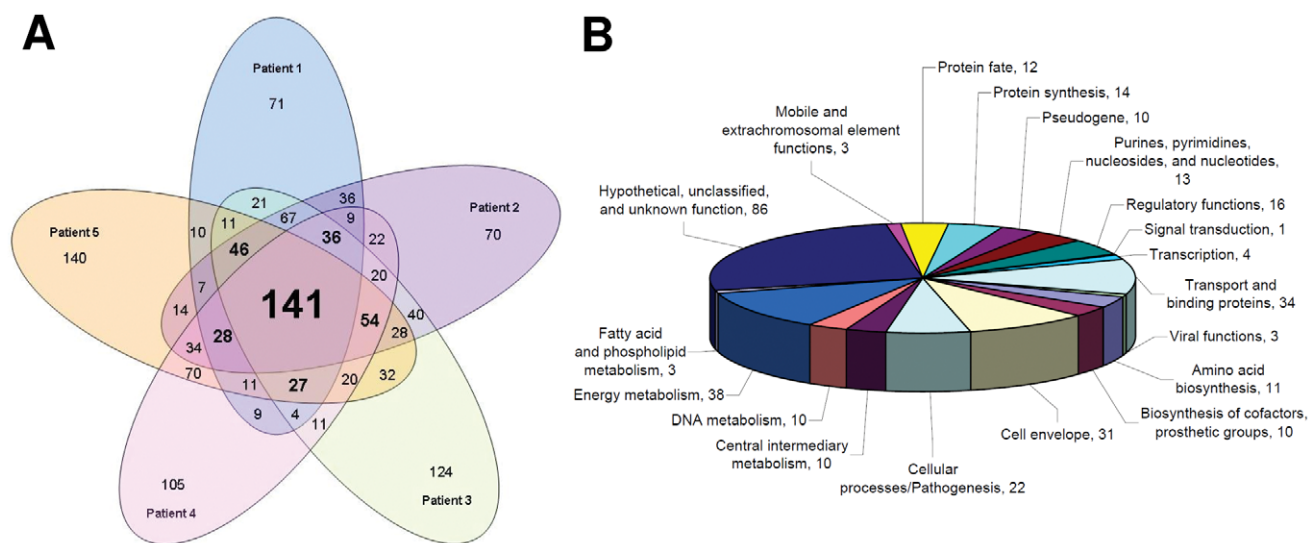


Figure 2. *S. Typhi* mRNA with different levels of detection *in vivo* versus *in vitro* bacterial samples by SCOTS-microarray analysis. (A) Venn diagram of the 1320 *S. Typhi* transcripts with significantly different levels of detection between *in vivo* and *in vitro* bacterial samples by patient, and (B) functional classification of proteins encoded by the 331 *S. Typhi* transcripts with significantly different levels of detection in 4 or more patients.

doi:10.1371/journal.pntd.0001419.g002

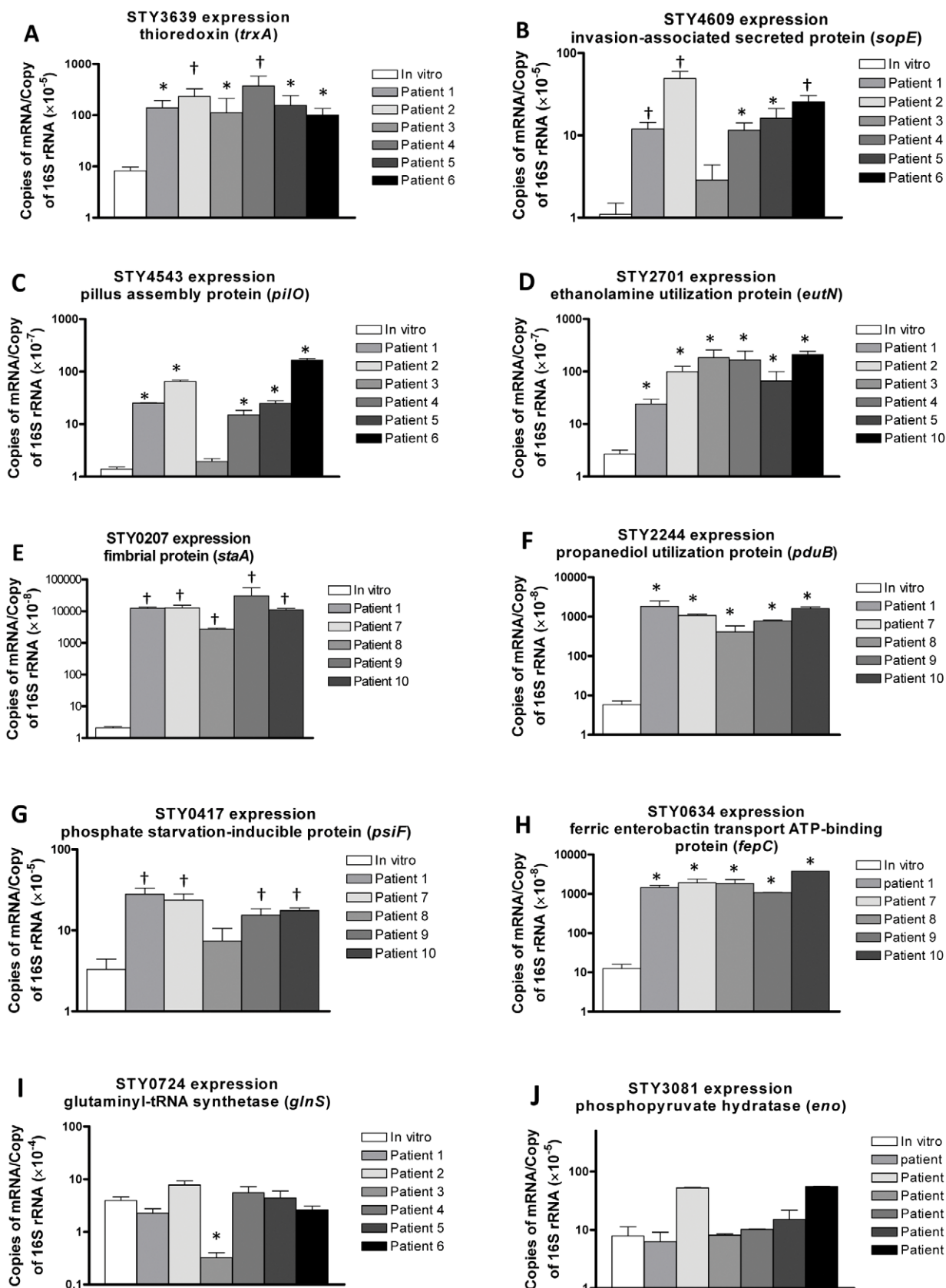


Figure 3. Quantitative RT-PCR *S. Typhi* mRNA expression profiles in human blood compared to *in vitro* bacterial samples. Quantitative real time-PCR analysis of *S. Typhi* genes (A–H) comparing RNA recovered from blood of bacteremic patients to an *in vitro* culture of the corresponding clinical isolate. Genes represented in A–H were identified by SCOTS-cDNA as being more highly expressed in human blood samples

than in *in vitro* grown organisms; genes I–J had equivalent levels of detection in *in vivo* versus *in vitro* samples by SCOTS. Mean copies of mRNA per copy of 16S rRNA and standard error of the mean are presented. * $p < 0.05$; † $p < 0.01$. RT-qPCR was performed on a minimum of 5 patients as sample quantity permitted.

doi:10.1371/journal.pntd.0001419.g003

detection of these transcripts in the blood of infected humans builds upon recent suggestions that the SPI-1 T3SS is involved in pathogenic events beyond intestinal epithelial cell invasion during enteric fever [23–25]. In addition to *sopE*, we also detected transcripts from the Type IV pilus operon encoded within SPI-7, including *pilL*, *pilO*, *pilQ*, *pilR*, *pilU*, and *pilV*, which facilitates invasion of *Salmonella* into epithelial cells and monocytes [26,27]. Identification of SPI-7 genes in our analysis is of particular interest since SPI-7 is absent from *S. Typhimurium* and *S. Paratyphi A*, but present in *S. Typhi*, *S. Paratyphi C*, and *S. Dublin* [28].

In addition to those associated with SPIs and the PhoPQ regulon, we detected transcripts from a number of virulence-associated *Salmonella* genes in human blood. These include aromatic amino acid biosynthesis pathway genes (*aroG*, *aroD*, *aroH*, *aroE*, *aroB*); mutations in this pathway have been the basis of live attenuated *S. Typhi* vaccines [29]. We also detected transcripts from genes involved in purine biosynthesis (*guaB*, *purG*, *purA*) [30] and divalent cation transport including Mg^{2+} (*corA*, *mgtBC*) [31–33], and Fe^{2+} and Mn^{2+} uptake systems (*sitBC* and *mntH*) [34] that have all been associated with virulence in *Salmonella*.

In order to adapt to the intracellular environment, *Salmonella* must alter its metabolism to available nutrient and energy sources. We detected transcripts of genes involved in the use of alternative carbon sources, the coenzyme B12-dependent 1,2-propranediol utilization pathway (encoded by the *pdu* operon), and the ethanolamine utilization pathway (encoded by the *eut* operon). We also found these operons to be up-regulated in our analysis of *S. Paratyphi A* genes detected in the blood of humans [12], and mutations in these operons result in attenuation of virulence in *S. Typhimurium* infection models [35–37]. We also identified transcripts expressed from genes encoding three NiFe-uptake hydrogenases that have been associated with virulence in *S. Typhimurium*, including hydrogenase A, B and D [38]. Prior studies have shown that the *hya* and *hyd* operons are upregulated in murine and human phagocytes; *hya* genes are required for survival within macrophages, and both *hya* and *hyd* genes were detected in mice using the RIVET (Resolvase In-Vivo Expression Technology) reporter system that identifies genes expressed *in vivo* [39]. Our analysis shows that these genes are also expressed by *S. Typhi* during human infection. Other potential virulence-associated genes that we identified included genes involved in thiamine biosynthesis (e.g. *thiG*, *thiJ*, *abpA*), biotin biosynthesis (e.g. *bioB*, *bioF*, *kbl*), iron acquisition via siderophore biosynthesis (e.g. *iroA* gene cluster, *fes*, *fepECDB*), and phosphate transport (*ugpBAEC* operon), many of which were also detected in our transcriptional analysis of *S. Paratyphi A* in infected humans [12].

In addition to survival in nutrient-limited conditions, *Salmonella* must also be able to survive the action of antimicrobial peptides, oxidative killing, and nitric oxide in various ecologic niches within the human body. We detected genes that may be involved in survival of stressful environments, including a number involved in antimicrobial resistance (e.g. *pqaB*, *virK*, *pmrF*, *smvA*, *bacA*, *emrA*, *mdtC*) [40–45], oxidative stress (e.g. *trxA*) [46], resistance to acid tolerance (e.g. *narX*YWV operon) [47], and genes involved in DNA recombination and repair (e.g. *recA*, *recBD*, *recN*, *recG*, *xthA*) [48]. Of note, the most highly represented group were genes currently annotated to encode hypothetical proteins or proteins designated as unknown or unclassified.

When comparing expression levels of *S. Typhi* genes detected in our analysis in humans to expression levels of *S. Typhi* genes in *in vitro* grown cultures, equilibrating for *S. Typhi* 16S rRNA, we noted differing levels of *S. Typhi* mRNA for 65% of the genes detected in humans. In total, 331 *S. Typhi* transcripts had significantly different levels of detection in at least 4 patients compared to *in vitro* cultures, and 141 had significant differences in all 5 patients compared to mRNA detected in *in vitro* cultures. Identified genes were involved in iron (*fepB*, *fepC*, *fepD*), thiamine (*thiG*), and biotin (*bioF*) metabolism; use of alternative carbon sources including ethanolamine (*eutB*, *eutC*, *eutD*, *eutA*, and *eutN*), oxalacetate (*oadAB* and *oadG*), and propranediol (*pduB* and *pduK*); and antimicrobial resistance (*bacA*, *mdtC*). We also identified these operons in our analysis of *S. Paratyphi A*, further supporting a potential role of these operons in the pathogenesis of enteric fever [12]. In addition, we identified 24 genes with significantly different levels of expression in *in vivo* compared to *in vitro* samples that are not present in the *S. Typhimurium* genome and may play an important role in *S. Typhi* pathogenesis, including genes encoded within the Type IV pilus cluster of SPI-7 (i.e. *pilO* and *pilL*), and fimbrial proteins *staA* and *steD*. Of note, the largest grouping of *S. Typhi* genes identified in our comparison encoded proteins of unknown or unclassified function.

Our findings are similar to prior *Salmonella* transcriptional analyses. We previously applied SCOTS-microarray analysis to *S. Paratyphi A* in the blood of infected humans, and the homologs of 75% of the bacterial transcripts identified in *S. Paratyphi A* infected patients were also identified in *S. Typhi* infected patients [12]. SCOTS analysis has also been previously applied to *S. Typhi* using an *ex vivo* macrophage model system by Faucher et al. [9]. Similar to our current analysis using blood of infected patients, the *ex vivo* analysis also detected transcripts of genes involved in intracellular survival including a number of genes encoded within SPI-2, *mgtBC* in SPI-3, the SPI-1 effector, *sopE*, and genes involved in antimicrobial peptide resistance. Both analyses suggested a role of SPI-1 beyond invasion of the intestinal epithelium and the potential role of alternative carbon sources in *S. Typhi* pathogenesis. In contrast to Faucher's analysis, we found higher levels of transcripts of genes involved in iron acquisition and transport *in vivo* including *fes*, *flu*, *fep*, *iro*, and *ent*. Our detection of these genes may reflect a greater complexity or degree of iron-limitation in the blood of infected humans versus in a cultured macrophage model system.

To our knowledge there has not been a prior analysis of *S. Typhi* gene expression across the transcriptome in humans. Our results highlight potential survival adaptations of *S. Typhi* within the human host, including expression of genes required for utilization of alternative carbon and energy sources, divalent cation transport, antimicrobial resistance, and oxidative stress resistance, as well as many genes whose function is currently unknown. Further study of these genes, especially those of unknown function, may further our understanding of *S. Typhi* pathogenesis and aid in vaccine, diagnostic, and/or drug target development.

Supporting Information

Table S1 qPCR Primer sequences.
(DOC)

Table S2 *S. Typhi* genes whose transcripts were detected in the blood of humans with typhoid fever. (XLS)

Author Contributions

Conceived and designed the experiments: AS RCC NS SMR JBH MSB MA FK AB AK SP DTL WAB RCL ELH AC TL SBC MM JEG FQ

References

- Crump JA, Luby SP, Mintz ED (2004) The global burden of typhoid fever. Bull World Health Organ 82: 346–353.
- Watson KG, Holden DW (2010) Dynamics of growth and dissemination of *Salmonella in vivo*. Cell Microbiol 12: 1389–1397.
- Song J, Willinger T, Rongvaux A, Eynon EE, Stevens S, et al. (2010) A mouse model for the human pathogen *Salmonella typhi*. Cell Host Microbe 8: 369–376.
- Haraga A, Ohlson MB, Miller SI (2008) *Salmonellae* interplay with host cells. Nat Rev Microbiol 6: 53–66.
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, et al. (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature 413: 852–856.
- Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, et al. (2001) Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature 413: 848–852.
- Daigle F, Graham JE, Curtiss R, 3rd (2001) Identification of *Salmonella typhi* genes expressed within macrophages by selective capture of transcribed sequences (SCOTS). Mol Microbiol 41: 1211–1222.
- Faucher SP, Curtiss R, 3rd, Daigle F (2005) Selective capture of *Salmonella enterica* serovar Typhi genes expressed in macrophages that are absent from the *Salmonella enterica* serovar Typhimurium genome. Infect Immun 73: 5217–5221.
- Faucher SP, Porwollik S, Dozois CM, McClelland M, Daigle F (2006) Transcriptome of *Salmonella enterica* serovar Typhi within macrophages revealed through the selective capture of transcribed sequences. Proc Natl Acad Sci U S A 103: 1906–1911.
- Graham JE, Clark-Curtiss JE (1999) Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc Natl Acad Sci U S A 96: 11554–11559.
- Daigle F, Hou JY, Clark-Curtiss JE (2002) Microbial gene expression elucidated by selective capture of transcribed sequences (SCOTS). Methods Enzymol 358: 108–122.
- Sheikh A, Charles RC, Rollins SM, Harris JB, Bhuiyan MS, et al. (2010) Analysis of *Salmonella enterica* serotype Paratyphi A gene expression in the blood of bacteremic patients in Bangladesh. PLoS Negl Trop Dis 4: e908.
- Giannella RA (1996) *Salmonella*. In: Baron S, ed. Medical Microbiology. Galveston/Texas: University of Texas Medical Branch at Galveston.
- Froussard P (1992) A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res 20: 2900.
- Rollins SM, Peppercorn A, Young JS, Drysdale M, Baresch A, et al. (2008) Application of in vivo induced antigen technology (IVIAT) to *Bacillus anthracis*. PLoS One 3: e1824.
- Charles RC, Harris JB, Chase MR, Lebrun LM, Sheikh A, et al. (2009) Comparative proteomic analysis of the PhoP regulon in *Salmonella enterica* serovar Typhi versus Typhimurium. PLoS One 4: e6994.
- Galan JE (1999) Interaction of *Salmonella* with host cells through the centisome 63 type III secretion system. Curr Opin Microbiol 2: 46–50.
- Groisman EA (2001) The pleiotropic two-component regulatory system PhoP-PhoQ. J Bacteriol 183: 1835–1842.
- Kato A, Groisman EA, Howard Hughes Medical Institute (2008) The PhoQ/PhoP regulatory network of *Salmonella enterica*. Adv Exp Med Biol 631: 7–21.
- Monsieurs P, De Keersmaecker S, Navarre WW, Bader MW, De Smet F, et al. (2005) Comparison of the PhoPQ regulon in *Escherichia coli* and *Salmonella typhimurium*. J Mol Evol 60: 462–474.
- Galan JE (1996) Molecular genetic bases of *Salmonella* entry into host cells. Mol Microbiol 20: 263–271.
- Hernandez LD, Hueffer K, Wenk MR, Galan JE (2004) *Salmonella* modulates vesicular traffic by altering phosphoinositide metabolism. Science 304: 1805–1807.
- Ibarra JA, Steele-Mortimer O (2009) *Salmonella*—the ultimate insider. *Salmonella* virulence factors that modulate intracellular survival. Cell Microbiol 11: 1579–1586.
- Giacomodonato MN, Uzzau S, Bacciu D, Caccuri R, Sarnacki SH, et al. (2007) SipA, SopA, SopB, SopD and SopE2 effector proteins of *Salmonella enterica* serovar Typhimurium are synthesized at late stages of infection in mice. Microbiology 153: 1221–1228.
- Knodler LA, Vallance BA, Celli J, Winfree S, Hansen B, et al. (2010) Dissemination of invasive *Salmonella* via bacterial-induced extrusion of mucosal epithelia. Proc Natl Acad Sci U S A 107: 17733–17738.
- Pan Q, Zhang XL, Wu HY, He PW, Wang F, et al. (2005) Aptamers that preferentially bind type IVB pili and inhibit human monocytic-cell invasion by *Salmonella enterica* serovar Typhi. Antimicrob Agents Chemother 49: 4052–4060.
- ETR. Performed the experiments: AS RCC NS FK AB AK SMR MA MSB SP DTL. Analyzed the data: TL AS RCC MM SP ETR. Contributed reagents/materials/analysis tools: ELH JBH RCL SBC ETR JEG MM SP FQ AC WAB TL. Wrote the paper: AS RCC NS SMR JBH MSB MA FK AB AK SP DTL WAB RCL ELH AC TL SBC MM JEG FQ ETR.
- Zhang XL, Tsui IS, Yip CM, Fung AW, Wong DK, et al. (2000) *Salmonella enterica* serovar Typhi uses type IVB pili to enter human intestinal epithelial cells. Infect Immun 68: 3067–3073.
- Pickard D, Wain J, Baker S, Line A, Chohan S, et al. (2003) Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7. J Bacteriol 185: 5055–5065.
- Stocker BA (1988) Auxotrophic *Salmonella typhi* as live vaccine. Vaccine 6: 141–145.
- McFarland WC, Stocker BA (1987) Effect of different purine auxotrophic mutations on mouse-virulence of a Vi-positive strain of *Salmonella dublin* and of two strains of *Salmonella Typhimurium*. Microb Pathog 3: 129–141.
- Blanc-Potard AB, Groisman EA (1997) The *Salmonella selC* locus contains a pathogenicity island mediating intramacrophage survival. EMBO J 16: 5376–5385.
- Papp-Wallace KM, Nartea M, Kehres DG, Porwollik S, McClelland M, et al. (2008) The CorA Mg²⁺ channel is required for the virulence of *Salmonella enterica* serovar Typhimurium. J Bacteriol 190: 6517–6523.
- Smith RL, Kaczmarek MT, Kucharski LM, Maguire ME (1998) Magnesium transport in *Salmonella typhimurium*: regulation of *mgtA* and *mgtCB* during invasion of epithelial and macrophage cells. Microbiology 144(Pt 7): 1835–1843.
- Zaharik ML, Cullen VL, Fung AM, Libby SJ, Kujat Choy SL, et al. (2004) The *Salmonella enterica* serovar Typhimurium divalent cation transport systems MntH and SitABCD are essential for virulence in an *Nramp1*^{G169} murine typhoid model. Infect Immun 72: 5522–5525.
- Klumpp J, Fuchs TM (2007) Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. Microbiology 153: 1207–1220.
- Stojiljkovic I, Baumber AJ, Heffron F (1995) Ethanolamine utilization in *Salmonella typhimurium*: nucleotide sequence, protein expression, and mutational analysis of the *chcA chcB eutE eutJ eutG eutH* gene cluster. J Bacteriol 177: 1357–1366.
- Conner CP, Heithoff DM, Julio SM, Sinsheimer RL, Mahan MJ (1998) Differential patterns of acquired virulence genes distinguish *Salmonella* strains. Proc Natl Acad Sci U S A 95: 4641–4645.
- Maier RJ, Olczak A, Maier S, Soni S, Gunn J (2004) Respiratory hydrogen use by *Salmonella enterica* serovar Typhimurium is essential for virulence. Infect Immun 72: 6294–6299.
- Zbell AL, Maier SE, Maier RJ (2008) *Salmonella enterica* serovar Typhimurium NiFe uptake-type hydrogenases are differentially expressed *in vivo*. Infect Immun 76: 4445–4454.
- Cain BD, Norton PJ, Eubanks W, Nick HS, Allen CM (1993) Amplification of the *bacA* gene confers bacitracin resistance to *Escherichia coli*. J Bacteriol 175: 3784–3789.
- Nishino K, Latifi T, Groisman EA (2006) Virulence and drug resistance roles of multidrug efflux systems of *Salmonella enterica* serovar Typhimurium. Mol Microbiol 59: 126–141.
- Parra-Lopez C, Baer MT, Groisman EA (1993) Molecular genetic analysis of a locus required for resistance to antimicrobial peptides in *Salmonella Typhimurium*. EMBO J 12: 4053–4062.
- Gunn JS, Ryan SS, Van Velkinburgh JC, Ernst RK, Miller SI (2000) Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar Typhimurium. Infect Immun 68: 6139–6146.
- Baker SJ, Gunn JS, Morona R (1999) The *Salmonella typhi* melittin resistance gene *pqaB* affects intracellular growth in PMA-differentiated U937 cells, polymyxin B resistance and lipopolysaccharide. Microbiology 145(Pt 2): 367–378.
- Detweiler CS, Monack DM, Brodsky IE, Mathew H, Falkow S (2003) *virK*, *somA* and *resC* are important for systemic *Salmonella enterica* serovar Typhimurium infection and cationic peptide resistance. Mol Microbiol 48: 385–400.
- Bjor E, Eriksson-Ygberg S, Aslund F, Rhen M (2006) Thioredoxin 1 promotes intracellular replication and virulence of *Salmonella enterica* serovar Typhimurium. Infect Immun 74: 5140–5151.
- Spector MP, Garcia del Portillo F, Bearson SM, Mahmud A, Magut M, et al. (1999) The *rpoS*-dependent starvation-stress response locus *stiA* encodes a nitrate reductase (*narZYW*) required for carbon-starvation-inducible thermotolerance and acid tolerance in *Salmonella typhimurium*. Microbiology 145(Pt 11): 3035–3045.
- Buchmeier NA, Lipps CJ, So MY, Heffron F (1993) Recombination-deficient mutants of *Salmonella typhimurium* are avirulent and sensitive to the oxidative burst of macrophages. Mol Microbiol 7: 933–936.