Analysis of Salmonella enterica Serotype Paratyphi A Gene Expression in the Blood of Bacteremic Patients in Bangladesh

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version
doi:10.1371/journal.pntd.0000908

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:4889510

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Analysis of *Salmonella enterica* Serotype Paratyphi A Gene Expression in the Blood of Bacteremic Patients in Bangladesh

Alaullah Sheikh\(^1,2,4,\) Richelle C. Charles\(^2,3,4,\) Sean M. Rollins\(^2,\) Jason B. Harris\(^2,3,\) Md. Saruar Bhuiyan\(^1,2,\) Farhana Khanam\(^1,\) Archana Bukka\(^4,\) Anuj Kalsy\(^2,\) Steffen Porwollik\(^5,\) W. Abdullah Brooks\(^1,\) Regina C. LaRoque\(^2,3,\) Elizabeth L. Hohmann\(^2,\) Alejandro Cravioto\(^1,\) Tanya Logvinenko\(^6,\) Stephen B. Calderwood\(^2,7,\) Michael McClelland\(^2,\) James E. Graham\(^4,\) Firdausi Qadri\(^1,\) Edward T. Ryan\(^2,3,8^a\)

\(^1\) 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; \(^5\) 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 Paratyphi A is a human-restricted cause of paratyphoid fever, accounting for up to a fifth of all cases of enteric fever in Asia.

**Methodology/Principal Findings:** In this work, we applied an RNA analysis method, Selective Capture of Transcribed Sequences (SCOTS), and cDNA hybridization-microarray technology to identify 800 Paratyphi A transcripts expressed by bacteria in the blood of three patients in Bangladesh. In total, we detected 1,798 *S. Paratyphi A* mRNAs expressed in the blood of infected humans (43.9% of the ORFome). Of these, we identified 868 in at least two patients, and 315 in all three patients. *S. Paratyphi A* Transcripts identified in at least two patients encode proteins involved in energy metabolism, nutrient and iron acquisition, vitamin biosynthesis, stress responses, oxidative stress resistance, and pathogenesis. A number of detected transcripts are expressed from PhoP and SlyA-regulated genes associated with intra-macrophage survival, genes contained within *Salmonella* Pathogenicity Islands (SPIs) 1–4, 6, 10, 13, and 16, as well as RpoS-regulated genes. The largest category of identified transcripts is that of encoding proteins with unknown function. When comparing levels of bacterial mRNA using *in vivo* samples collected from infected patients to samples from *in vitro* grown organisms, we found significant differences for 347, 391, and 456 *S. Paratyphi A* Transcripts in each of three individual patients (approximately 9.7% of the ORFome). Of these, expression of 194 transcripts (4.7% of ORFs) was concordant in two or more patients, and 41 in all patients. Genes encoding these transcripts are contained within SPI-1, 3, 6 and 10, PhoP-regulated genes, involved in energy metabolism, nutrient acquisition, drug resistance, or uncharacterized genes. Using quantitative RT-PCR, we confirmed increased gene expression *in vivo* for a subset of these genes.

**Conclusion/Significance:** To our knowledge, we describe the first microarray-based transcriptional analysis of a pathogen in the blood of naturally infected humans.


Editor: Elizabeth Angelica Leme Martins, Instituto Butantan, Brazil

Received July 21, 2010; Accepted November 8, 2010; Published December 7, 2010

Copyright: © 2010 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 (www.icddrb.org) and grants from the National Institutes of Health, including the National Institute of Allergy & Infectious Diseases (www.niaid.nih.gov – AI072599, AI077863 [JTR]; A0558535 [SBC]; AI052237, AI073971, AI075093, AI083964, AI035557, AI077645, AI083646, and BARD 15-4267-09 [MCC]; A148635 [JEG]); the PneumoADIP Project of the Johns Hopkins Bloomberg School of Public Health (www.preventpneumo.org - WAB); a ICDDR,B-US Centers for Disease Control and Prevention Cooperative Agreement (www.cdc.gov - WAB); ICDDR,B core funds (FQ, WAB); The Thrasher Research Fund (www.thrasherresearch.org - WAB); a Training Grant in Vaccine Development from the Fogarty International Center (www.fic.nih.gov - TW05572 [AS, MSB, FQ]; an American Recovery and Reinvestment Act (ARRA) FIC Post-doctoral Fellowship in Global Infectious Diseases (TW05572-S1 [RCC]); Career Development Awards from the Fogarty International Center (K01 TW007409 [JBI], K01 TW07144 [RCL]); and NIAID (K08 AI089721 [RCL]); and a Physician Scientist Early Career Award from the Howard Hughes Medical Institute (www.hhmi.org - 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: alaullah@icddrb.org

\(^*\) These authors contributed equally to this work.

\(^\dagger\) These authors also contributed equally to this work.
Author Summary

Salmonella enterica serotype Paratyphi A (S. Paratyphi A) is a significant emerging food and water-borne pathogen that currently accounts for one fifth of all cases of enteric fever in many areas of Asia. S. Paratyphi A only infects humans, and the lack of an appropriate animal model has limited the study of S. Paratyphi A infection. In this study, we report the application of an RNA analysis method, Selective Capture of Transcribed Sequences (SCOTS), to evaluate which S. Paratyphi A genes are expressed directly in the blood of infected humans. Our results provide insight into the bacterial adaptations and modifications that S. Paratyphi A may need to survive within infected humans and suggest that similar approaches may be applied to other pathogens in infected humans and animals.

Introduction

Salmonella enterica serotype Paratyphi A (S. Paratyphi A) is an emerging food and water-borne pathogen that currently accounts for approximately 1 in 5 cases of enteric fever in South Asia [1,2]. Although previously thought to cause a milder illness than Salmonella enterica serotype Typhi (S. Typhi), recent studies suggest that S. Paratyphi A causes a clinical syndrome quite similar to that caused by S. Typhi [1,3]. Over the past decade, S. Paratyphi A isolation rates have increased throughout South Asia, along with antimicrobial resistance [1,3-6]. The reasons for this emergence of S. Paratyphi A are unclear, and may relate in part to secondary effects of vaccine programs targeting S. Typhi [6]. No commercially available vaccine protects against S. Paratyphi A infection, although some protection against S. Paratyphi B is provided by the oral live attenuated typhoid vaccine strain Ty2a (a derivative of S. Typhi wild type strain Ty2) [1,7,8]. Increasing rates of infection, the lack of a commercially available vaccine effective against S. Paratyphi A infection, and steadily increasing resistance of S. Paratyphi A to antimicrobial agents make S. Paratyphi A infection a growing public health concern.

Following oral ingestion, S. Paratyphi A organisms invade intestinal epithelial cells, are taken up by gut-associated lymphoreticular tissues, and enter the systemic circulation. In the bloodstream, a majority of organisms reside within professional phagocytic cells, while the remainder are extracellular [9]. S. Paratyphi A studies have been limited by both a lack of an adequate animal model, and the low number of microorganisms present in the blood of infected individuals (estimated at 0.1–10,000 Salmonella organisms per ml of blood) [9-11]. Most of what is known about the pathogenesis of S. Paratyphi A has, therefore, been extrapolated from studies with S. Typhi and S. Typhimurium, and comparative genomic analyses [12-21].

To begin to address pathogen-host interactions during this unique human-restricted infection, we applied a transcript capture and amplification technique, Selective Capture of Transcribed Sequences (SCOTS) [22], along with microarray technology, to assess whether we could detect S. Paratyphi A mRNA directly in the blood of bacteremic patients. To identify genes whose expression might be potentially regulated in vivo, we also compared the relative level of bacterial genes detected in vivo to those detected using the same technique on in vitro grown organisms. We then used RT-qPCR to validate expression of genes detected in our screen.

SCOTS has previously been used in gene expression studies of a number of organisms, including analyses of S. Typhi, S. Typhimurium, and Mycobacterium tuberculosis, using ex vivo human macrophage models [13,14,16,23]. SCOTS has also been applied to evaluate bacterial gene expression in Helicobacter pylori in gastric mucosa biopsies of infected humans, and Heminophilus ducreyi in pustules of infected humans [24,25]. To our knowledge, however, no previous study has used this approach to assess bacterial gene expression directly in the blood of infected humans.

Methods

Ethics statement

This study was conducted according to the principles expressed in the Declaration of Helsinki. We obtained written consent from each patient prior to participation. Written informed consent was obtained from parents authorizing the participation of their children in the study and assent was obtained from children greater than 5 years old. This study was approved by the ethical and research review committees of the ICDDR,B and the Human Research Committee of Massachusetts General Hospital.

Study subject selection, sample collection, and recovery of organisms

Individuals (3–59 years of age) presenting to the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) Hospital or Dhaka Medical College Hospital with fever of 3–7 days duration (≥39°C), without an obvious focus of infection, and lacking an alternate diagnosis were eligible for enrollment. After obtaining written consent, we collected venous blood (5 ml from children <5 years, and 10 ml from all others), and immediately placed 2 ml of blood into TRizol (Invitrogen Life Technologies, Carlsbad, CA) at a 1 (blood):2 (TRizol) volume ratio. We mixed these samples and stored them at −70°C for later analysis. We next cultured 3–5 ml of day 0 blood using a BacT/Alert automated system, sub-culturing positive bottles on MacConkey agar, and identifying isolates using standard biochemical tests and reaction with Salmonella specific antisera [26]. Following collection of blood, patients were initially treated with oral ciprofloxacin or cefixime, or injectable ceftriaxone, and antibiotics were continued for up to 14 days at the discretion of the attending physician.

cDNA synthesis and amplification

To create cDNA of organisms in the blood of bacteremic patients (in vivo sample), we used TRizol preserved blood samples of patients whose day 0 culture subsequently grew S. Paratyphi A. To generate corresponding in vitro cDNA samples, we grew each patient’s bacterial isolate in Luria Bertani (LB) broth until mid-log growth phase (OD600 0.4–0.6), and then immediately placed these samples into TRizol at a 1 (mid-log culture):2 (TRizol) volume ratio. We recovered total RNA from TRizol preserved in vivo and in vitro samples per manufacturer’s instructions (Invitrogen), and treated with DNase I on RNAse columns (Qiagen). We converted 5 μg of total extracted RNA into cDNA using random priming (T-PCR) to obtain a representative amplifiable double-stranded cDNA population as described by Froussard et al. [27], with modifications as previously described by Graham et al. [23]. Briefly, we used Superscript III (Invitrogen) to synthesize first strand cDNA with K9 primer (for in vivo sample) or F9 primer (for in vitro sample) primers with a defined 5′ end terminal sequence and a random nonamer at the 3′ end (Supplementary Table S1). We then synthesized second strands using the same primers and Klenow fragment (Invitrogen) according to manufacturer’s instruction. We then PCR-amplified double stranded cDNA with K9 primer lacking random terminal residues for in vivo samples, and F9 primer for in vitro samples for a total of 30 cycles (94°C for...
1 minute, 55°C for 1 minute, 72°C for 45 second cycle, with an initial denaturation at 94°C for 2 minutes).

Genomic DNA biotinylation and rRNA blocking-plasmid construction
SCOTS requires capturing bacterial cDNA from a mixture of host and microbial nucleic acid by solution-phase hybridization to biotinylated bacterial genomic DNA (gDNA) [23]. To generate gDNA, we grew one of the S. Paratyphi A clinical isolates to mid-logarithmic growth phase in LB at 37°C, and purified gDNA using Easy DNA (Invitrogen). We then biotinylated extracted gDNA with photobiotin acetate (Sigma), and sonicated as previously described [22]. To minimize capture of cDNA of ribosomal RNA by biotinylated gDNA, we blocked ribosomal RNA encoding sequences in the gDNA by pre-hybridizing with sonicated DNA fragments from plasmids encoding 16S and 23S rRNA, and then added this denatured mixture to the mixture of denatured biotinylated cDNA samples, separately, as previously described [23]. Briefly, we mixed denatured bacterial genomic DNA (gDNA) [23], and added this denatured mixture to cDNA samples, hybridizing samples overnight at 67°C. We captured biotinylated cDNA-gDNA hybrids using streptavidin-coated magnetic beads (Dynal M-280). After washing samples, we eluted captured cDNA with NaOH, PCR-amplified S. Paratyphi A 16S and 23S rRNA genes, and cloned these products into pKK223-3 to generate pRibDNA_PTA.

SCOTS
We performed three rounds of SCOTS on in vivo and in vitro cDNA samples, separately, as previously described [23]. Briefly, we mixed denatured biotinylated S. Paratyphi A gDNA with blocking pRibDNA_PTA, and added this denatured mixture to cDNA samples, hybridizing samples overnight at 67°C. We captured biotinylated cDNA-gDNA hybrids using streptavidin-coated magnetic beads (Dynal M-280). After washing samples, we eluted captured cDNA with NaOH, PCR-amplified in vivo and in vitro cDNA samples with K9 or F9 primers, respectively, and purified products using Qiagen PCR column purification kits. We performed three hybridization and amplification cycles to obtain bacterial cDNAs for microarray hybridization. cDNA hybridization-microarray analysis
We differentially labelled in vivo and in vitro SCOTS-cDNAs for each of the three patients with S. Paratyphi A bacteremia [28], and added these products to activated Salmonella ORF microarray glass slides (version STv7S; McClelland Laboratory, Vaccine Research Institute of San Diego, CA, http://www.sdibr.org/Faculty mcclelland/mcclelland-lab), as previously described [13,29]. Microarrays contained gene-specific PCR-products of 4,271 ORFs from Salmonella enterica strains including ~97% identical orthologues of more than 95% of annotated S. Paratyphi A SARB42 genes. We performed labeling and hybridization in duplicate, with dye reversals. We quantified signal intensities using ScanArray software (ScanArray express, version 3.0.1). To assess potential biases in identifying potentially differentially in vivo-expressed genes, we also probed slides comparing SCOTS-product generated from in vivo samples to labeled cDNA from in vitro samples to which SCOTS was not applied. Additionally, we directly compared in vivo cDNAs slides to which SCOTS was and was not applied using analyses described below and similar to Faucher et al. [13], which confirmed that no significant biases were introduced using this approach. For analyses, we subtracted local background from spot signal intensities, and considered a cDNA for an ORF detected in a particular sample if it met the following criteria: 1) the median signal intensity of at least 2 of its 3 replicate spots on the array was at least ten median absolute deviations greater than the median of spots on the microarray corresponding to genes absent from the S. Paratyphi A genome, and 2), this criterion was met on greater than 75% of slides for that subject. We then evaluated differences in expression in in vivo versus in vitro grown organisms for all genes detected in in vivo samples. Using LOESS-normalized, log-transformed data, we employed repeated measures ANOVA (to within slide replicate spots) with fixed type (in vivo versus in vitro) and dye effects with Benjamini-Hochberg correction. We only considered array features with a coefficient of variation in signal intensity of less than 50% within an array. We considered significant variations in signal intensity as determined by ANOVA indication of genes potentially 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 GSE22958. Functional classification of genes was based on J. Craig Venter Institute annotations (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi).

Quantitative PCR analysis
We selected a subset of genes identified in our microarray analyses for confirmation by quantitative real time PCR (RT-qPCR). We compared steady state mRNA levels in specimens from two of the patients analyzed by SCOTS-array hybridization, and two additional patients bacteremic with S. Paratyphi A that were not included in the SCOTS analyses. We compared in vivo expression levels to those present in in vitro samples using three bench-top culture replicates of the S. Paratyphi A isolate from Patient 3 grown to mid-logarithmic phase in LB, as described above. We included in our selection criteria several genes of interest with high baseline signals and fold-changes by SCOTS-array analysis. For comparison, we used SPA3294 (encoding 50S ribosomal protein L5) as a housekeeping gene with no differential detection by SCOTS-array analysis. We had insufficient sample from patient #1 to perform RT-qPCR for all selected genes, and for this individual, we performed RT-qPCR analysis for only three of the six selected genes. We used SuperScript II (Invitrogen) with random hexamers (Sigma), according to manufacturer’s instructions to generate cDNA, 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 [30]. Primers are listed in Table S1. For each sample, we used no-template controls and samples lacking reverse transcriptase as baseline reactions. After initial denaturation at 95°C for 3 min, the RT-PCR cycle was as follows: denaturation at 94°C for 30 seconds (s), extension at 58°C for 30 s, extension at 72°C for 1 minute, followed by a plate read. We repeated the cycle 40 times, set the calculated threshold cycle (CT) in the low/linear portion of product curves, and quantified gene copy numbers using pGEM-T Easy-based plasmids (Promega) containing the gene of interest [30]. We calculated control gene copy numbers using plasmid size and A260 readings, and normalized gene copy numbers against cDNA copies of 16S rRNA. We assessed singularity of product species and size by melting curve analysis.

Results
Detection of S. Paratyphi A mRNA in the blood of infected humans
We isolated serotype Paratyphi A in the blood of 5 of 89 individuals who met study criteria. In total, we detected transcripts of 1798 S. Paratyphi A genes in the blood of infected humans, or 43.9% of ORFs in the S. Paratyphi ATCC9150 genome. Of these, we detected 868 transcripts in at least two patients, and 315 in all three patients (Figure 1A and Table S2). Detected transcripts are predicted to encode products that could be categorized into a number of functional groups (Figure 2A). The largest grouping was “unknown/unclassified” (338 transcripts; 18.8% of detected mRNA/cDNAs; 8.3% of S. Paratyphi A ORFs). Other large
groupings included transcripts of genes associated with pathogenesis (lphOQ, ptoS), nutrient acquisition, energy metabolism, biosynthesis of the essential vitamins biotin (bioABF, thiFI) and thiamine (thiC), iron acquisition (genes in the iroA cluster, fepD, yheA) and stress responses (katE, groEL, groES, dnaK) including oxidative stress resistance (katE, umuCD). Forty-three of the transcripts corresponded to genes contained within Salmonella Pathogenicity Islands (SPIs) 1–4, 6, 10, 13, and 16 (annotated in column E of Table S2).

Comparison of levels of S. Paratyphi A transcripts detected using in vivo versus in vitro samples

mRNA/cDNA of 11 genes were detected only in in vivo samples of at least two patients, and not detected in any in vitro sample (Table 1). These included transcripts of genes involved in biosynthesis of essential vitamins such as biotin (bioF), utilization of alternative carbon sources including ethanolamine (cat operon) and propane-diol (pdu operon), iron acquisition (fepD), and resistance to antimicrobial agents (yelO, also known as mcdC). When considering genes whose expression was detected by SCOTS array in both in vivo and in vitro samples, we identified transcripts of 910 genes (22.3% of ORFs) with significant and at least 2-fold difference in signal detection between in vivo versus in vitro samples: 347, 367, and 443 in patients 1, 2, and 3, respectively (Figure 1B and annotated in column Y, AS, and AE of Table S2). We detected transcripts of 206 genes in at least two patients, expression of 194 of which were concordant (4.7% ORFs), and 41 in all three patients (Table 2). We categorized the 194 concordant gene products detected in at least two patients into a number of functional groups (Figure 2B). The largest grouping was “unknown/unclassified” (44 genes). We identified transcripts of 12 genes were located within SPIs (SPI-1, 3, 6, 10, 16 [12,51–34], a number of transcripts corresponded to genes within the PhoP/Q regulatory system involved in intramorphophage survival [20,35–37], and five genes were in phages unique to S. Paratyphi A, including SPA-2 and SPA-3-P2 [12].

Quantitative Real Time-PCR analysis

To further analyze expression levels of genes identified in our screen, we used quantitative RT-PCR to assess relative steady state mRNA levels for some of the genes of interest identified by SCOTS-array. We chose five genes with high baseline signals and difference in detection signal by SCOTS-array analysis. These included SPA0410 (outR), encoding an ethanolamine ammonia lyase; SPA1451 (sesE), encoding a secreted effector protein and located within SPI-2; SPA2748 (spaP), encoding a secreted protein located within SPI-1; SPA3315 (yheL), encoding a sulfur transfer complex subunit; and SPA3373, a putative cytoplasmic protein. We also assessed SPA3294 transcript levels, encoding ribosomal protein L3, as a representative housekeeping gene. We performed RT-qPCR analyses using blood collected at the initial clinical encounter and immediately preserved in TRIZol. We had sufficient sample to perform RT-qPCR on initial blood samples of SCOTS patients 1 and 3, and performed additional RT-qPCR analysis on two additional patients bacteremic with S. Paratyphi A whose blood samples were not analyzed by SCOTS-cDNA hybridization (patients 4 and 5). As shown in figure 3, we found significantly increased expression of all five analyzed candidate genes in vivo, but not for housekeeping gene SPA3294.

Discussion

S. Paratyphi A now accounts for a fifth of all cases of enteric fever in many areas of Asia, and it is therefore a significant and emerging global public health concern. S. Paratyphi A in vivo biology is difficult to study due to a lack of an animal model that fully replicates human infection. In this study, we used an mRNA/cDNA capture and amplification technique, microarray technology, and quantitative PCR to detect S. Paratyphi A transcripts in the blood of bacteremic patients in Bangladesh. We identified a subset of S. Paratyphi A genes with higher level of transcription in vivo compared to in vitro samples, suggesting possible in vivo induction of these genes, and we confirmed increased in vivo expression of a subset of these genes using RT-qPCR. SCOTS and SCOTS/microarray technologies have previously been used to analyze S. Typhimurium and S. Typhi gene expression using ex vivo models [13,14,16,18], but to our knowledge, no group has used this or a similar technology to assess bacterial gene expression directly in the blood of bacteremic patients, and no group has assessed S. Paratyphi A gene expression in various conditions.

We could categorize the 868 S. Paratyphi A genes expressed in the blood stream of two or more bacteremic patients into a number of groupings that may reflect transcriptional adaptations
of *S. Paratyphi A* to the *in vivo* environment. These include adaptations involved in intra-macrophage survival, for example altering energy metabolism, nutrient acquisition, and resisting both oxidative killing and antimicrobial peptides. Forty-three expressed genes were located within known *Salmonella* Pathogenicity Islands including SPI 1-4, 6, 10, 13, and 16 [12,31–34].

Figure 2. Functional classifications of products of *S. Paratyphi A* genes whose mRNA/cDNA was detected in this study. Functional classification of products of transcripts (A) detected in *in vivo* samples in two or more patients, and (B) associated with significantly different levels of detection in *in vivo* versus *in vitro* samples in two or more patients using SCOTS-array.

doi:10.1371/journal.pntd.0000908.g002
SPI-1 encodes a type III secretion system involved in intestinal epithelial cell invasion [38]. Our identification of SPI-1 gene transcripts in the blood stream of infected humans, including genes of the sip and spa gene clusters, suggests a potential role for these genes in eukaryotic cell invasion outside of the intestinal epithelium. SPI-2 encodes a type III secretion system (TTSS) that is involved in survival of Salmonella within macrophages, including creating and maintaining Salmonella-containing vacuoles (SCV) [39]. Among the SPI-2 genes detected in our analyses were TTSS-associated effector genes contained within the ssa and sse operons. SPI-3 also contains a number of genes involved in pathogenesis. We detected in vivo expression of the mgtB(C) operon, encoding a magnesium transporter, within this island, and magnesium availability is a key signal for Salmonella in the intracellular environment [40]. Transcription in this operon is regulated by PhoPQ and is required for survival within macrophages and low Mg²⁺ conditions [40]. We also identified the transcript of the SPI-3-associated gene misL, an autotransporter possibly involved in cellular adhesion [41]. Within SPI-4, we detected in vivo expression of sseE, an adhesin that aids in efficient translocation of SPI-1 effectors through involvement in apical membrane ruffling of epithelial cells [42]. SPI-6 appears to be involved in intra-macrophage survival, including encoding proteins involved in metabolic pathways, nutrient acquisition, and utilization of alternate carbon sources within macrophages [43]. Our identification of ybeT involved in carbohydrate transport and metabolism is consistent with this model. SPI-6 also encodes a number of fimbrial proteins involved in adherence and virulence in a number of Salmonella animal models [44], and we identified transcripts of tssB/C and safC, encoding fimbrial proteins, supporting the proposal that these factors may play a role at various stages of Salmonella infection. Faucher et al. identified SPI-10 in their application of SCOTS to S. Typhi within ex vivo macrophages [14], and we also detected a number of transcripts of S. Paratyphi A genes within SPI-10, many of which are uncharacterized. Interestingly, SPI-10 is present in S. Typhi and S. Paratyphi A, but absent from most strains of S. Typhimurium and all strains of S. Paratyphi B and S. Paratyphi C examined to date [43]. We also detected transcripts of three genes (asaG, evaT, SPA2997) contained within the conserved portion of SPI-13, and gtB, involved in O-antigen glycosylation, and contained within SPI-16 [31].

We detected mRNA/cDNA of a number of genes expressed under the control of major virulence regulatory systems in Salmonella, including the PhoP, RpoS, and SlyA systems involved in intra-macrophage survival and virulence [35,46,47]. Within the PhoPQ cascade, we detected transcripts of phoPQ itself, lpxO, a dioxygenase involved in lipid synthesis [20]; mgtBC, mgtA, a possible pseudogene in S. Paratyphi A [36]; virK encoding the VirK virulence protein [37]; and bshF involved in biotin synthesis [20]. Within the RpoS regulatory cascade, we detected transcripts of rpoS itself, katE and 3klk involved in resistance to oxidative stress [48,49], and narZTV involved in nitrate reduction [48,50]. Of note, mutations in the narZ operon are associated with decreased virulence of S. Typhimurium in mice [50]. Within the SlyA regulatory system also involved in intra-macrophage survival of Salmonella, we detected expression of gnoEL, a chaperone protein also regulated by PhoP [40,51].

Many of the genes that we identified as in vivo expressed and potentially differentially expressed (compared to ex vivo conditions) are associated with adaptation to the likely nutrient-altered environment of the macrophage. For instance, we detected mRNA of 7 of the 17 genes of the evt operon involved in ethanolamine utilization, providing alternate sources of carbon and/or nitrogen [52]. We also detected expression of the pduBCJK genes involved in alternate carbon source propanediol utilization [43], and genes involved in citrate and tartrate fermentation, including oxaloacetate decarboxylase genes oadA and oadB [53]. Of note, oadA mutants of Legionella pneumophila have impaired replication and survival within macrophages [54]. We detected transcripts of dsmC, encoding a dimethyl sulfoxide reductase involved in bacterial survival in anaerobic conditions, similar to that probably encountered within Salmonella-containing vacuoles, and genes involved in the response to phosphate-limited conditions, including the phosphate transporter genes mgtAE and pstCA-phoU [55].

S. Paratyphi are becoming increasingly resistant to antimicrobial agents, especially in Asia [4], and all S. Paratyphi A strains isolated in this study were resistant to nalidixic acid and intermediately susceptible to ciprofloxacin, despite the fact that no antibiotics were administered prior to collection of blood in this study. In our analysis, we detected mRNA/cDNA of several bacterial genes in vivo samples involved in resistance to a number of antimicrobial agents, including nalidixic acid, novobiocin, tetracycline, and norfloxicin [56]. Mechanisms of resistance encoded by detected genes included alteration of target molecules such as the protein encoded by basA that confers resistance to bacitracin [57], and multidrug efflux systems of the resistance-nodulation-division (RND)-type systems (acrB, sfpC, ybeO/mdc) and the major facilitator system (emrA, emrE) [56,58].

Of the 1798 in vivo-expressed genes identified, 910 had statistically significant differences in signal detection when comparing in vivo to in vitro samples (approximately 50% of detected transcripts; 22.2% of S. Paratyphi A ORFs); a figure in concordance with the 36% identified in an ex vivo macrophage model of S. Typhi by Faucher et al. [13]. 194 were concordantly differentially detected in at least two patients, and 41 in all three patients. Many of these 41 genes are involved in energy metabolism, including evaT, asaG, and ybeT; survival in metal ion limiting conditions including the Mn²⁺ transport-associated gene mntF; and biosynthesis of essential molecules, including thiC involved in thiamine synthesis. Interestingly, we also identified transcripts of spaP at higher levels of detection in all in vivo versus in vitro samples, and subsequently confirmed this increased expression in vivo using RT-qPCR. spaP is contained within SPI-1 and is thought to be involved in invasion of epithelial cells. Our results

<table>
<thead>
<tr>
<th>Table 1. S. Paratyphi A (SPA) genes whose transcripts were detected only in in vivo samples of at least two patients and not in any in vitro sample.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPA locus</td>
</tr>
<tr>
<td>SPA0405</td>
</tr>
<tr>
<td>SPA0612</td>
</tr>
<tr>
<td>SPA0737</td>
</tr>
<tr>
<td>SPA0826</td>
</tr>
<tr>
<td>SPA1856</td>
</tr>
<tr>
<td>SPA1957</td>
</tr>
<tr>
<td>SPA2142</td>
</tr>
<tr>
<td>SPA2483</td>
</tr>
<tr>
<td>SPA2570</td>
</tr>
<tr>
<td>SPA2574</td>
</tr>
<tr>
<td>SPA3456</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pntd.0000908.t001
<table>
<thead>
<tr>
<th>Main Role</th>
<th>SPA locus</th>
<th>SPA gene</th>
<th>SPA description</th>
<th>Pt1 FC</th>
<th>Pt 1 pv</th>
<th>Pt2 FC</th>
<th>Pt 2 pv</th>
<th>Pt3 FC</th>
<th>Pt 3 pv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogenesis/Cellular processes</td>
<td>SPA0452</td>
<td>mntH</td>
<td>Manganese transport protein MntH</td>
<td>3.11</td>
<td>0.0459</td>
<td>2.06</td>
<td>0.0265</td>
<td>2.75</td>
<td>0.0196</td>
</tr>
<tr>
<td></td>
<td>SPA2748</td>
<td>spaP</td>
<td>Secretory protein (associated with virulence)</td>
<td>10.93</td>
<td>0.0027</td>
<td>1.36</td>
<td>0.0400</td>
<td>2.03</td>
<td>0.0379</td>
</tr>
<tr>
<td>Amino acid biosynthesis</td>
<td>SPA0426</td>
<td>cytM</td>
<td>Cysteine synthase B</td>
<td>5.39</td>
<td>0.0340</td>
<td>3.66</td>
<td>0.0468</td>
<td>2.24</td>
<td>0.0323</td>
</tr>
<tr>
<td>Biosynthesis of cofactors, prosthetic groups, and carriers</td>
<td>SPA2972</td>
<td>yggH</td>
<td>Conserved hypothetical protein</td>
<td>7.43</td>
<td>0.0027</td>
<td>1.47</td>
<td>0.0445</td>
<td>1.51</td>
<td>0.0335</td>
</tr>
<tr>
<td>Cell envelope</td>
<td>SPA4001</td>
<td>thiC</td>
<td>Thiamine biosynthesis protein</td>
<td>3.48</td>
<td>0.0099</td>
<td>2.67</td>
<td>0.0253</td>
<td>2.78</td>
<td>0.0496</td>
</tr>
<tr>
<td>Central intermediary metabolism</td>
<td>SPA0545</td>
<td>nuoK</td>
<td>NADH dehydrogenase I chain k</td>
<td>9.54</td>
<td>0.0047</td>
<td>5.70</td>
<td>0.0295</td>
<td>3.97</td>
<td>0.0193</td>
</tr>
<tr>
<td>Energy metabolism</td>
<td>SPA0406</td>
<td>eutJ</td>
<td>Putative ethanolamine utilization protein EutJ</td>
<td>7.73</td>
<td>0.0029</td>
<td>7.96</td>
<td>0.0391</td>
<td>3.04</td>
<td>0.0250</td>
</tr>
<tr>
<td></td>
<td>SPA0410</td>
<td>eutB</td>
<td>Ethanolamine ammonia-lyase heavy chain</td>
<td>5.07</td>
<td>0.0031</td>
<td>3.05</td>
<td>0.0322</td>
<td>2.40</td>
<td>0.0228</td>
</tr>
<tr>
<td></td>
<td>SPA0828</td>
<td>pduG</td>
<td>PduG protein</td>
<td>10.54</td>
<td>0.0030</td>
<td>2.91</td>
<td>0.0244</td>
<td>3.86</td>
<td>0.0220</td>
</tr>
<tr>
<td></td>
<td>SPA3367</td>
<td>gntR</td>
<td>Gluconate utilization operon repressor</td>
<td>9.75</td>
<td>0.0027</td>
<td>1.16</td>
<td>0.0403</td>
<td>2.09</td>
<td>0.0206</td>
</tr>
<tr>
<td></td>
<td>SPA3527</td>
<td>yiaR</td>
<td>Putative sugar-phosphate isomerase</td>
<td>3.62</td>
<td>0.0046</td>
<td>7.61</td>
<td>0.0363</td>
<td>6.29</td>
<td>0.0497</td>
</tr>
<tr>
<td></td>
<td>SPA3706</td>
<td>atpA</td>
<td>ATP synthase alpha subunit</td>
<td>3.76</td>
<td>0.0047</td>
<td>2.90</td>
<td>0.0487</td>
<td>1.03</td>
<td>0.0335</td>
</tr>
<tr>
<td>Fatty acid and phospholipid</td>
<td>SPA1655</td>
<td>acpP</td>
<td>Acyl carrier protein</td>
<td>9.57</td>
<td>0.0068</td>
<td>3.04</td>
<td>0.0237</td>
<td>3.05</td>
<td>0.0392</td>
</tr>
<tr>
<td>metabolism</td>
<td>SPA1739</td>
<td>yccD</td>
<td>Conserved hypothetical protein</td>
<td>2.87</td>
<td>0.0041</td>
<td>2.45</td>
<td>0.0278</td>
<td>3.82</td>
<td>0.0461</td>
</tr>
<tr>
<td>Hypothetical proteins</td>
<td>SPA3373</td>
<td>secF</td>
<td>Protein-export membrane protein SecF</td>
<td>5.30</td>
<td>0.0030</td>
<td>1.49</td>
<td>0.0292</td>
<td>1.01</td>
<td>0.0348</td>
</tr>
<tr>
<td>Mobile and extrachromosomal</td>
<td>SPA3288</td>
<td>rpmD</td>
<td>50S ribosomal subunit protein L30</td>
<td>7.81</td>
<td>0.0029</td>
<td>4.42</td>
<td>0.0275</td>
<td>3.47</td>
<td>0.0217</td>
</tr>
<tr>
<td>element functions</td>
<td>SPA3289</td>
<td>rpsE</td>
<td>30S ribosomal subunit protein S5</td>
<td>7.16</td>
<td>0.0042</td>
<td>4.19</td>
<td>0.0276</td>
<td>4.03</td>
<td>0.0196</td>
</tr>
<tr>
<td>Protein fate</td>
<td>SPA3290</td>
<td>rplR</td>
<td>50S ribosomal subunit protein L18</td>
<td>5.42</td>
<td>0.0028</td>
<td>4.94</td>
<td>0.0272</td>
<td>4.03</td>
<td>0.0213</td>
</tr>
<tr>
<td></td>
<td>SPA3291</td>
<td>rplF</td>
<td>50S ribosomal subunit protein L6</td>
<td>3.70</td>
<td>0.0043</td>
<td>1.98</td>
<td>0.0332</td>
<td>4.53</td>
<td>0.0225</td>
</tr>
<tr>
<td></td>
<td>SPA3304</td>
<td>rplW</td>
<td>50S ribosomal subunit L23</td>
<td>11.35</td>
<td>0.0036</td>
<td>5.71</td>
<td>0.0429</td>
<td>5.62</td>
<td>0.0273</td>
</tr>
<tr>
<td></td>
<td>SPA3305</td>
<td>rplD</td>
<td>50S ribosomal subunit L4</td>
<td>13.12</td>
<td>0.0094</td>
<td>3.77</td>
<td>0.0245</td>
<td>2.88</td>
<td>0.0185</td>
</tr>
<tr>
<td></td>
<td>SPA4208</td>
<td>rpsF</td>
<td>30S ribosomal protein S6</td>
<td>9.06</td>
<td>0.0025</td>
<td>5.87</td>
<td>0.0234</td>
<td>4.46</td>
<td>0.0205</td>
</tr>
<tr>
<td>Regulatory functions</td>
<td>SPA0293</td>
<td>yfiH</td>
<td>Putative transcriptional regulator</td>
<td>4.78</td>
<td>0.0367</td>
<td>2.40</td>
<td>0.0259</td>
<td>2.81</td>
<td>0.0338</td>
</tr>
<tr>
<td></td>
<td>SPA0905</td>
<td>yedE</td>
<td>Putative membrane protein</td>
<td>3.68</td>
<td>0.0056</td>
<td>5.55</td>
<td>0.0294</td>
<td>3.51</td>
<td>0.0261</td>
</tr>
<tr>
<td>Transport and binding</td>
<td>SPA0056</td>
<td>oadA</td>
<td>Oxaloacetate decarboxylase alpha chain</td>
<td>7.21</td>
<td>0.0037</td>
<td>2.45</td>
<td>0.0259</td>
<td>2.38</td>
<td>0.0201</td>
</tr>
<tr>
<td></td>
<td>SPA0918</td>
<td>yecC</td>
<td>Putative ABC transport ATP-binding protein</td>
<td>2.43</td>
<td>0.0035</td>
<td>4.01</td>
<td>0.0256</td>
<td>2.64</td>
<td>0.0262</td>
</tr>
<tr>
<td></td>
<td>SPA1531</td>
<td>celB</td>
<td>PTS system, cellobiose-specific IIC component</td>
<td>9.19</td>
<td>0.0028</td>
<td>2.88</td>
<td>0.0274</td>
<td>2.39</td>
<td>0.0294</td>
</tr>
<tr>
<td></td>
<td>SPA2035</td>
<td>kdpA</td>
<td>Potassium-transporting ATPase A chain</td>
<td>6.80</td>
<td>0.0032</td>
<td>5.58</td>
<td>0.0268</td>
<td>2.95</td>
<td>0.0499</td>
</tr>
<tr>
<td></td>
<td>SPA3336</td>
<td>muted</td>
<td>Mannitol-specific enzyme II of phosphotransferase system</td>
<td>6.24</td>
<td>0.0035</td>
<td>1.11</td>
<td>0.0285</td>
<td>4.39</td>
<td>0.0272</td>
</tr>
<tr>
<td>Unclassified</td>
<td>SPA0411</td>
<td>eutC</td>
<td>Ethanolamine ammonia-lyase light chain</td>
<td>8.68</td>
<td>0.0217</td>
<td>3.42</td>
<td>0.0295</td>
<td>1.50</td>
<td>0.0450</td>
</tr>
<tr>
<td>Unknown function</td>
<td>SPA0409</td>
<td>eutA</td>
<td>Putative ethanolamine utilization protein EutA</td>
<td>8.48</td>
<td>0.0028</td>
<td>9.61</td>
<td>0.0286</td>
<td>4.76</td>
<td>0.0265</td>
</tr>
<tr>
<td>Pseudogene</td>
<td>SPA2109</td>
<td>dipB</td>
<td>Sensor kinase DipB pseudogene</td>
<td>4.64</td>
<td>0.0333</td>
<td>1.85</td>
<td>0.0333</td>
<td>1.75</td>
<td>0.0252</td>
</tr>
<tr>
<td></td>
<td>SPA3223</td>
<td>yiaR</td>
<td>Putative membrane transport protein pseudogene</td>
<td>11.35</td>
<td>0.0044</td>
<td>2.08</td>
<td>0.0262</td>
<td>2.80</td>
<td>0.0249</td>
</tr>
<tr>
<td></td>
<td>SPA3638</td>
<td>rhpC</td>
<td>Regulatory protein pseudogene</td>
<td>7.43</td>
<td>0.0029</td>
<td>3.61</td>
<td>0.0246</td>
<td>2.09</td>
<td>0.0335</td>
</tr>
<tr>
<td></td>
<td>SPA3755</td>
<td>rhlB</td>
<td>Putative ATP-dependent RNA helicase pseudogene</td>
<td>5.03</td>
<td>0.0341</td>
<td>1.43</td>
<td>0.0315</td>
<td>2.18</td>
<td>0.0260</td>
</tr>
</tbody>
</table>

SPA: S. Paratyphi A.  
PV: P value.  
doi:10.1371/journal.pntd.0000908.t002
suggest a potential role of SpaP in eukaryotic cell invasion beyond the intestinal epithelium.

In conclusion, we have used a capture-amplification and microarray approach to assess gene expression for a human-restricted pathogen, S. Paratyphi A, directly in the blood of infected humans. We detected transcripts of many genes contained within known Salmonella pathogenicity islands and genes controlled by the PhoPQ, RpoS and SlyA regulons required for intramacrophage survival. We detected expression of many genes involved in energy metabolism, nutritional acquisition, protein

Figure 3. Quantitative-PCR mRNA expression profiles. Quantitative-PCR mRNA expression profiles during in vitro and in vivo growth of S. Paratyphi A of genes associated with significantly different levels of detection in in vivo versus in vitro samples using SCOTS-array analysis. Quantitative RT-PCR gene profiles of S. Paratyphi A genes (A–E) comparing RNA recovered from blood of bacteremic patients to in vitro cultures, and a house-keeping gene with no differential expression by SCOTS-microarray analysis (F). Mean copies of mRNA per copy of 16S rRNA, standard error of the mean (SEM), and fold-induction (G) are presented. * p<0.05; † p<0.01; ‡ p<0.001.

doi:10.1371/journal.pntd.0000908.g003
synthesis, fatty acid and phospholipid metabolism, pathogenesis, transport and bind, regulation, SOS responses, and antimicrobial resistance. These functional categories may reflect bacterial modifications required for survival within infected humans. Of note, we also identified expression of a large number of genes with currently unknown function. We further identified a subset of genes whose transcripts had altered detection in vivo versus in vitro samples, suggesting potential regulation of these genes within the human host, and we confirmed induction of a subset of these genes in vivo. The variability between patients that we observed may relate in part to the low level of bacterial mRNA present, variations introduced by our capture and amplification technology, differences in infecting strains or growth phase in vivo, and differences in bacterial location (intra-macrophase versus extracellular). Despite this variability, our results have given insight into bacterial responses in humans infected with S. Paratyphi A, and have identified genes for future analysis, including drug target development. These results suggest that similar approaches may be applied to other pathogens in infected humans and animals.

References


Supporting Information

Table S1 Sequences of primers used in this study. Found at: doi:10.1371/journal.pntd.0000908.s001 (0.04 MB DOC)

Table S2 8. Paratyphi A genes whose mRNA/cDNA was detected in the blood of infected humans in this study. Found at: doi:10.1371/journal.pntd.0000908.s002 (0.63 MB XLS)

Author Contributions

Conceived and designed the experiments: AS RCC SMR JBH MSB WAB RCL ELH AC MM JEG FQ ETR. Performed the experiments: AS SMR MSB FK AB. Analyzed the data: AS RCC SMR JBH FK RCL TL SBC MM JEG FQ ETR.

www.plosntds.org 9 December 2010 | Volume 4 | Issue 12 | e908

SCOTS Paratyphi A