Genetic determinants in *Salmonella enterica* serotype Typhimurium required for overcoming stressors in the host environment

Rabindra K. Mandal[1*], Tieshan Jiang[1] and Young Min Kwon[1, 2]

[1]Department of Poultry Science, University of Arkansas, Fayetteville, AR, 72701

[2]Cell and Molecular Biology Program, University of Arkansas, Fayetteville, AR, 72701

*Corresponding author's current address:

Department of Microbiology and Immunology

School of Medicine

University of Louisville

Louisville, KY 40202

Phone: (479) 276-2203

Email: rabindra.mandal@louisville.edu

Abstract

Salmonella enterica serovar Typhimurium (*S.* Typhimurium), a non-typhoidal *Salmonella* (NTS), result in a range of diseases, including self-limiting gastroenteritis, bacteremia, enteric fever, and focal infections representing a major disease burden worldwide. There is still a significant portion of *Salmonella* genes whose functional basis to overcome host innate defense mechanisms, consequently causing disease in host, largely remains unknown. Here, we have applied a high-throughput transposon sequencing (Tn-seq) method to unveil the genetic factors required for the growth or survival of *S.* Typhimurium under various host stressors simulated *in vitro*. A highly saturating Tn5 library of *S.* Typhimurium 14028s was subjected to selection during growth in the presence of short chain fatty acid (100 mM propionate), osmotic stress (3% NaCl) or oxidative stress (1 mM H_2O_2) or survival in extreme acidic pH (30 min in pH3) or starvation (12 days in 1X PBS). We have identified an overlapping set of 339 conditionally essential genes (CEGs) required by *S.* Typhimurium to overcome these host insults. Interestingly, entire eight genes encoding F_0F_1-ATP synthase subunit proteins were required for fitness in all five stresses. Intriguingly, total 88 genes in *Salmonella* pathogenicity island (SPI), including SPI-1, SPI-2, SPI-3, SPI-5, SPI-6 and SPI-11 are also required for fitness under the *in vitro* conditions evaluated in this study. Additionally, by comparative analysis of the genes identified in this study and the genes previously shown to be required for *in vivo* fitness, we identified novel genes (*marBCT*, *envF*, *barA*, *hscA*, *rfaQ*, *rfbI* and putative proteins STM14_1138, STM14_3334, STM14_4825, and STM_5184) that has compelling potential to be exploited as vaccine development and/or drug target to curb the *Salmonella* infection.

Key Words: *Salmonella*, host stress, Tn-seq, conditionally essential genes, *in vitro* fitness

Introduction

Non-typhoidal *Salmonella* (NTS), a Gram-negative bacterial pathogen, causes 93 million enteric infections, 155,000 diarrheal deaths, and 3.4 million blood stream infection worldwide annually (Ao et al., 2015; Majowicz et al., 2010). Gram-negative bacterial pathogens, including NTS, are developing resistance against antimicrobial agents including the last resort antibiotics at a startling rate, creating a global crisis in human health. Scientists fear the impending global epidemic of untreatable infections and return to a pre-antibiotic era where a common infection and minor injury can be lethal (Liu et al., 2015; McKenna, 2013; Spencer, 2015; World Health Organization (WHO). Thus, there is an urgent need to identify genetic factors of pathogenic microorganisms that can serve as targets to develop novel strategies to combat infectious diseases (Medini et al., 2008; van Opijnen and Camilli, 2012). Nonetheless, the insufficiency of the genome-wide data that provide links between genotype and the infection-related phenotypes of bacteria is the major roadblock to discover suitable targets for development of the effective strategies to control infection.

Salmonella enterica serotype Typhimurium (*S.* Typhimurium) is one of the leading cause of NTS (Carden et al., 2015; Crim et al., 2015). Despite *Salmonella* infection has an enormous global burden on disease worldwide and availability of complete genome sequence of *S.* Typhimurium LT2 nearly one and half decade (2002) ago, the phenotypic basis of *S.* Typhimurium genes required for *in vivo* survival is still unknown for a large portion of the genes (Feasey et al., 2012; McClelland et al., 2001). Researchers have tried to delve into the pathogenesis of *S.* Typhimurium using different variations of high throughput screening of transposon mutants, with a limited number of mutants based on a negative selection (Kwon et al., 2016). Chan et al., (2005) had discovered 157 and 264 genes required by *S.* Typhimurium strain SL1344 for acute

63 infection in mice (A-Mice) and survival inside macrophage (MΦ), respectively using a
64 microarray-based tracking method (Chan et al., 2005). Lawley et al., (2006) used the same
65 method to identify 118 genes of *S.* Typhimurium SL1344 required for long-term persistent
66 infection in mice (P-Mice) using the spleen samples collected after 28 day post infection (Lawley
67 et al., 2006). Additionally, Chaudhuri et al. (2013) have comprehensively assigned a core set of
68 611 genes of *S.* Typhimurium strain ST4/74 required for effective colonization in the calf, pig,
69 and chicken (Chaudhuri et al., 2013). Recently, Silva-Valenzuela et al., identified 224 mutants of
70 *S.* Typhimurium 14028s that were negatively selected using two pools of single gene deletion
71 mutants from spleen and liver at 2 days post infection in mice (Sp-Liv) (Silva-Valenzuela et al.,
72 2015). Previously, our laboratory conducted Tn-seq screening to identify an overlapping set of
73 105 coding genes of *S.* Typhimurium 14028s required for *in vitro* growth in diluted Luria-
74 Bertani (LB) medium, LB medium plus bile acid and LB medium at 42°C (Khatiwara et al.,
75 2012). However, there is still a gap in the above approaches to correlate *in vivo* and *in vitro*
76 survival or growth genes required by *S.* Typhimurium that will help delve into biochemical and
77 molecular basis of virulence and potentially pave a roadmap towards the efficient development
78 of novel vaccines, antibiotics, and control strategies.

79 In this study, we conducted transposon sequencing (Tn-seq) analysis of *S.* Typhimurium 14028s
80 under the five *in vitro* conditions mimicking host stressors found during enteric and systemic
81 infection. Tn-seq is a powerful tool for functional analysis of bacterial genomes based on the use
82 of random transposon mutagenesis and next generation sequencing technology (Kwon et al.,
83 2016; Van Opijnen et al., 2009; Van Opijnen and Camilli, 2013). We have applied a highly
84 efficient method for Tn-seq library preparation that requires only small amount of DNA without
85 the need for enzymatic digestion or physical shearing of genomic DNA (Dawoud et al., 2014;

Karash et al., 2017; Mandal and Kwon, 2017; Mandal et al., 2017). To cause enteric infection *S.* Typhimurium has to overcome gastrointestinal host insult such as low acidic pH in the stomach, osmotic and short chain fatty acid (SCFAs) in intestine (Ha et al., 1998; Nava et al., 2005; Sleator and Hill, 2002; Smith, 2003). Eventually, for systemic infection, *S.* Typhimurium has to vanquish macrophage stress such as oxidative stress, starvation as well as hyperosmotic condition (Lee et al., 2014; Rosenberger and Finlay, 2003; van der Heijden et al., 2015). We hypothesized that the comparative analysis of the comprehensive sets of the *in vitro* fitness genes (for stress resistance, this study and previous) and *in vivo* (required for enteric and systemic infection in the host) will allow better understanding of the biochemical or phenotypic basis of the genetic requirements of *S.* Typhimurium for host infection and provide enhanced resolution to link genotype to phenotype. Thus, we performed a comparative study between the *in vivo* and *in vitro* fitness genes from previous studies and this study, respectively.

Material and Methods

Bacterial strains and growth conditions

S. Typhimurium 14028s, a spontaneous mutant resistant to nalidixic acid (NA), was grown in Luria-Bertani (LB) plate or LB medium (BD Difco, Sparks, MD) on shaking rack at 225 rpm and incubated at 37°C unless otherwise indicated. Nalidixic acid (NA, ICN Biomedicals Inc., Aurora OH, USA) and Kanamycin (Km, Shelton Scientific, Inc. CT, USA) were used at 25 µg/ml and 50 µg/ml respectively. *S.* Typhimurium was stored in 50% glycerol at -80°C.

Construction of Transposon mutant library

To prepare electrocompetent cells, *S.* Typhimurium was grown overnight in 10 ml LB medium with NA and was diluted 100 fold in 10 ml 2xYT (BD Difco, Sparks, MD, USA) medium with NA and incubated for 3 h on a shaking rack. Bacterial cells were washed 6 times with wash solution (10% glycerol). Centrifugation was done at 8,000 rpm for 1 min at refrigeration temperature (4ºC). The bacterial pellet was mixed gently in 60 µl of wash solution preventing aeration. One µl of the EZ-Tn5 <KAN-2> Tnp transposome complex (Epicentre BioTechnologies, Madison, WI, USA) was added to electrocompetent *S.* Typhimurium cells and incubated on ice for 10 min. Then, the mixture was gently transferred to ice cold cuvette avoiding the formation of any air bubble and electroporated at 2450 V. Immediately, 500 µl of SOC was added and incubated for 90 minutes on a shaking rack at 37°C. The reaction was plated on LB plates supplemented with NA and Km to recover the transformants. With three electroporations we were able to collect 350,000 Tn5 mutants and stored them in LB medium with 50% glycerol at -80ºC (Figure 1).

In vitro growth assay of transposon mutant library

In vitro selection of transposon mutant library was done as described by Opijnen and Camilli, (2010) (van Opijnen et al., 2014) with some modifications. Briefly, transposon mutant library was thawed on ice and an aliquot of 300 µl was added to 60 ml LB broth with NA and Km (OD_{600} = 0.131). The library was incubated at 37°C on a shaking rack for 30 min (OD_{600} = 0.135) and centrifuged at 5,500 rpm for 8 min at room temperature. The transposon mutant library pellet was resuspended in 50 ml 1X phosphate buffer saline (PBS) (OD_{600} = 0.143) and CFU ($4X10^7$/ml) was measured (t_1). This step was included to prepare the mutant cells adapted to LB medium and shorten the lag phase in the following selective conditions. Ten ml aliquot were

saved from t_1 as an input pool (IP1). Above procedure was repeated to make a technical replicate of IP1 as input pool 2 (IP2). An aliquot of 0.5 ml from t_1 was inoculated to 10 ml LB (LB), LB with 3% NaCl (NaCl), LB with 100mM propionate with pH adjusted to pH7 (PA), LB with 1mM H_2O_2 (H_2O_2). The initial OD_{600} of inoculated medium was 0.009. We then incubated the libraries on a shaking rack (225 rpm) at 37°C with variable incubation time ranging from 3.75 h to 7 h (t_2) to a mid-logarithmic. The final OD_{600} of all output pools was very similar around 0.64 at time point t_2. Input pool and output pool libraries were centrifuged and the pellet was stored at -80°C for DNA extraction (Figure 1).

In vitro survival assay of transposon mutant library

To identify genes negatively selected during starvation, an aliquot of 0.5 ml from t_1 was transferred to 10 ml PBS and incubated at 37°C on shaking rack for 12 days. On the 12th day, the tube was centrifuged and the pellet was dissolved in 1 ml PBS. 100 µl aliquot was incubated on LB plate (NA + Km) overnight at 37°C. The cells were collected in PBS and stored at -80°C for DNA extraction. Whereas for survival in pH3, 0.5 ml from t_1 was exposed to LB medium adjusted at pH3 for 30 min at 37°C and immediately transferred to 40 ml PBS. The cells were centrifuged at 8000 rpm for 8 min and pellet was mixed in 1ml PBS. An aliquot of 250 µl was plated on LB plate (NA + Km) overnight at 37°C. Colonies were collected in PBS and stored at -80°C for DNA extraction (Figure 1).

DNA library preparation for Illumina sequencing

Genomic DNA (gDNA) from the bacterial cell pellet of input library and output libraries stored at -80°C was extracted using QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA) following

149 manufacturer's protocol. The purity and concentration were checked using Qubit 2.0
150 Fluorometer (Life Technologies, Carlsbad, CA) with Qubit Assay Kits (dsDNA BR Assay)
151 following the manufacturer's manual.

152 The sample for Illumina sequencing was prepared as previously described (Dawoud et al., 2014;
153 Mandal and Kwon, 2017; Mandal et al., 2017; Mandal, 2016). All the DNA primers (Table S5)
154 used for Tn-seq library were custom designed using Primer3 (v. 0.4.0) (Untergasser et al., 2012)
155 and ordered from Integrated DNA Technologies (Coralville, Iowa). The simplified diagram for
156 preparation of Tn-seq amplicon library is shown Figure S1A. Briefly, Tn5-junctions at the right
157 end of transposon was enriched from gDNA extracted from input and output library. The single
158 primer linear extension was done with EZ-Tn5 primer3 using Taq DNA polymerase (New
159 England Biolabs, Ipswich, MA, USA). The 50 µl linear PCR extension reaction constituted:
160 Nuclease-free water – 40 µl (volume adjusted according to gDNA volume), Thermopol buffer
161 (10X) – 5 µl, dNTPs (2.5 mM each) – 1 µl, EZ-Tn5 primer3 (20 µM) – 1 µl, gDNA library (50
162 ng/ul) – 2 µl (~100 ng), and Taq DNA polymerase – 1 µl (added during PCR). The PCR cycle
163 consisted of manual hot start with the initial denaturation at 95°C for 2 min, and addition of Taq
164 DNA polymerase followed by 50 cycles of 95°C for 30 s, 62°C for 45 s, and 72°C for 10 s,
165 which was then followed by a hold at 4°C. The linear PCR products were then purified with
166 MinElute PCR purification kit (Qiagen, Valencia, CA, USA) and eluted in 10 µl of elution buffer
167 (EB) following the manufacturer's protocol. Then deoxycytosine homopolymer tail (C-tail) was
168 added to the linear extension purified PCR product using Terminal Transferase (TdT, New
169 England Biolabs, Ipswich, MA, USA) enzyme following previous protocol (Lazinski and
170 Camilli, 2013). The C-tailing reaction consisted: DNA (linear extension product from linear
171 PCR) – 10 µl, TdT Buffer (10X) – 2 µl, $CoCl_2$ (2.5 mM) – 2 µl, dCTP (10 mM) – 2.4 µl, ddCTP

(1mM) – 1 µl, Nuclease-free H_2O – 1.6 µl, and Terminal Transferase – 1 µl, making a total volume of 20 µl. The reaction was incubated at 37°C for 1 h followed by heat inactivation of the enzyme at 75°C for 20 min on a thermocycler. The C- tailed products were purified using MinElute PCR purification kit and eluted to 10 µl.

Subsequently, C-tailed PCR product was enriched with exponential PCR. PCR reaction constituted: nuclease-free H_2O – 35 µl, Thermopol Buffer (10X) – 5 µl, dNTPs (2.5 mM each) – 4 µl, IR2 BC primer (with Illumina adapter and barcode, 10 µM) – 2 µl, HTM primer (with Illumina adapter, 20 µM) – 1 µl, C-tailed DNA – 2 µl, and Taq DNA Polymerase (NEB) – 1 µl, making a total volume of 50 µl. The manual hot start PCR cycle comprised of 95°C for 2 min, followed by 25 cycles of 95°C for 30s, 58°C for 45s, and 72°C for 20s, trailed by a final extension at 72°C for 10 min.

Finally, the exponential PCR products were pulse heated at 65°C for 15 min and ran on 1.5% agarose gel. Tn-seq library had smear pattern whereas gDNA of *S.* Typhimurium (negative control) had almost no amplification (Figure S1B). Gel was excised ranging from 300-500 bp and DNA was extracted using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). The purity and concentration of DNA were measured using Qubit 2.0 Fluorometer. An equal amount (~ 10 ng) of DNA (gel-purified products) from each library were mixed together and sent for next generation sequencing, Illumina HiSeq 2000 single end read 100 cycles (Center for Genome Research and Biocomputing, Oregon State University, Corvallis).

Analysis of Transposon sequencing data

Raw reads from HiSeq Illumina sequencing were de-multiplexed based on the barcodes to their respective libraries using custom Perl script. The barcode and transposon sequence were trimmed

194 off from 5' end. Consequently, the remaining sequence was Tn5-junction sequences
195 with/without poly C-tail. Only 20 bp from the Tn5-junction were kept discarding most of the
196 poly C-tails. The reads were then aligned against *S.* Typhimurium 14028s complete genome
197 (NC_016856.1) using Bowtie version 0.12.7(Langmead and Salzberg, 2012). The aligned
198 sequence (SAM mapping file) were fed to ARTIST pipeline to identify conditionally essential
199 genes (CEGs) using Con-ARTIST (Pritchard et al., 2014). Briefly, Tn5 insertion frequency was
200 assigned to the *S.* Typhimurium 14028s genome divided into 100 bp window size. Uncorrected
201 raw data (non-normalized) of input and output libraries were used to normalize the control data
202 (IP1) to account for the random loss of mutants in output pool. Then, reads were compared
203 between the matching input and output pool using a Mann-Whitney U test (MWU). The MWU
204 results were used train hidden Markov model (HMM) to predict the likelihood of loci that were
205 not required for growth in either condition, essential under both conditions, enriched in output
206 library and window depleted in output library ($p < 0.01$). The insertions were only considered in
207 the central 80% of the gene to avoid any polar effect of transposon insertion. The cutoffs for
208 depleted loci and enriched loci were >8 fold and >2 fold, respectively.

209 **Comparative analysis of conditionally essential genes (CEGs) between *in vitro* and *in vivo***
210 **stressors**

211 We compared the *in vitro* essential genes identified in this study and our previous study
212 (Khatiwara et al. 2012) with the previously identified *in vivo* fitness genes. CEGs for acute
213 infection of mice (A-Mice), macrophage survival (MΦ) (Chan et al., 2005) and persistent
214 infection of mice (P-Mice) (Lawley et al., 2006) were previously identified in *S.* Typhimurium
215 strain SL1344 background. Additionally, *Salmonella* genes required for gastrointestinal

colonization of pig, calf and chicken were identified in *S.* Typhimurium strain ST4/74 (Chaudhuri et al., 2013), and those for intraperitoneal infection of mice (Sp-Liv) were reported in *S.* Typhimurium strain 14028s background (Silva-Valenzuela et al., 2015). The CEGs of different strain were searched for the corresponding orthologous genes in *S.* Typhimurium strain 14028s background using Prokaryotic Genome Analysis Tool (PGAT) (Brittnacher et al., 2011). To get insight into the phenotypic basis of CEGs required for *in vivo* intestinal colonization of pig, calf and chicken, these CEGs were compared with CEGs of *in vitro* host stressors found in the gut (PA, NaCl, PH3, Bile, and LB42). Similarly, for the phenotypic basis of CEGs for *in vivo* systemic infection (A-Mice, MΦ, P-Mice and Sp-Liv) were compared to *in vitro* macrophage stressors (H_2O_2, NaCl, Starvation, dLB, and PH3). Only the CEGs that were common between at least one of the *in vitro* host stressors and at least one of *in vivo* infection were identified and included for the comparative analysis.

Results and Discussion

Overall evaluation of resulting Tn-seq profiles

We have constructed a highly saturated transposon mutant library of *S.* Typhimurium 14028s with approximately 350,000 transposon mutants created via transformation of EZ-Tn5 transposome complex to electrocompetent cells. The complex Tn5 library, input pool 1 (IP1) was then subjected to negative selection under the *in vitro* stress conditions encountered during enteric and systemic infection as described in Materials and Methods. Input pool 2 (IP2) was the technical replicate of IP1 to evaluate the reproducibility of our Tn-seq method (Figure 1). Tn-seq amplicon library for Illumina sequencing was prepared for each of the input and output pools (Figure S1A and S1B). This efficient Tn-seq protocol was developed in our laboratory that offers

238 distinctive advantages over other Tn-seq library preparation methods, including a low amount
239 (~100 ng) of DNA required, and no need for physical shearing or restriction digestion(Dawoud
240 et al., 2014; Karash et al., 2017; Kwon et al., 2016; Mandal and Kwon, 2017).

241 Illumina sequencing using HiSeq 3000 produced 163,943,475 reads from a single flow cell lane.
242 The raw reads were demultiplexed allowing a perfect match for the barcodes used (Table S1)
243 with exception of up to two mismatches within Tn5 mosaic end (ME) using a custom Perl script.
244 H_2O_2 (19,250,956) had the highest number of reads followed by IP1 (10,842,764), Starvation
245 (9,518,226), IP2 (6,345,173), LB (5,004,934), PH3 (3,841,401), PA (2,113,033) and NaCl
246 (1,970,072) (Figure 2A).

247 After demultiplexing, Illumina reads were trimmed of barcode and transposon sequences. The
248 Tn5-junction sequences of 20bp were extracted and mapped to the complete genome of *S.*
249 Typhimurium 14028s (NC_016856.1) using Bowtie. The overall alignment rate throughout all
250 Tn5 libraries were 85.19% (SE ± 1.79). Additionally, we looked for the unique insertion sites in
251 the genome in each library. IP1 had the highest number of unique insertions (186,621) followed
252 by LB (157,915), H_2O_2 (149,752), IP2 (149,740), PA (127,722), NaCl (125,918), Starvation
253 (118,607) and PH3 (92,008) (Figure 2A). Similarly, H_2O_2 had the highest average read per
254 unique insertion site in the genome (96.007 ± 1.11) with 40 median reads, whereas NaCl had the
255 lowest (13.53 ± 0.99) with 5 median reads (Figure 2A).

256 Pre-aligned reads of the Tn5 library in default SAM mapping file format were fed to 'Analysis of
257 high-Resolution Transposon-Insertion Sequences Technique' (ARTIST) pipeline (Pritchard et
258 al., 2014). Tn5 insertions were mapped into 100 bp genome-wide windows. We observed the
259 highest Spearman correlation coefficients (a commonly used numerical measure to describe a
260 statistical relationship

p between two variables) between IP1 and IP2, and IP1 and LB (0.98, $p < 0.0001$). However, there was lower Spearman correlation of IP1 with NaCl (0.97, $p < 0.0001$), PA (0.96, $p < 0.0001$), and H_2O_2 (0.93, $p < 0.0001$). We observed the lowest correlation of IP1 with PH3 and Starvation (0.84 and 0.91 respectively, $p < 0.0001$) (Figure 2B). These relationships corroborate well with the Tn5 library selection strategies employed, with a higher correlation for the selections based on growth fitness (NaCl, PA, and H_2O_2) and a lower correlation for the selections based on survival (PH3 and Starvation).

Besides, we looked for the occurrence of any hot spots of Tn5 insertion in the sample libraries. We found an even distribution of Tn5 insertion reads across the libraries throughout the genome. Some of the genomic regions lacking insertions have white stripes that are clearly visible (Figure S2) across all the samples that represent essential loci in the *S.* Typhimurium 14028s genome.

Identification of Conditionally essential genes (CEGs)

In this study, we used two strategies to identify conditionally essential genes (CEGs) of *S.* Typhimurium to overcome host stressors. The first strategy was a negative selection of complex Tn5 mutant libraries based on growth fitness for mild stressors (3% NaCl, 100 mM propionate, 1 mM H_2O_2) and the second one was based on survival of Tn5 mutant libraries for harsher stressors (12 days starvation and PH3) as shown in Figure 1.

The ARTIST pipeline can identify if genes are entirely essential or domain essential in a given condition. In our study only a few of the genes were identified as domain essential and the majority of them were entirely essential. For simplicity, we assigned both categories of the genes entirely essential and domain essential into one category, conditionally essential genes (CEGs). We deliberately compared the each of the output pool PA, NaCl, and H_2O_2 with both IP1 and LB

283 separately. As expected, most of the CEGs were overlapped with these two comparisons. For the
284 conditions PA, NaCl, and H_2O_2, we considered the common set of identified CEGs via the
285 comparison of output library with both IP1 and LB as CEGs for each condition. However, the
286 output libraries for PH3, and Starvation were compared only with IP1 because the selection of
287 the Tn5 library was based on survived mutants and the mutant cells did not multiply during
288 selection in liquid media.

289 We identified an overlapping set of 339 CEGs that are required for fitness of *S.* Typhimurium
290 14028s in at least one of the five conditions (Figure 3A). Starvation had the highest CEGs (241),
291 followed by PH3 (103), NaCl (60), H_2O_2 (40) and PA (19) as shown in Table S2 and S3. This
292 might likely reflect that starvation is a severe stressor involving diverse genetic pathways for
293 survival, while PA is a mild stressor for the fitness of *S.* Typhimurium. More than a half of CEGs
294 were on the lagging strand (56.63%), which is somewhat contrary to the responsive genes in
295 *Escherichia coli* and *Streptococcus pneumoniae* (Nichols et al., 2011; van Opijnen and Camilli,
296 2012). We assigned a functional role to 96 CEGs that were putative proteins and 21 CEGs
297 belonging to hypothetical proteins. The stress tolerant proteins commonly identified in at least 2
298 of the *in vitro* stressors included ATP synthase, a transcriptional regulator, 3-dehydoroquinate
299 synthase, site-specific tyrosine recombinase *xerC*, flavin mononucleotide phosphatase, ribulose-
300 phosphate 3-epimerase, and DNA-dependent helicase II among others (Table S2 and S3).

301 Intriguingly, we found many genes in the *Salmonella* pathogenicity islands (SPI) were required
302 for fitness in the presence of the *in vitro* stressors used in this study. Numerous genes in SPI-1,
303 SPI-2, SPI-3, SPI-5, SPI-6, and SPI-11 were required for resistance against Starvation (n=68),
304 NaCl (n=28), and PH3 (n=27) (Table S4). However, no SPI genes were required for fitness in
305 PA and H_2O_2. SPI-5 and SPI-11 genes were only conditionally essential in PH3 (n=4 and 6,

306 respectively), while SPI-3 genes in NaCl (n=7) and SPI-6 genes in starvation (n=7). Tn-seq
307 profiles for SPI-1 region is shown in Figure S3A as an example.

308 For a broader insight into pathways involved in stress resistance, we assigned each CEGs to the
309 cluster of orthologous groups (COG) using eggNOG database (evolutionary genealogy of genes:
310 Non-supervised Orthologous Groups) (Jensen et al., 2008). The CEGs having top hit for the
311 COG in the *S.* Typhimurium LT2 were kept and CEGs with no orthologous group were allotted
312 to group XX (Figure 3B; Table S3). In overall, 21.83% of CEGs belonged to category "function
313 unknown" followed by "intracellular trafficking, secretion, and vesicular transport" (10.91%),
314 "energy production and conversion" (9.44%), and "no orthologs found" (8.26%) among others.
315 A substantial portion of CEGs (30.6%) falling into either "function unknown or "no orthologs
316 found" shows that our data set is rich in novel genotype-phenotype relationships.

317 Additionally, we were interested to see if any CEGs identified in our study fell into the essential
318 genomes of *S.* Typhimurium in other strain backgrounds. Essential genomes of *S.* Typhimurium
319 strain SL3261 (selected on LB agar) (Barquist et al., 2013) and *S.* Typhimurium strain LT2
320 (selected on rich medium) (Knuth et al., 2004; Zhang et al., 2004) were compared with the CEGs
321 of *S.* Typhimurium 14028s identified in this study. Genes in different strain background were
322 looked for the corresponding orthologous genes in *S.* Typhimurium 14028s background.
323 Interestingly, 10 and 15 CEGs in this study were shared with the essential genes of *S.*
324 Typhimurium SL3261 and LT2, respectively (Table S5; Figure S4). This indicates that these
325 genes that are essential in other strain backgrounds are dispensable in *S.* Typhimurium 14028s
326 strain background.

327 **Molecular and phenotypic basis of CEGs in *S.* Typhimurium**

Next, we delved into the genetic and biochemical mechanisms related to the CEGs identified in our study. For convenience, we split the section into specific CEGs, required for fitness in only one stressor, and common CEGs shared in at least two stressors out of five host stressors.

CEGs specifically required for propionate (100 mM PA) stress resistance. CEGs specific for fitness of *S.* Typhimurium in propionate were *yiiD* and *sdhAD*. YiiD is a putative acetyltransferase protein (Read coverage shown in Figure 4C). Acetylation, a post-translation modification of protein was previously shown to enable prokaryotes to increase stress resistance (Ma and Wood, 2011). Additionally, succinate dehydrogenase flavoprotein (*sdhA*) and cytochrome b566 (*sdhD*) subunit proteins were up-regulated by intestinal SCFA in *S.* Typhimurium (Lawhon, 2002). Chowdhury and Shimizu (2008) reported that *sdhA* in the tricarboxylic acid cycle (TCA) were highly induced during temperature upshift in *E. coli* (Hasan and Shimizu, 2008).

CEGs specifically required for osmotic (3% NaCl) stress resistance. Twenty-six resistance genes of *S.* Typhimurium were required for fitness in osmotic stress (3% NaCl) alone. Protein-protein network analysis using STRING database (http://string-db.org) against *S. enterica* LT2 showed three distinct clustering of genes, SPI-3 (*mgtBC*, *misL*, *cigR*, *slsA*, *fidL* and *marT*), two-component system (*dcuBRS*) and sodium ion transport (*yihPO*) along with other nodes (http://bit.ly/2bCKGVG). SPI-3 genes are important for intracellular replication inside phagosome where *Salmonella* experience hyperosmotic stress (Schmidt and Hensel, 2004). The virulence proteins *mgtC* and *mgtB*, Mg^{2+} transporter were expressed five-fold when *S.* Typhimurium was exposed to 0.3 M NaCl (Lee and Groisman, 2012). MisL, an autotransporter protein is an intestinal colonization factor (activated by *marT*, a transcriptional regulator) that binds to extracellular matrix fibronectin in an animal host and is also involved in adhesion to

plant tissue (Dorsey et al., 2005; Kroupitski et al., 2013). Deletion of *cigR* in *S.* Pullorum resulted in a significantly decreased biofilm formation and increased virulence (Yin et al., 2016). Additionally, Figureueira et al., showed Δ*cigR* strain of *S.* Typhimurium had attenuated replication in mouse bone marrow-derived macrophage (Figueira et al., 2013).

yihPO genes are essential for capsule assembly that is required by *Salmonella* for environmental stress persistence such as desiccation (Gibson et al., 2006). The absence of *ompL* (ortholog of *yshA*) leads to solvent hypersensitivity as it helps in the stabilization of cell wall integrity protecting from solvent penetrance as a physical barrier (Murinova and Dercova, 2014). In *E. coli*, the genes under the control of *dcuS-dcuR*, a two-component system, were not affected upon a hyperosmotic shock (Weber and Jung, 2002). However, *dcuBRS* were conditionally essential in *S.* Typhimurium for fitness during osmotic stress. Putative cytoplasmic protein (STM14_4542, STM14_4828, and STM14_5175), putative inner membrane protein (STM14_4824 and STM14_5184) and putative hydrolase (STM14_4823) were also required for osmotic stress tolerance.

CEGs specifically required for oxidative (1 mM H_2O_2) stress resistance. We identified 16 specific resistance genes required for fitness of *S.* Typhimurium in the presence of 1 mM H_2O_2 and the functional protein association network analysis among the genes was constructed using STRING against *S. enterica* LT2 (http://bit.ly/2bsVKXF). Major resistance genes were those involved in two-component system (*glnD*, *rpoN*, *arcA* (STM4598), and *arcB* (STM3328)), DNA recombination (*recJ*, and *xerD*), and metal ion transport (*corA*, and *trkA*).

Hydrogen peroxide kills *E. coli* cells with two distinct modes, mode-1 killing occurs at a lower concentration of H_2O_2 due to DNA damage and mode-2 killing occurs at a higher concentration of H_2O_2 due to damage of other structures like proteins and lipids(Imlay and Linn, 1986).

374 Nucleic acid metabolic process genes involved in oxidative stress resistance were *recJ*, *xerD*,
375 *sun,* and *rpoN*. *RecJ* protein, a single-stranded DNA (ssDNA)-specific 5'-3'
376 exonuclease/deoxribophophodiesterase, plays a role in homologous recombination, mismatch
377 repair, and base excision repair (Wakamatsu et al., 2011). In *E. coli*, *xerD* knockout mutants are
378 hypersensitive to tightly bound DNA-protein complexes (TBCs) that block replication forks *in*
379 *vivo* (Henderson and Kreuzer, 2015). *RpoN*, the alternative sigma factor 54 (σ^{54}), an important
380 regulator of stress resistance and virulence genes in many bacterial species (Riordan et al., 2010).
381 σ^{54} is involved in carbon/nitrogen limitation, nucleic acid damage, cell envelope, and nitric oxide
382 stress (Hartman et al., 2016). However, Hwang *et al.,* 2011 found that *rpoN* mutant in
383 *Campylobacter jejuni* was more resistant to 1 mM H_2O_2 (Hwang et al., 2011).

384 Besides, cellular component genes crucial for fitness in H_2O_2 stress were *dsbC*, *glmS*, *trkA*, *corA*
385 including *sun* and *xerD*. *DsbC*, a protein essential for disulfide bond isomerization in the
386 periplasm, has a new role in *E. coli* in protection against oxidative stress (Denoncin et al., 2014).
387 In *E. coli GlmS* plays an important role in cell wall synthesis thus providing protection against
388 cell envelope stress response (Zhou et al., 2009). HscB, a chaperone-encoding gene is
389 upregulated after exposure to oxidative stress in *Burkholderia pseudomallei* (Jitprasutwit et al.,
390 2014). YbgF, an outer membrane vesicle protein, increases the survival of bacteria during
391 exposure to stress or from toxic unfolded proteins by releasing the unwanted periplasmic
392 component (Gogol et al., 2011).

393 ***CEGs specifically required for higher acidic (pH 3) stress resistance.*** We found 49 specific
394 stress resistance genes required only for survival of *S*. Typhimurium in extreme acidic condition
395 (pH 3) among other stressors. Formate dehydrogenase (*fdoHI*, and *fdhDE*) curli proteins
396 (*csgBDEFG*), virulence and envelope proteins (SPI-2: *orf245*, *orf408*, *ssaB*; SPI-5: *pipBC*, *sopB*,

and SPI-11: *envEF*, *pagCD*, *msgA*, STM14_1486 where *ssaB*, *pipB,* and *sopB* are effector proteins), and biopolymer transport protein (*exbD* and *exbB*) were clustered in functional protein association network analysis using STRING (http://bit.ly/2bCLVnL).

Formate dehydrogenase catalyzes the oxidation of formate (HCOO-) to CO_2 and H^+. The released electrons from this reaction are used by two cytoplasmic protons to form dihydrogen thus consuming net protons, consequently, counteracting acidification (Leonhartsberger et al., 2002). Curli are major complex extra-cellular proteinaceous matrix produced by *Enterobacteriaceae* that helps pathogenic bacteria like *Salmonella* in adhesion to surfaces, cell aggregation, and biofilm formation (Barnhart and Chapman, 2006). Acidic pH strongly enhances biofilm formation in *Streptococcus agalactiae* (D'Urzo et al., 2014). We hypothesize that curli fibers might potentially protect bacteria from severe acid stress through the physical barrier and likely by the generation of alkaline compounds as in oral biofilms (Cotter and Hill, 2003). PhoP regulates SPI-11 genes such as *envEF*, *pagCD*, and *msgA* where later three are required by *Salmonella* to survive low pH within macrophage (Gunn et al., 1995; Lee et al., 2013). In *Helicobacter pylori,* only the organism to colonize in the acidic human stomach, *ExbB*/*ExbD*/*TonB* complex is required for acid survival and periplasmic buffering (Marcus et al., 2013). Additionally, survival of *ΔexbD* was diminished compared to wild type at pH 3 in *E. coli* (Ahmer et al., 1995). The *metC* gene encoding a key enzyme in methionine biosynthesis, required for the generation of homocysteine, pyruvate, and ammonia, play a crucial role in bacterial acid stress responses (Reid et al., 2008).

CEGs specifically required for starvation stress resistance. Out of 261 *Salmonella* fitness genes essential for starvation stress, 160 genes were explicitly important for resistance against starvation stress among the five infection-relevant conditions in this study

(http://mcaf.ee/k0uhrm). Major enriched gene pathways were oxidative phosphorylation, pathogenesis, two-component system, and lipopolysaccharide biosynthetic process among others. NADH dehydrogenase, the first component of the respiratory chain, subunit proteins (*nuoCEFGHLMN*) were required for fitness of *Salmonella* during long-term carbon starvation. *Salmonella* defective in NADH dehydrogenase enzyme exhibits defective energy-dependent proteolysis during carbon starvation (Archer et al., 1993). Proteolysis of unbound or unemployed proteins helps bacteria to access nutrients as an important survival strategy during carbon starvation (Michalik et al., 2009). SPI-1 (*hilACD*, *iagB*, *invH*, *orgAC*, *prgHIJK*, STM14_3500, and STM14_3501) and SPI-2 (*ssaMNOPQRSTV*, *sscB*, and *sseDEF*) encoding type III secretion system (T3SS) and SPI-6 (*safABCD*, *sinR*, STM14_0359, and *ybeJ*) encoding type VI (T6SS) secretion system were required for *in vitro* survival in long-term starvation stress. *Salmonella* usually requires SPI-1 genes for the invasion of intestinal epithelial cells (Klein et al., 2000). HilACD regulates SPI-1 invasion gene expression during multiple environmental conditions including stationary phase, pH, osmolality, oxygen tension, and short chain fatty acids (Olekhnovich and Kadner, 2007). SPI-2 genes are expressed under *in vitro* starvation conditions indicating the use of nutritional deprivation as a signal (Hensel, 2000). T6SS has been hypothesized to confer a growth advantage to bacteria in environmental niches where bacterial competition for nutrient is critical for survival (Brunet et al., 2015).

Two-component systems (TCs), a basic stimulus-response coupling mechanism, enable microbes to respond to various stimuli such as pH, osmolality, quorum signals, or nutrient availability and regulate their cellular functions (Freeman et al., 2013). TCs required for fitness during starvation conditions were *envZ*/*OmpR*, *cpxA*/*cpxR*, sensory histidine kinase protein (*phoQ*), and kdpD (Figure S3B). EnvZ/*OmpR* regulates the synthesis of porin proteins (*ompF* and *OmpC*) that are

443 important for the survival of *E. coli* in sea water under starvation stress condition (Darcan et al.,
444 2009). It is believed that carbon starvation causes cell envelope stress. Bacchelor et al., (2005)
445 found *cpxA*/*cpxR* in *E. coli* regulates the expression of prions *ompF* and *ompC*, a major
446 component of the outer membrane. However, Kenyon et al., (2002) showed the starvation stress
447 of *S.* Typhimurium do not require cpxR-regulated extra-cytoplasmic functions (Batchelor et al.,
448 2005; Kenyon et al., 2002). PhoQ and *kdpD* plays a role in Mg^{2+} and K^{+} homeostasis
449 respectively, critical to the virulence and intracellular survival of *S.* Typhimurium (Freeman et
450 al., 2013; Kato and Groisman, 2008).

451 The outer membrane of Gram-negative bacteria contains phospholipids and lipopolysaccharides
452 (LPS). LPS molecules act as a permeability barrier to prevent the entry of toxic compounds and
453 allow the entry of nutrient molecules (Schakermann et al., 2013). LPS biosynthetic process genes
454 required for fitness in starvation conditions were *rfbABCD*, *rfbUNMKP, galF*, *udg*, *wzxE,* and
455 *wzzB*. Starvation of carbon energy source activates envelope stress response in *S.* Typhimurium
456 (Rowley et al., 2006). Additionally, *pstSCAB* coding for the Pst ABC transporter catalyzes the
457 uptake of inorganic phosphate (Lüttmann et al., 2012). Mutations in the Pst system results in
458 structural modifications of lipid A and an imbalance in unsaturated fatty acids consequently
459 leading to increase in outer membrane permeability making *E. coli* more vulnerable to
460 environmental stresses including antimicrobial peptide and low pH (Lüttmann et al., 2012).

461 Additional genes required for starvation stress resistance were *aroGH*, ytfMNP (ytfM - outer
462 membrane protein), *stcB* (putative periplasmic outer chaperone protein). Furthermore, other
463 envelope proteins were outer membrane lipoproteins (*stcD* and *yifL*), putative outer membrane
464 proteins (*stcC*, STM14_0404, and ytfM), and putative inner membrane proteins (STM14_0398,
465 STM14_0402, STM14_2763, STM14_4741, STM14_4742, STM14_4745, STM14_4880, *ydiK*

and *yjeT*). Similarly, putative cytoplasmic proteins required for starvation stress were STM14_2759, STM14_4743, STM14_5374, *ydiL,* and *ytfP*.

CEGs required for tolerance to multiple stressors. We found 12 *Salmonella* genes required for stress resistance in either three or four of the *in vitro* host stresses in our study as shown in STRING protein-protein interaction network (http://bit.ly/2btx1zg). The enriched GO biological process / KEGG pathways were ncRNA processing (*gidAB* and *mnmE*), DNA metabolic process (*dam*, *uvrD* (SOS response), *xerC*), and biosynthesis of amino acids (*aroB* and *rpe* - microbial metabolism in diverse environments). In addition, other responsive proteins include ATP synthase subunit protein (*atpI*), putative permease (STM14_4659), inner membrane protein (*damX*), and flavin mononucleotide phosphatase.

DamX, *dam*, *rpe*, *aroB*, *uvrD*, and *yigB* were required for fitness in PH3, Starvation, and H_2O_2. Disruption of *damX* in *S. enterica* causes bile sensitivity (López-Garrido and Casadesús, 2010). DNA adenine methylation gene (*dam*) plays an important role in bacterial gene expression and virulence (Low et al., 2001). Dam mutants of *S. enterica* are extremely attenuated in mouse (Jakomin et al., 2008). The gene *aroB* encodes dehyroquinate synthase, a part of shikimate pathway, is essential for bacteria and absent in mammals (de Mendonca et al., 2007). In prokaryote species, *uvrD* is involved in maintaining genomic stability and helps DNA lesion repair, mismatch repair, nucleotide excision repair and recombinational repair (Kang and Blaser, 2006). Overproduction of *yigB* produced higher-level persister, cells that exhibit multidrug tolerance, in *E. coli* (Hansen et al., 2008). However, deletion of *gidB* (glucose-inhibited division gene B) confers high-level antimicrobial resistance in *Salmonella* and has compromised overall bacterial fitness compared to wildtype (Mikheil et al., 2012). GidA (together with *mnmE*) is

488 responsible for the proper biosynthesis of 5-methylaminomethtyl-2-thouridine of tRNAs and
489 deletion causes attenuation in bacterial pathogenesis (Shippy and Fadl, 2014b).

490 **ATP synthase genes are obligatory for *Salmonella* fitness during *in vitro* host stressors**

491 ATP synthase (F_1F_0-ATPase) is a ubiquitous enzyme largely conserved across all domains of
492 life. All the eight genes encoding ATP synthase subunit proteins were required for fitness of *S.*
493 Typhimurium in every 5 *in vitro* conditions of our study (Figure 3A and 4A). F_1F_0-ATP synthase
494 complex is required for ATP production from ADP and Pi. ATP synthase also regulates pH
495 homeostasis in bacteria (*Listeria monocytogenes* and *S.* Typhimurium) at the expense of ATP
496 (Balemans et al., 2012). In *Streptococcal faecalis*, upregulation of F_1F_0-ATPase promotes ATP-
497 dependent H+ extrusion under acidic conditions. However, in *E. coli* the expression of ATP
498 synthase is decreased under acidic condition (Krulwich et al., 2011). ATP synthase in
499 *Mycobacterium* and *Staphylococcus* has been validated as a promising target for new
500 antimicrobial drugs (Balemans et al., 2012; Lu et al., 2014).

501 **Mechanistic basis of *Salmonella in vivo* fitness genes required for enteric and systemic**
502 **infection**

503 The network diagrams shown in Figure 5 and Figure 6 show all the genes that are commonly
504 important for fitness under at least one of the *in vitro* and *in vivo* conditions. The genes that were
505 important only either in the *in vitro* or *in vivo* conditions were excluded in the diagram.
506 Numerous *in vivo* fitness genes have been identified in previous studies, indicating that they are
507 required by *S.* Typhimurium to overcome host defenses. However, for a large portion of them the
508 mechanistic bases why they are required in particular *in vivo* niches remain unknown. The
509 information on the common requirements of the genes shown in these networks (Figure 5 and 6)

for both at least one well-defined *in vitro* stress and *in vivo* infection model is valuable in the sense that it provides novel insights on the type of selective pressures *S.* Typhimurium might be facing during infection in the host.

Enteric infection. We have identified an overlapping set of 135 CEGs that are commonly required to cause enteric infection in at least one of the host [pig, calf, and chicken (Chaudhuri et al., 2013)] and for fitness in one of the *in vitro* host stressors [LB42, Bile (Khatiwara et al., 2012), PH3, PA, and NaCl] encountered during enteric infection (Figure 5; Table S6). Genes in SPI-1 (*invABCEIJ*, *sicAP*, *sipABCD*, *spaOPQRS*, *sptP*) and SPI-3 (*cigR*, *marT*, *mgtBC*, *misL*, *slsA*) were required for fitness in NaCl and all host. However, genes encoding SPI-2 (*sseCG*), SPI-5(*slsA*, *pipC*) and SPI-11(*envEF*) were essential for fitness only one *in vitro* stressor PH3 and intestinal colonization in 3 hosts. Other enriched pathways were lipopolysaccharide biosynthesis (*rfaIJKLQY* and *rfbBDKMNP*), oxidative phosphorylation (ATP synthase genes and *sdhA*), and biosynthesis of amino acids (*aroABD, rpe and metC*) including others as shown in STRIN protein-protein interaction against *S. enterica* LT2 (http://mcaf.ee/wzljud).

High osmolality, low oxygen, and late log phase induce *hilA* expression *in vitro* that in turn regulates the expression of SPI-1 genes (Lostroh and Lee, 2001). Interestingly, we identified SPI-1 genes as fitness genes required for *in vitro* NaCl stressor. Similarly, lipopolysaccharide (LPS) biosynthetic process genes were enriched in LB42, Bile and in pig, calf, and chicken for fitness during enteric infection. LPS, a critical factor in the virulence of gram-negative bacterial infection is required for intestinal colonization, resistance to killing by macrophage, swarming motility, serum resistance and bile stress (Khatiwara et al., 2012; Kong et al., 2011). CsgBA (curli subunit protein) mutant of *S.* Typhimurium was attenuated to elicit fluid accumulation in bovine ligated ileal loops (Tükel et al., 2005) and are required for fitness in PH3 including *csgF*

and *csgG*. Additionally, putative proteins STM14_1138, STM14_1486, STM14_1981, STM14_3333 and STM14_4826, STM14_4828, STM14_5184, STM14_5185 (hypothetical protein) were required for fitness *in vitro* acidic and osmotic stress respectively and enteric infection in the entire three host.

Systemic infection. We compared the CEGs that are at least shared between the one of the *in vitro* host stressors, H_2O_2, NaCl, PH3, Starvation and dLB (Khatiwara et al., 2012), encountered inside MΦ and *in vivo* systemic infections (MΦ (Chan et al., 2005), A-Mice (Chan et al., 2005), P-Mice (Lawley et al., 2006), Sp-Liv (Silva-Valenzuela et al., 2015) and identified an overlapping set of 130 genes (Figure 6; Table S7) shown in protein-protein interaction network using STRING (http://mcaf.ee/p34rjn). SPI-1 genes (*hilACD*, *iacP*, *iagB*, *invABCEFGI*, *orgA*, *prgHIJK*, *sicA*, *sipABC*, *spaOPQRS*) encoding TTSS were essential for fitness in NaCl, Starvation, MΦ survival and systemic infection. Additionally, SPI-2 genes (*ssaBCDEGIJKLMNOPQRSTV*, *orf245*, *orf408*, *sscAB*, *sseCDEF*, *ssrA*, STM14_1706) encoding TTSS were required for fitness in PH3, starvation, MΦ survival and systemic infection. Similarly, SPI-3 genes (*marBCT*) were required for fitness in NaCl, MΦ survival, and persistent infection in mice (P-Mice). SPI-11 genes (envF, pagCD) were required for fitness in PH3, MΦ survival, and P-Mice.

Other than SPI genes, the majorly enriched genes were nucleic acid metabolic process (*dam*, *trpS*, *MnmE*, *truA*, *serc*, *csgD*, *ompR* and *cra*), lipopolysaccharide biosynthetic process (*rfbABCNPU*, *rfaB*, *udg*, *galF*), oxidative phosphorylation (ATP synthase genes, NADH dehydrogenase genes), two component system (*ompR*, *barA*, *phoQ*, *glnDL*, *pagKO*) among others (Figure 6). Gene *dam* was required for fitness in H_2O_2, NaCl, A-Mice, and Sp-Liv. XerC and *rpe* were required for H_2O_2, PH3, Starvation and Sp-Liv. Interestingly, *pagK* were not

identified as CEG in A-Mice, P-Mice, Sp-Liv but in PH3, Starvation, and MΦ. Putative genes either essential for one of *in vitro* or *in vivo* systemic infection were STM14_1138, STM14_4880, STM14_4992, STM14_5184, STM14_2759, STM14_2807, STM14_3334, STM14_4825, STM14_5299, and STM14_5300.

Limitations of the study

This study has some limitations. Firstly, this study was exploratory in nature. Thus, prior knowledge of CEGs regarding stress tolerance were discussed where possible rather than performing phenotypic study of single gene knockout mutants. Secondly, Tn-seq approach is prone to false positive and false negative results. However, assessment for either false positive or false negative was not performed. Additionally, domain essential genes might have increased the chance of false positive discovery which were categorized as CEGs. Lastly, the scope of comparative study was limited to the CEGs identified in this study that were compared with the previously identified CEGs either *in vitro* or *in vivo* conditions which were mainly identified using Tn-seq approach. Nevertheless, all the CEGs identified in stress conditions had significantly lower reads compared to the control group, strongly supporting true conditional essentiality of the CEGs identified in this study. Most importantly, our goal was to provide comprehensive framework for mechanistic basis of genes required for *in vivo* fitness.

Conclusion

A recent study by Kroger et. al. (2013) presented transcriptomes of *S.* Typhimurium in 22 distinct infection-relevant environmental conditions *in vitro*. The study found induction of *Salmonella* pathogenicity islands *in vitro* conditions such as early stationary phase, anaerobic growth, oxygen shock, nitric oxide shock as well as in pH3, NaCl, bile, and peroxide shock

among others (Kröger et al., 2013). However, transcription of a gene does not necessarily indicate the need of that gene function for fitness in a given particular condition. The transcript can be a leaky expression or required for fitness in the upcoming environment in a cost effective way through predictive adaptation, phenomena where bacteria are able to anticipate and pre-emptively respond to the regular environmental fluctuations (temporally distributed stimuli) that confers a considerable fitness advantage for the survival of an organism (Mitchell et al., 2009; Ta gkopoulos et al., 2008). Traditionally, it is believed that "central dogma of life" i.e. flow of information from DNA to RNA to proteins are highly concordant. However, there is a modest correlation between levels of transcripts and corresponding proteins (Foss et al., 2007; Fu et al., 2009; Ghazalpour et al., 2011). Thus, functional genomics screening such as Tn-seq is expected to reveal more direct functional aspects of the genes involved in responding to the current stresses.

In this report, we were able to map genotype to phenotype links providing the mechanistic basis of the genetic requirements for fitness for an overlapping set of 221 virulence genes for *in vivo* fitness (Figure S5). These CEGs were required for fitness in at least one of the *in vitro* host stressors (PA, NaCl, PH3, Starvation, Bile, LB42 and dLB), and enteric infection (calf, chicken and pig), or systemic infection (mice including intracellular survival inside macrophage). Forty-four common CEGs were required to cause both systemic and enteric infections (*in vivo* fitness) and *in vitro* fitness (Figure S5 and Table 1). Common SPI genes for *in vivo* and *in vitro* fitness were SPI-1 (*invABCEI*, *sicA*, *sipABD*, *spaOPQRS*), SPI-2(*sseC*), SPI-3(*marT*, *mgtCB*) and SPI-11(*envF*). *Salmonella* genes other than SPI essential for fitness under *in vitro* stresses and *in vivo* survival were *atpAEF*, *lepA*, *dam*, *pstB*, *xerC*, *manA*, *phoQ*, *rfaQ*, *rfbBIP*, *rpe*, *trmE*, *rfbIP*,

ompR, *csgF*, *recG*, *hscA*, *barA*, and putative genes STM14_1138, STM14_3334, STM14_4825, and STM14_5184 (Table 1).

Interestingly, most of the common forty-four genes required for *in vitro* and *in vivo* (enteric and systemic infection) fitness have been implicated in vaccine or drug target development against broad spectrum of bacteria. Such as ATP synthase genes (Balemans et al., 2012; Lu et al., 2014), *dam* (Garcia-Del Portillo et al., 1999), *pstB* (Garmory and Titball, 2004), *phoQ* (Miller and Mekalanos, 1998), *ompR* (Dougan et al., 1996), *xerC* (Hur et al., 2011), and *rfbBPN* (Sturm and Timmis, 1986), *manA* (Amineni et al., 2010), *rpe* (Edwards et al., 2004), *lepA* (Patton, 2007), *csgF* (Cegelski et al., 2008), trmE (Shippy and Fadl, 2014a), and SPI-1 and SPI-2 (Matulová et al., 2012) have been used as vaccine development or drug target (Table 1). Thus, there lies a great potential to explore genes *marBCT*, *envF*, *barA*, *hscA*, *rfaQ*, *rfbI* and putative proteins STM14_1138, STM14_3334, STM14_4825, and STM_5184 as novel therapeutic and intervention strategy to curb *Salmonella* infection.

Acknowledgements

We would like to thank Arkansas High Performance Computing Center (AHPCC) - University of Arkansas for their computational support.

616 **Table 1: *Salmonella* genes required for *in vitro* and *in vivo* (enteric and systemic) fitness.**

Category	Genes	Conditions (*In vitro*, enteric, and, systemic)	COG	Protein name
SPI Genes				
SPI-1*	*invA*	Na, S, C, P, Ch, MΦ, SL	U	needle complex export protein
	invB	Na, S, C, P, Ch, MΦ, SL	U	secretion chaperone
	invC	Na, S, C, P, MΦ, PM	NU	ATP synthase SpaL
	invE	Na, S, C, P, Ch, MΦ, SL	U	invasion protein
	invI	Na, S, C, P, Ch, MΦ, SL	S	needle complex assembly protein
	sicA	Na, S, C, P, Ch, MΦ, SL	S	secretion chaperone
	sipA	Na, S, C, P, Ch, MΦ	D	secreted effector protein
	sipB	Na, S, C, P, Ch, MΦ, SL	U	translocation machinery component
	sipD	Na,S, C, P, Ch, MΦ	S	translocation machinery component
	spaO	Na, S, C, P, Ch, MΦ, PM, SL	U	surface presentation of antigens protein SpaO
	spaP	Na, S, C, P, Ch, MΦ, AM, SL	U	surface presentation of antigens protein SpaP
	spaQ	Na, S, C, P, Ch, SL	U	needle complex export protein
	spaR	Na, S, C, P, Ch, MΦ, PM, SL	U	needle complex export protein
	spaS	Na, S, C, P, Ch, MΦ	U	surface presentation of antigens protein SpaS
SPI-2*	*sseC*	pH, S, C, P, Ch, MΦ, SL	S	translocation machinery component
PI-3	*marT*	Na, C, P, Ch, PM	K	putative transcriptional regulator
SPI-11	*envF*	pH, C, Ch, MΦ	XX	putative envelope lipoprotein
Non-SPI genes				
Two-component system	*ompR**	B, S, C, P, MΦ, SL	T	osmolality response regulator
	*phoQ**	B, S, dLB, C, Ch, AM, SL	E	sensor protein PhoQ
	barA	pH, S, C, P, Ch, MΦ	T	hybrid sensory histidine kinase BarA
O antigen biosynthetic process	*rfbB*	B, S, C, P, Ch, AM, SL	M	dTDP-glucose-4,6-dehydratase
	rfbP	B, S, L4, C, P, Ch, PM	M	undecaprenol-phosphate galactosephosphotransferase/O-antigen transferase
	rfbN	B, S, C, P, Ch, SL	M	rhamnosyl transferase
ATP synthase genes*	*atpA*	PA, Na, pH, H2, S, C, P, Ch, SL	C	F0F1 ATP synthase subunit alpha
	atpE	PA, Na, pH, H2, S, P, Ch, MΦ, SL	C	F0F1 ATP synthase subunit C

617

Table 1: Continued…

Category	Genes	Conditions (*in vitro*, enteric, and, systemic)	COG	Protein name
	atpF	PA, Na, pH, H2, S, C, Ch, MΦ	C	F0F1 ATP synthase subunit B
Mismatch repair	*dam**	pH, H2, S, C, P, Ch, AM, SL	L	DNA adenine methylase
chromosome segregation	*xerC**	PA, pH, H2, S, C, P, Ch, SL	L	site-specific tyrosine recombinase XerC
Fructose and Mannose Metabolism	*manA**	B, L4, dLB, C, P, Ch, SL	G	mannose-6-phosphate isomerase
Carbon metabolism	*rpe**	pH, H2, S, C, P, Ch, SL	G	ribulose-phosphate 3-epimerase
Homologous recombination	*recG*	B, dLB, C, P, Ch, AM	L	ATP-dependent DNA helicase RecG
ABC transporter	*pstB**	B, L4, S, dLB, C, P, Ch, SL	P	phosphate transporter subunit
Translational elongation	*lepA**	B, L4, S, dLB, C, P, Ch, MΦ, SL	M	GTP-binding protein LepA
Iron-sulfur cluster assembly	*hscA*	pH, H2, C, Ch, MΦ, SL	F	chaperone protein HscA
Others	*csgF**	pH, C, P, Ch, MΦ, PM	U	curli assembly protein CsgF
	rfaQ	B, L4, dLB, C, P, Ch, SL	M	lipopolysaccharide core biosynthesis protein
	rfbI	B, S, dLB, C, P, Ch, MΦ	GM	CDP-6-deoxy-delta-3,4-glucoseen reductase
	*trmE**	PA, H2, S, C, P, Ch, SL	S	tRNA modification GTPase TrmE
Putative Protein	STM14_1138	pH, C, P, Ch, MΦ	E	putative transcriptional regulator
	STM14_3334	pH, C, P, SL	S	putative DNA/RNA helicase
	STM14_4825	Na, C, P, Ch, MΦ	H	coproporphyrinogen III oxidase
	STM_5184	Na, C, P, Ch, SL	S	putative inner membrane protein

Genes marked with asterisk (*) have been implicated in vaccine development or drug target against a wide range of bacteria. The genes listed are required for both *in vitro* and *in vivo* fitness (enteric and systemic infection) i.e. conditions listed in Figure 5 and Figure 6. COG- Cluster of orthologous groups (same as Figure 3), SPI- *Salmonella* pathogenicity island, Na- NaCl, S-Starvation, C – Cattle, P- Pig, Ch- Chicken, MΦ- Macrophage, SL- Sp-Liv, pH- PH3, B- Bile, L4- LB42, AM- A-Mice, PM- P-Mice, SL- Sp-Liv.

REFERENCES

Ahmer, B.M., Thomas, M.G., Larsen, R.A., and Postle, K. (1995). Characterization of the exbBD operon of escherichia coli and the role of ExbB and ExbD in TonB function and stability. *J. Bacteriol.* 177:16, 4742-4747.

Amineni, U., Pradhan, D., and Marisetty, H. (2010). In silico identification of common putative drug targets in leptospira interrogans. *Journal of Chemical Biology.* 3:4, 165-173.

Ao, T.T., Feasey, N.A., Gordon, M.A., Keddy, K.H., Angulo, F.J., and Crump, J.A. (2015). Global burden of invasive nontyphoidal salmonella disease, 2010(1). *Emerg. Infect. Dis.* 21:6, 10.3201/eid2106.140999. doi:10.3201/eid2106.140999.

Archer, C.D., Wang, X., and Elliott, T. (1993). Mutants defective in the energy-conserving NADH dehydrogenase of salmonella typhimurium identified by a decrease in energy-dependent proteolysis after carbon starvation. *Proc. Natl. Acad. Sci. U. S. A.* 90:21, 9877-9881.

Balemans, W., Vranckx, L., Lounis, N., Pop, O., Guillemont, J., Vergauwen, K., Mol, S., Gilissen, R., Motte, M., Lancois, D*., et al.* (2012). Novel antibiotics targeting respiratory ATP synthesis in gram-positive pathogenic bacteria. *Antimicrob. Agents Chemother.* 56:8, 4131-4139. doi:10.1128/AAC.00273-12.

Barnhart, M.M., and Chapman, M.R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol.* 60, 131-147. doi:10.1146/annurev.micro.60.080805.142106.

Barquist, L., Langridge, G.C., Turner, D.J., Phan, M.D., Turner, A.K., Bateman, A., Parkhill, J., Wain, J., and Gardner, P.P. (2013). A comparison of dense transposon insertion libraries in the salmonella serovars typhi and typhimurium. *Nucleic Acids Res.* 41:8, 4549-4564. doi:10.1093/nar/gkt148.

Batchelor, E., Walthers, D., Kenney, L.J., and Goulian, M. (2005). The escherichia coli CpxA-CpxR envelope stress response system regulates expression of the porins ompF and ompC. *J. Bacteriol.* 187:16, 5723-5731. doi:187/16/5723.

Brittnacher, M.J., Fong, C., Hayden, H.S., Jacobs, M.A., Radey, M., and Rohmer, L. (2011). PGAT: A multistrain analysis resource for microbial genomes. *Bioinformatics.* 27:17, 2429-2430. doi:10.1093/bioinformatics/btr418.

Brunet, Y.R., Khodr, A., Logger, L., Aussel, L., Mignot, T., Rimsky, S., and Cascales, E. (2015). H-NS silencing of the salmonella pathogenicity island 6-encoded type VI secretion system limits salmonella enterica serovar typhimurium interbacterial killing. *Infect. Immun.* 83:7, 2738-2750. doi:10.1128/IAI.00198-15.

Carden, S., Okoro, C., Dougan, G., and Monack, D. (2015). Non-typhoidal salmonella typhimurium ST313 isolates that cause bacteremia in humans stimulate less inflammasome activation than ST19 isolates associated with gastroenteritis. *Pathog. Dis.* 73:4, 10.1093/femspd/ftu023. Epub 2014 Dec 24. doi:10.1093/femspd/ftu023.

Cegelski, L., Marshall, G.R., Eldridge, G.R., and Hultgren, S.J. (2008). The biology and future prospects of antivirulence therapies. *Nature Reviews Microbiology.* 6:1, 17-27.

Chan, K., Kim, C.C., and Falkow, S. (2005). Microarray-based detection of salmonella enterica serovar typhimurium transposon mutants that cannot survive in macrophages and mice. *Infect. Immun.* 73:9, 5438-5449. doi:73/9/5438.

Chaudhuri, R.R., Morgan, E., Peters, S.E., Pleasance, S.J., Hudson, D.L., Davies, H.M., Wang, J., van Diemen, P.M., Buckley, A.M., and Bowen, A.J. (2013). Comprehensive assignment of roles for salmonella typhimurium genes in intestinal colonization of food-producing animals. *PLoS Genet.* 9:4, e1003456.

Cotter, P.D., and Hill, C. (2003). Surviving the acid test: Responses of gram-positive bacteria to low pH. *Microbiol. Mol. Biol. Rev.* 67:3, 429-53, table of contents.

Crim, S.M., Griffin, P.M., Tauxe, R., Marder, E.P., Gilliss, D., Cronquist, A.B., Cartter, M., Tobin-D'Angelo, M., Blythe, D., and Smith, K. (2015). Preliminary incidence and trends of infection with pathogens transmitted commonly through food—Foodborne diseases active surveillance network, 10 US sites, 2006–2014. *MMWR Morb. Mortal. Wkly. Rep.* 64, 495-499.

Darcan, C., Ozkanca, R., Idil, O., and Flint, K.P. (2009). Viable but non-culturable state (VBNC) of escherichia coli related to EnvZ under the effect of pH, starvation and osmotic stress in sea water. *Pol. J. Microbiol.* 58:4, 307-317.

Dawoud, T.M., Jiang, T., Mandal, R.K., Ricke, S.C., and Kwon, Y.M. (2014). Improving the efficiency of transposon mutagenesis in salmonella enteritidis by overcoming host-restriction barriers. *Mol. Biotechnol.* 56:11, 1004-1010.

de Mendonca, J.D., Ely, F., Palma, M.S., Frazzon, J., Basso, L.A., and Santos, D.S. (2007). Functional characterization by genetic complementation of aroB-encoded dehydroquinate synthase from mycobacterium tuberculosis H37Rv and its heterologous expression and purification. *J. Bacteriol.* 189:17, 6246-6252. doi:JB.00425-07.

Denoncin, K., Vertommen, D., Arts, I.S., Goemans, C.V., Rahuel-Clermont, S., Messens, J., and Collet, J.F. (2014). A new role for escherichia coli DsbC protein in protection against oxidative stress. *J. Biol. Chem.* 289:18, 12356-12364. doi:10.1074/jbc.M114.554055.

Dorsey, C.W., Laarakker, M.C., Humphries, A.D., Weening, E.H., and Bäumler, A.J. (2005). Salmonella enterica serotype typhimurium MisL is an intestinal colonization factor that binds fibronectin. *Mol. Microbiol.* 57:1, 196-211.

Dougan, G., Chatfield, S.N., Higgins, C.F., and Dorman, C.J. (1996). *Vaccines Comprising Attenuated Salmonella having Mutations in the Ompr Genes.*

D'Urzo, N., Martinelli, M., Pezzicoli, A., De Cesare, V., Pinto, V., Margarit, I., Telford, J.L., Maione, D., and Members of the DEVANI Study Group. (2014). Acidic pH strongly enhances in vitro biofilm formation by a subset of hypervirulent ST-17 streptococcus agalactiae strains. *Appl. Environ. Microbiol.* 80:7, 2176-2185. doi:10.1128/AEM.03627-13.

Edwards, A., Dharamsi, A., Vedadi, M., Domagala, M., Pinder, B., Alam, M.Z., Nethery, K., Houston, S., Kimber, M., and Clarke, T. (2004). Crystal structures of bacterial ribulose-phosphate 3-epimerases.

Feasey, N.A., Dougan, G., Kingsley, R.A., Heyderman, R.S., and Gordon, M.A. (2012). Invasive non-typhoidal salmonella disease: An emerging and neglected tropical disease in africa. *The Lancet.* 379:9835, 2489-2499.

Figueira, R., Watson, K.G., Holden, D.W., and Helaine, S. (2013). Identification of salmonella pathogenicity island-2 type III secretion system effectors involved in intramacrophage replication of S. enterica serovar typhimurium: Implications for rational vaccine design. *MBio.* 4:2, e00065-13. doi:10.1128/mBio.00065-13.

Foss, E.J., Radulovic, D., Shaffer, S.A., Ruderfer, D.M., Bedalov, A., Goodlett, D.R., and Kruglyak, L. (2007). Genetic basis of proteome variation in yeast. *Nat. Genet.* 39:11, 1369-1375.

Freeman, Z.N., Dorus, S., and Waterfield, N.R. (2013). The KdpD/KdpE two-component system: Integrating K homeostasis and virulence. *PLoS Pathog.* 9:3, e1003201.

Fu, J., Keurentjes, J.J., Bouwmeester, H., America, T., Verstappen, F.W., Ward, J.L., Beale, M.H., De Vos, R.C., Dijkstra, M., and Scheltema, R.A. (2009). System-wide molecular evidence for phenotypic buffering in arabidopsis. *Nat. Genet.* 41:2, 166-167.

Garcia-Del Portillo, F., Pucciarelli, M.G., and Casadesus, J. (1999). DNA adenine methylase mutants of salmonella typhimurium show defects in protein secretion, cell invasion, and M cell cytotoxicity. *Proc. Natl. Acad. Sci. U. S. A.* 96:20, 11578-11583.

Garmory, H.S., and Titball, R.W. (2004). ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infect. Immun.* 72:12, 6757-6763. doi:72/12/6757.

Ghazalpour, A., Bennett, B., Petyuk, V.A., Orozco, L., Hagopian, R., Mungrue, I.N., Farber, C.R., Sinsheimer, J., Kang, H.M., and Furlotte, N. (2011). Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet.* 7:6, e1001393.

Gibson, D.L., White, A.P., Snyder, S.D., Martin, S., Heiss, C., Azadi, P., Surette, M., and Kay, W.W. (2006). Salmonella produces an O-antigen capsule regulated by AgfD and important for environmental persistence. *J. Bacteriol.* 188:22, 7722-7730. doi:188/22/7722.

Gogol, E.B., Rhodius, V.A., Papenfort, K., Vogel, J., and Gross, C.A. (2011). Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc. Natl. Acad. Sci. U. S. A.* 108:31, 12875-12880. doi:10.1073/pnas.1109379108.

Gunn, J.S., Alpuche-Aranda, C.M., Loomis, W.P., Belden, W.J., and Miller, S.I. (1995). Characterization of the salmonella typhimurium pagC/pagD chromosomal region. *J. Bacteriol.* 177:17, 5040-5047.

Ha, S., Maciorowski, K., Kwon, Y., Jones, F., and Ricke, S. (1998). Survivability of indigenous microflora and a salmonella typhimurium marker strain in poultry mash treated with buffered propionic acid. *Anim. Feed Sci. Technol.* 75:2, 145-155.

Hansen, S., Lewis, K., and Vulic, M. (2008). Role of global regulators and nucleotide metabolism in antibiotic tolerance in escherichia coli. *Antimicrob. Agents Chemother.* 52:8, 2718-2726. doi:10.1128/AAC.00144-08.

Hartman, C.E., Samuels, D.J., and Karls, A.C. (2016). Modulating salmonella typhimurium's response to a changing environment through bacterial enhancer-binding proteins and the RpoN regulon. *Frontiers in Molecular Biosciences.* 3, 41.

Hasan, C.M.M., and Shimizu, K. (2008). Effect of temperature up-shift on fermentation and metabolic characteristics in view of gene expressions in escherichia coli. *Microbial Cell Factories.* 7:1, 1.

Henderson, M.L., and Kreuzer, K.N. (2015). Functions that protect escherichia coli from tightly bound DNA-protein complexes created by mutant EcoRII methyltransferase. *PLoS One.* 10:5, e0128092. doi:10.1371/journal.pone.0128092.

Hensel, M. (2000). Salmonella pathogenicity island 2. *Mol. Microbiol.* 36:5, 1015-1023.

Hur, J., Xiang, Z., Feldman, E.L., and He, Y. (2011). Ontology-based brucella vaccine literature indexing and systematic analysis of gene-vaccine association network. *BMC Immunology.* 12:1, 1.

Hwang, S., Jeon, B., Yun, J., and Ryu, S. (2011). Roles of RpoN in the resistance of campylobacter jejuni under various stress conditions. *BMC Microbiol.* 11, 207-2180-11-207. doi:10.1186/1471-2180-11-207.

Imlay, J.A., and Linn, S. (1986). Bimodal pattern of killing of DNA-repair-defective or anoxically grown escherichia coli by hydrogen peroxide. *J. Bacteriol.* 166:2, 519-527.

Jakomin, M., Chessa, D., Baumler, A.J., and Casadesus, J. (2008). Regulation of the salmonella enterica std fimbrial operon by DNA adenine methylation, SeqA, and HdfR. *J. Bacteriol.* 190:22, 7406-7413. doi:10.1128/JB.01136-08.

Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T., and Bork, P. (2008). eggNOG: Automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res.* 36, D250-4. doi:gkm796.

Jitprasutwit, S., Ong, C., Juntawieng, N., Ooi, W.F., Hemsley, C.M., Vattanaviboon, P., Titball, R.W., Tan, P., and Korbsrisate, S. (2014). Transcriptional profiles of burkholderia pseudomallei reveal the direct and indirect roles of sigma E under oxidative stress conditions. *BMC Genomics.* 15, 1.

Kang, J., and Blaser, M.J. (2006). UvrD helicase suppresses recombination and DNA damage-induced deletions. *J. Bacteriol.* 188:15, 5450-5459. doi:188/15/5450.

Karash, S., Liyanage, R., Qassab, A., Lay, J.O., and Kwon, Y.M. (2017). A comprehensive assessment of the genetic determinants in salmonella typhimurium for resistance to hydrogen peroxide using proteogenomics. *Scientific Reports.* 7, 17073.

Kato, A., and Groisman, E.A. (2008). The PhoQ/PhoP regulatory network of Salmonella enterica. In Bacterial Signal Transduction: Networks and Drug Targets, Springer), pp. 7-21.

Kenyon, W.J., Sayers, D.G., Humphreys, S., Roberts, M., and Spector, M.P. (2002). The starvation-stress response of salmonella enterica serovar typhimurium requires σE-, but not CpxR-regulated extracytoplasmic functions. *Microbiology.* 148:1, 113-122.

Khatiwara, A., Jiang, T., Sung, S.S., Dawoud, T., Kim, J.N., Bhattacharya, D., Kim, H.B., Ricke, S.C., and Kwon, Y.M. (2012). Genome scanning for conditionally essential genes in salmonella enterica serotype typhimurium. *Appl. Environ. Microbiol.* 78:9, 3098-3107. doi:10.1128/AEM.06865-11.

Klein, J.R., Fahlen, T.F., and Jones, B.D. (2000). Transcriptional organization and function of invasion genes within salmonella enterica serovar typhimurium pathogenicity island 1, including the prgH, prgI, prgJ, prgK, orgA, orgB, and orgC genes. *Infect. Immun.* 68:6, 3368-3376.

Knuth, K., Niesalla, H., Hueck, C.J., and Fuchs, T.M. (2004). Large-scale identification of essential salmonella genes by trapping lethal insertions. *Mol. Microbiol.* 51:6, 1729-1744.

Kong, Q., Yang, J., Liu, Q., Alamuri, P., Roland, K.L., and Curtiss, R.,3rd. (2011). Effect of deletion of genes involved in lipopolysaccharide core and O-antigen synthesis on virulence and immunogenicity of salmonella enterica serovar typhimurium. *Infect. Immun.* 79:10, 4227-4239. doi:10.1128/IAI.05398-11.

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S.K., Hammarlöf, D.L., Canals, R., Grissom, J.E., Conway, T., and Hokamp, K. (2013). An infection-relevant transcriptomic compendium for salmonella enterica serovar typhimurium. *Cell Host & Microbe.* 14:6, 683-695.

Kroupitski, Y., Brandl, M., Pinto, R., Belausov, E., Tamir-Ariel, D., Burdman, S., and Sela, S. (2013). Identification of salmonella enterica genes with a role in persistence on lettuce leaves

during cold storage by recombinase-based in vivo expression technology. *Phytopathology.* 103:4, 362-372.

Krulwich, T.A., Sachs, G., and Padan, E. (2011). Molecular aspects of bacterial pH sensing and homeostasis. *Nature Reviews Microbiology.* 9:5, 330-343.

Kwon, Y.M., Ricke, S.C., and Mandal, R.K. (2016). Transposon sequencing: Methods and expanding applications. *Appl. Microbiol. Biotechnol.* 100:1, 31-43.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods.* 9:4, 357-359.

Lawhon, S.D. (2002). Genetic and environmental regulation of virulence genes in salmonella enterica serovar typhimurium.

Lawley, T.D., Chan, K., Thompson, L.J., Kim, C.C., Govoni, G.R., and Monack, D.M. (2006). Genome-wide screen for salmonella genes required for long-term systemic infection of the mouse. *PLoS Pathog.* 2:2, e11.

Lazinski, D.W., and Camilli, A. (2013). Homopolymer tail-mediated ligation PCR: A streamlined and highly efficient method for DNA cloning and library construction. *BioTechniques.* 54:1, 25-34. doi:10.2144/000113981.

Lee, E., and Groisman, E.A. (2012). Tandem attenuators control expression of the salmonella mgtCBR virulence operon. *Mol. Microbiol.* 86:1, 212-224.

Lee, Y.H., Kim, S., Helmann, J.D., Kim, B., and Park, Y.K. (2013). RaoN, a small RNA encoded within salmonella pathogenicity island-11, confers resistance to macrophage-induced stress. *Microbiology.* 159:7, 1366-1378.

Lee, E.J., Choi, J., and Groisman, E.A. (2014). Control of a salmonella virulence operon by proline-charged tRNA(pro). *Proc. Natl. Acad. Sci. U. S. A.* 111:8, 3140-3145. doi:10.1073/pnas.1316209111.

Leonhartsberger, S., Korsa, I., and Bock, A. (2002). The molecular biology of formate metabolism in enterobacteria. *J. Mol. Microbiol. Biotechnol.* 4:3, 269-276.

Liu, Y., Wang, Y., Walsh, T.R., Yi, L., Zhang, R., Spencer, J., Doi, Y., Tian, G., Dong, B., and Huang, X. (2015). Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in china: A microbiological and molecular biological study. *The Lancet Infectious Diseases.* 16:2, 161-168.

López-Garrido, J., and Casadesús, J. (2010). The DamX protein of escherichia coli and salmonella enterica. *Gut Microbes.* 1:4, 285-288.

Lostroh, C.P., and Lee, C.A. (2001). The salmonella pathogenicity island-1 type III secretion system. *Microb. Infect.* 3:14, 1281-1291.

Low, D.A., Weyand, N.J., and Mahan, M.J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect. Immun.* 69:12, 7197-7204. doi:10.1128/IAI.69.12.7197-7204.2001.

Lu, P., Lill, H., and Bald, D. (2014). ATP synthase in mycobacteria: Special features and implications for a function as drug target. *Biochimica Et Biophysica Acta (BBA)-Bioenergetics.* 1837:7, 1208-1218.

Lüttmann, D., Göpel, Y., and Görke, B. (2012). The phosphotransferase protein EIIANtr modulates the phosphate starvation response through interaction with histidine kinase PhoR in escherichia coli. *Mol. Microbiol.* 86:1, 96-110.

Ma, Q., and Wood, T.K. (2011). Protein acetylation in prokaryotes increases stress resistance. *Biochem. Biophys. Res. Commun.* 410:4, 846-851.

Majowicz, S.E., Musto, J., Scallan, E., Angulo, F.J., Kirk, M., O'Brien, S.J., Jones, T.F., Fazil, A., Hoekstra, R.M., and International Collaboration on Enteric Disease 'Burden of Illness' Studies. (2010). The global burden of nontyphoidal salmonella gastroenteritis. *Clin. Infect. Dis.* 50:6, 882-889. doi:10.1086/650733.

Mandal, R.K., Jiang, T., and Kwon, Y.M. (2017). Essential genome of campylobacter jejuni. *BMC Genomics.* 18, 616.

Mandal, R.K., and Kwon, Y.M. (2017). Global screening of salmonella enterica serovar typhimurium genes for desiccation survival. *Frontiers in Microbiology.* 8, 1723.

Mandal, R.K. (2016). Genetic determinants of salmonella and campylobacter required for in vitro fitness. Doctoral Dissertation. University of Arkansas.

Marcus, E.A., Sachs, G., and Scott, D.R. (2013). The role of ExbD in periplasmic pH homeostasis in helicobacter pylori. *Helicobacter.* 18:5, 363-372.

Matulová, M., Havlickova, H., Sisak, F., and Rychlik, I. (2012). Vaccination of chickens with salmonella pathogenicity island (SPI) 1 and SPI2 defective mutants of salmonella enterica serovar enteritidis. *Vaccine.* 30:12, 2090-2097.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., and Du, F. (2001). Complete genome sequence of salmonella enterica serovar typhimurium LT2. *Nature.* 413:6858, 852-856.

McKenna, M. (2013). CDC threat report:'We will soon be in a post-antibiotic era'. *Wired Magazine.*

Medini, D., Serruto, D., Parkhill, J., Relman, D.A., Donati, C., Moxon, R., Falkow, S., and Rappuoli, R. (2008). Microbiology in the post-genomic era. *Nature Reviews Microbiology.* 6:6, 419-430.

Michalik, S., Liebeke, M., Zühlke, D., Lalk, M., Bernhardt, J., Gerth, U., and Hecker, M. (2009). Proteolysis during long-term glucose starvation in staphylococcus aureus COL. *Proteomics.* 9:19, 4468-4477.

Mikheil, D.M., Shippy, D.C., Eakley, N.M., Okwumabua, O.E., and Fadl, A.A. (2012). Deletion of gene encoding methyltransferase (gidB) confers high-level antimicrobial resistance in salmonella. *J. Antibiot.* 65:4, 185-192.

Miller, S.I., and Mekalanos, J.J. (1998). *Salmonella Vaccines.*

Mitchell, A., Romano, G.H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., Dahan, O., and Pilpel, Y. (2009). Adaptive prediction of environmental changes by microorganisms. *Nature.* 460:7252, 220-224.

Murinova, S., and Dercova, K. (2014). Response mechanisms of bacterial degraders to environmental contaminants on the level of cell walls and cytoplasmic membrane. *Int. J. Microbiol.* 2014, 873081. doi:10.1155/2014/873081.

Nava, G., Bielke, L., Callaway, T., and Castaneda, M. (2005). Probiotic alternatives to reduce gastrointestinal infections: The poultry experience. *Animal Health Research Reviews.* 6:01, 105-118.

Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K.M., Lee, K.J., and Wong, A. (2011). Phenotypic landscape of a bacterial cell. *Cell.* 144:1, 143-156.

Olekhnovich, I.N., and Kadner, R.J. (2007). Role of nucleoid-associated proteins hha and H-NS in expression of salmonella enterica activators HilD, HilC, and RtsA required for cell invasion. *J. Bacteriol.* 189:19, 6882-6890. doi:JB.00905-07.

Patton, B.S. (2007). Applications and mechanisms of colicin E1. Doctoral Dissertation. Iowa State University.

Pratt, D., Chen, J., Welker, D., Rivas, R., Pillich, R., Rynkov, V., Ono, K., Miello, C., Hicks, L., Szalma, S. and Stojmirovic, A. (2015). NDEx, the network data exchange. *Cell systems*, *1*:4, 302-305.

Pritchard, J.R., Chao, M.C., Abel, S., Davis, B.M., Baranowski, C., Zhang, Y.J., Rubin, E.J., and Waldor, M.K. (2014). ARTIST: High-resolution genome-wide assessment of fitness using transposon-insertion sequencing. *PLoS Genet.* 10:11, e1004782.

Reid, A.N., Pandey, R., Palyada, K., Whitworth, L., Doukhanine, E., and Stintzi, A. (2008). Identification of campylobacter jejuni genes contributing to acid adaptation by transcriptional profiling and genome-wide mutagenesis. *Appl. Environ. Microbiol.* 74:5, 1598-1612. doi:10.1128/AEM.01508-07.

Riordan, J.T., Tietjen, J.A., Walsh, C.W., Gustafson, J.E., and Whittam, T.S. (2010). Inactivation of alternative sigma factor 54 (RpoN) leads to increased acid resistance, and alters locus of enterocyte effacement (LEE) expression in escherichia coli O157: H7. *Microbiology.* 156:3, 719-730.

Rosenberger, C.M., and Finlay, B.B. (2003). Phagocyte sabotage: Disruption of macrophage signalling by bacterial pathogens. *Nature Reviews Molecular Cell Biology.* 4:5, 385-396.

Rowley, G., Spector, M., Kormanec, J., and Roberts, M. (2006). Pushing the envelope: Extracytoplasmic stress responses in bacterial pathogens. *Nature Reviews Microbiology.* 4:5, 383-394.

Schakermann, M., Langklotz, S., and Narberhaus, F. (2013). FtsH-mediated coordination of lipopolysaccharide biosynthesis in escherichia coli correlates with the growth rate and the alarmone (p)ppGpp. *J. Bacteriol.* 195:9, 1912-1919. doi:10.1128/JB.02134-12.

Schmidt, H., and Hensel, M. (2004). Pathogenicity islands in bacterial pathogenesis. *Clin. Microbiol. Rev.* 17:1, 14-56.

Shippy, D.C., and Fadl, A.A. (2014a). Bacterial tRNA modification enzymes: Potential role in biology and virulence. *Biology and Medicine.* 2014.

Shippy, D.C., and Fadl, A.A. (2014b). tRNA modification enzymes GidA and MnmE: Potential role in virulence of bacterial pathogens. *International Journal of Molecular Sciences.* 15:10, 18267-18280.

Silva-Valenzuela, C.A., Molina-Quiroz, R.C., Desai, P., Valenzuela, C., Porwollik, S., Zhao, M., Hoffman, R.M., Andrews-Polymenis, H., Contreras, I., and Santiviago, C.A. (2015). Analysis of two complementary single-gene deletion mutant libraries of salmonella typhimurium in intraperitoneal infection of BALB/c mice. *Frontiers in Microbiology.* 6.

Sleator, R.D., and Hill, C. (2002). Bacterial osmoadaptation: The role of osmolytes in bacterial stress and virulence. *FEMS Microbiol. Rev.* 26:1, 49-71. doi:S0168644501000717.

Smith, J.L. (2003). The role of gastric acid in preventing foodborne disease and how bacteria overcome acid conditions. *Journal of Food Protection®.* 66:7, 1292-1303.

Spencer, B. (2015). Superbug resistant to ALL antibiotics reaches europe: Danish patient becomes infected with untreatable form of salmonella that is 'probably already in britain'. 2016.

Sturm, S., and Timmis, K. (1986). Cloning of the rfb gene region of shigella dysenteriae 1 and construction of an rfb-rfp gene cassette for the development of lipopolysaccharide-based live anti-dysentery vaccines. *Microb. Pathog.* 1:3, 289-297.

Tagkopoulos, I., Liu, Y.C., and Tavazoie, S. (2008). Predictive behavior within microbial genetic networks. *Science.* 320:5881, 1313-1317. doi:10.1126/science.1154456.

Tükel, Ç., Raffatellu, M., Humphries, A.D., Wilson, R.P., Andrews-Polymenis, H.L., Gull, T., Figueiredo, J.F., Wong, M.H., Michelsen, K.S., and Akçelik, M. (2005). CsgA is a pathogen-associated molecular pattern of salmonella enterica serotype typhimurium that is recognized by Toll-like receptor 2. *Mol. Microbiol.* 58:1, 289-304.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40:15, e115-e115.

van der Heijden, J., Bosman, E.S., Reynolds, L.A., and Finlay, B.B. (2015). Direct measurement of oxidative and nitrosative stress dynamics in salmonella inside macrophages. *Proc. Natl. Acad. Sci. U. S. A.* 112:2, 560-565. doi:10.1073/pnas.1414569112.

Van Opijnen, T., Bodi, K.L., and Camilli, A. (2009). Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods.* 6, 767.

Van Opijnen, T., and Camilli, A. (2013). Transposon insertion sequencing: A new tool for systems-level analysis of microorganisms. *Nature Reviews Microbiology.* 11, 435.

van Opijnen, T., Lazinski, D.W., and Camilli, A. (2014). Genome-Wide fitness and genetic interactions determined by Tn-seq, a High-Throughput massively parallel sequencing method for microorganisms. *Current Protocols in Molecular Biology.* 1-7.16. 24.

van Opijnen, T., and Camilli, A. (2012). A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Res.* 22:12, 2541-2551. doi:10.1101/gr.137430.112.

Wakamatsu, T., Kim, K., Uemura, Y., Nakagawa, N., Kuramitsu, S., and Masui, R. (2011). Role of RecJ-like protein with 5'-3' exonuclease activity in oligo(deoxy)nucleotide degradation. *J. Biol. Chem.* 286:4, 2807-2816. doi:10.1074/jbc.M110.161596.

Weber, A., and Jung, K. (2002). Profiling early osmostress-dependent gene expression in escherichia coli using DNA macroarrays. *J. Bacteriol.* 184:19, 5502-5507.

World Health Organization (WHO). *WHO Antimicrobial Resistance: Global Report on Surveillance, April 2014.*

Yin, J., Xia, J., Tao, M., Xu, L., Li, Q., Geng, S., and Jiao, X. (2016). Construction and characterization of a cigR deletion mutant of salmonella enterica serovar pullorum. *Avian Pathol.* 45:5, 569-575.

Zhang, R., Ou, H.Y., and Zhang, C.T. (2004). DEG: A database of essential genes. *Nucleic Acids Res.* 32, D271-2. doi:10.1093/nar/gkh024.

Zhou, L., Zhao, M., Wolf, R.Z., Graham, D.E., and Georgiou, G. (2009). Transcriptional regulation of the escherichia coli gene rraB, encoding a protein inhibitor of RNase E. *J. Bacteriol.* 191:21, 6665-6674. doi:10.1128/JB.00344-09.

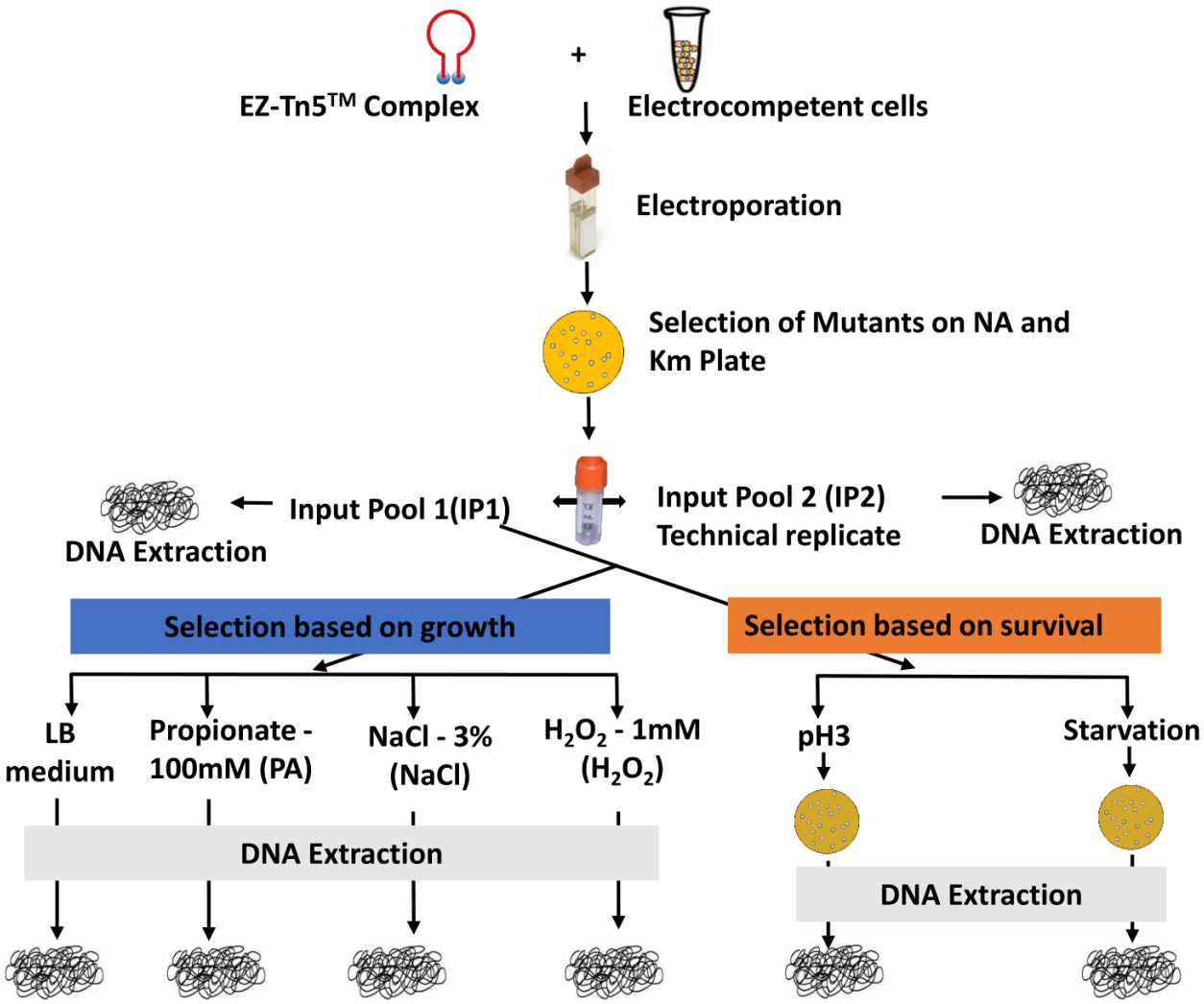


Figure 1. Schematic overview of the experimental design. A highly saturating Tn5 library was constructed through electroporation of EZ-Tn5 transposome complex to *S.* Typhimurium 14028s. Approximately 350,000 Tn5 mutants were collected on LB (Km + NA) plates. Complex Tn5 mutant library (IP1) was selected based on growth [LB medium (LB), 100 mM Propionate in LB medium (PA), 3% NaCl in LB medium (NaCl), and 1mM Hydrogen peroxide in LB medium (H_2O_2)] and survival [exposed to pH3 for 30 min (PH3) and incubated for 12 days in 1X PBS (Starvation)]. Input pool 2 (IP2) was a technical replicate of input pool 1 (IP1).

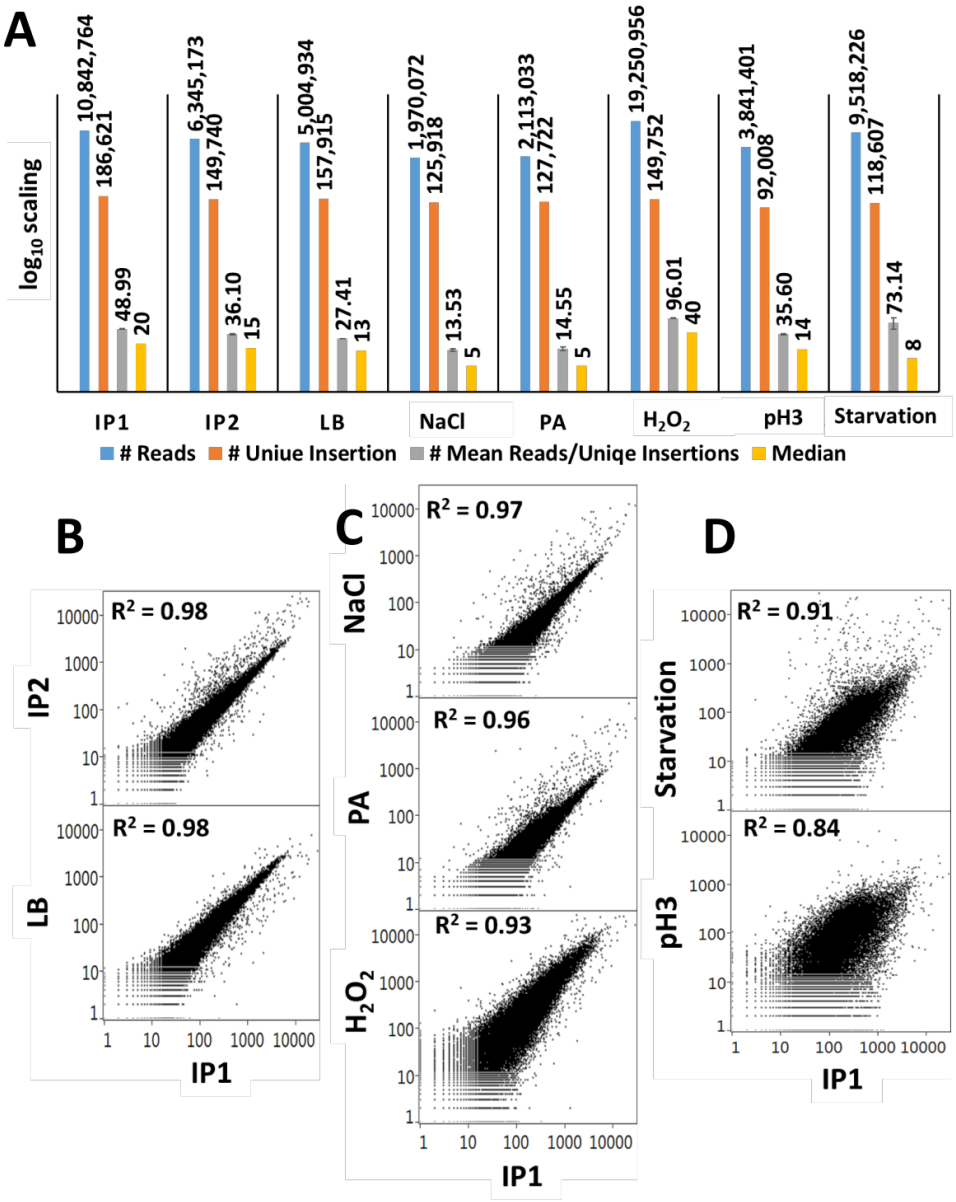


Figure 2. Summary of Illumina sequencing reads and correlation between Tn5 mutant libraries. A) Bar graph shows the number of Illumina sequencing reads distribution in each Tn5 libraries after sorting according to the barcode (blue color), unique insertions (orange color), mean reads per unique insertions (grey color), and median reads for each unique insertions (yellow color). B) Scatter plot displays the Spearman correlation (R^2) among the Tn5 mutant libraries based on read count per 100 bp window across the genome ($p < 0.0001$).

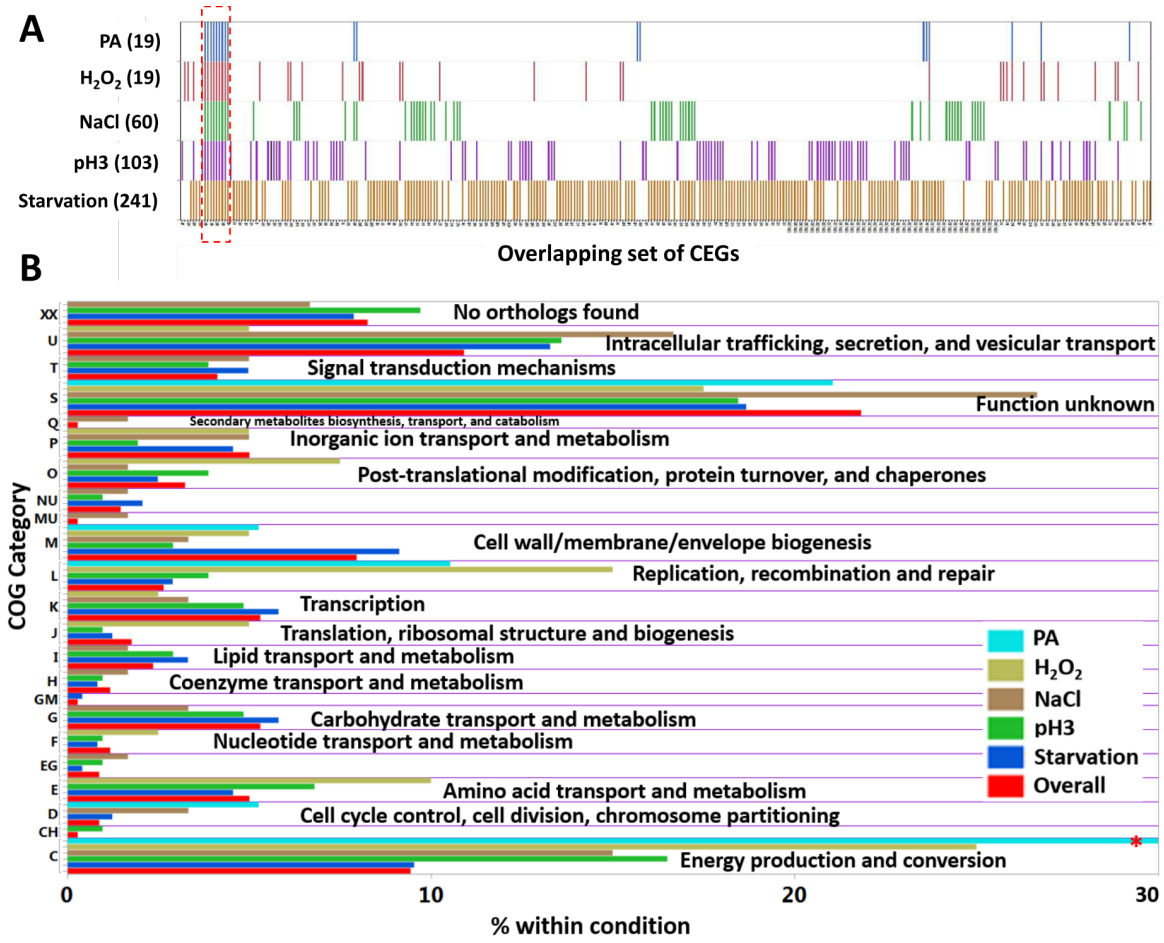


Figure 3. Conditionally essential genes (CEGs) of *S.* Typhimurium 14028s and cluster of orthologous groups (COG). A) Distribution of the overlapping set of 339 CEGs identified in the 5 conditions. Numbers inside the bracket indicate number of CEGs identified. Red dashed box indicates the CEGs (ATP synthase genes) common to all 5 conditions. B) Functional assignments of CEGs into COG category. Overall is the COG assigned to all the 339 CEGs. (Red asterisk (*): Abundance of COG C in PA was 57.89 %).

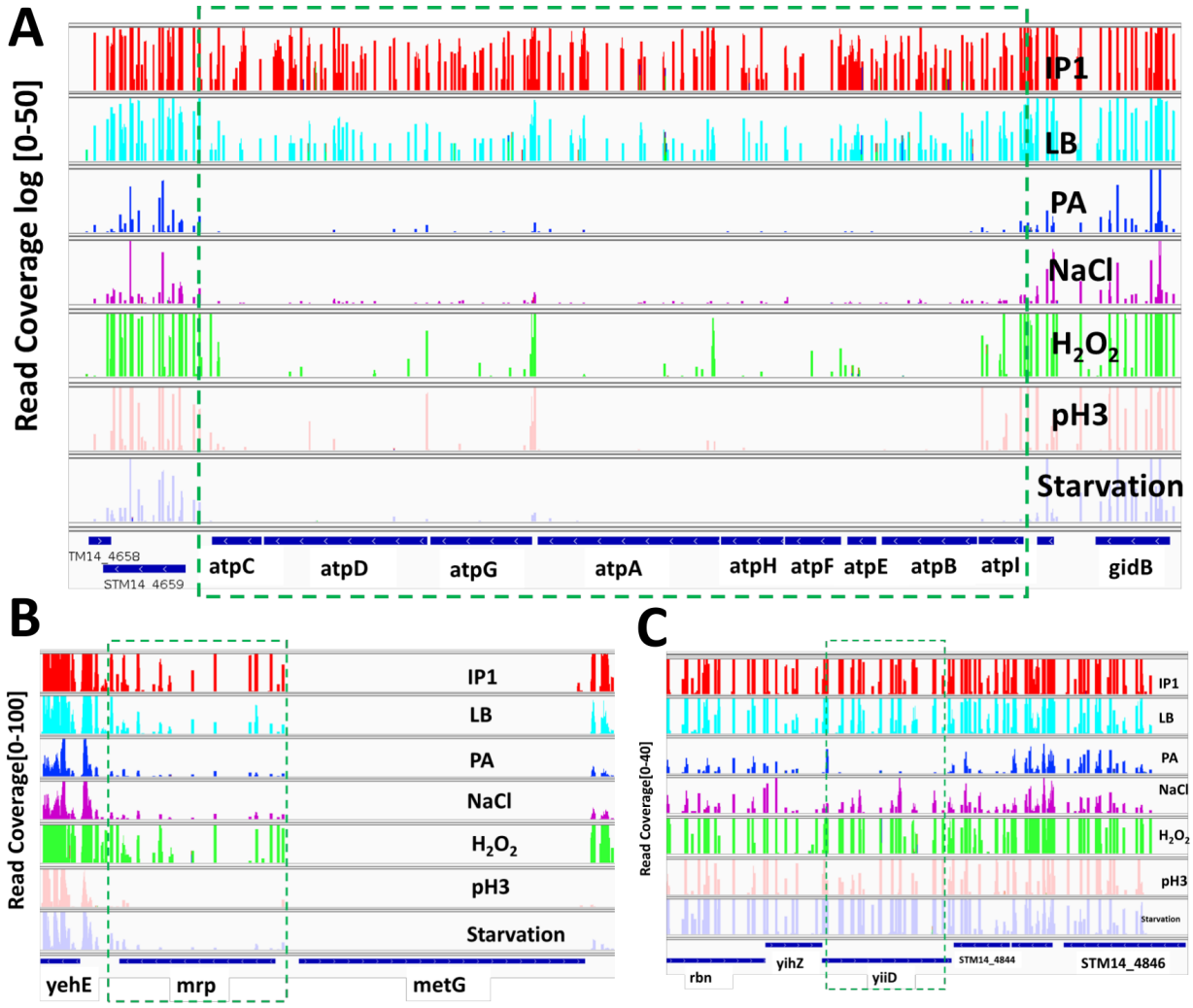


Figure 4. Tn-seq profiles for selected genes across 7 conditions. Y-axis: Numbers in the bracket indicates the raw read coverage. A) ATP synthase genes conditionally essential in all the 5 conditions (PA, NaCl, H_2O_2, PH3 and Starvation. B) Gene *mrp* essential in PH3 and Starvation. C) Gene *yiiD* essential in PA only.

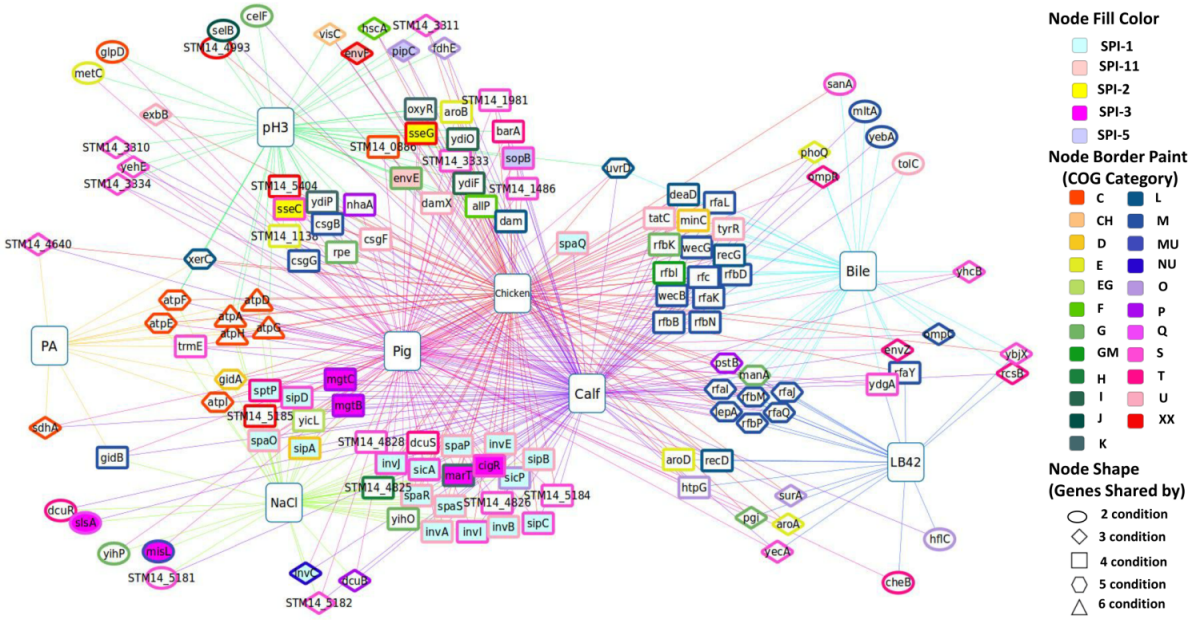


Figure 5. Genotype-phenotype network connections illustrating mechanistic basis of *S.* Typhimurium genetic factors required for enteric infection [*In vitro* vs *in vivo* (Enteric)]. Large square nodes indicate various conditions (studies) and small nodes are fitness genes. Each node (gene) is at least shared by one of the in vitro condition i.e. stressors encountered by *Salmonella* during enteric infection (PA, PH3, NaCl, Bile, and LB42) and at least one of the *in vivo* enteric condition (Pig, Calf, and Chicken). The interactive network through the Network Data Exchange (NDEx) is available at www.ndexbio.org/#/network/027b067d-e209-11e8-aaa6-0ac135e8bacf (Pratt et al. 2015).

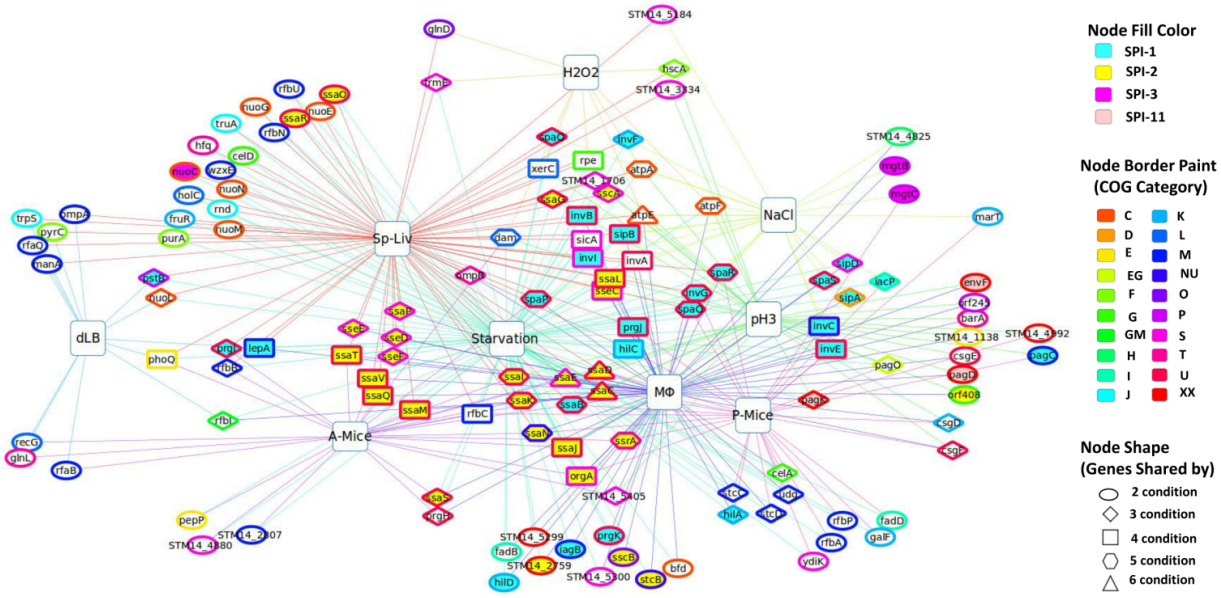


Figure 6. Genotype-phenotype network connections illustrating mechanistic basis of *S.* Typhimurium genetic factors required for systemic infection [*In vitro* vs *In vivo* (Systemic)]. Large square nodes indicate various conditions (studies) and small nodes are fitness genes. Each node (gene) is at least shared by one of the *in vitro* condition i.e. stressors encountered by *Salmonella* inside macrophage (NaCl, H_2O_2, PH3, Starvation, and dLB) and at least one of the *in vivo* systemic condition (MΦ, Sp-Liv, P-Mice, and A-Mice). The interactive network through the Network Data Exchange (NDEx) is available at www.ndexbio.org/#/network/5e78ad70-e209-11e8-aaa6-0ac135e8bacf (Pratt et al. 2015).

Supplemental Materials:

Genetic determinants in *Salmonella enterica* serotype Typhimurium required for overcoming stressors in the host environment

Rabindra K. Mandal[1,3*], Tieshan Jiang[1] and Young Min Kwon[1, 2]

[1]Department of Poultry Science, University of Arkansas, Fayetteville, AR, 72701

[2]Cell and Molecular Biology Program, University of Arkansas, Fayetteville, AR, 72701

3*Corresponding author present address:

Department of Microbiology and Immunology

School of Medicine

University of Louisville

Louisville, KY 40202

Phone: (479) 276-2203

Email: rabindra.mandal@louisville.edu

Figure S1. **Preparation of Tn-seq amplicon library for Illumina sequencing.** A) Genomic DNA of Tn5 mutant library was linealry extended using Tn-specific primer 1 (Ez-Tn5 primer3 in Table S1). Then C-tail was attached to the 3' end of purified single-stranded DNA. The C-tailed product was purified and exponential PCR was performed using Tn-specific primer 2 (Barcoded primers in Table S1) and C-tail specific primer (HTM-Primer in Table S1) with Illumina adapter attached to primers. B) Exponentially amplified DNA was than run on 1.5% agarose gel. DNA from 300bp to 500bp was extracted from the gel and sent for Illumina sequencing. [M: Hi-Lo DNA marker; 1, 2, 3, 4: Tn5 mutant libraries; and C: negative control (gDNA of the wild type *S.* Typhimurium 14028s)].

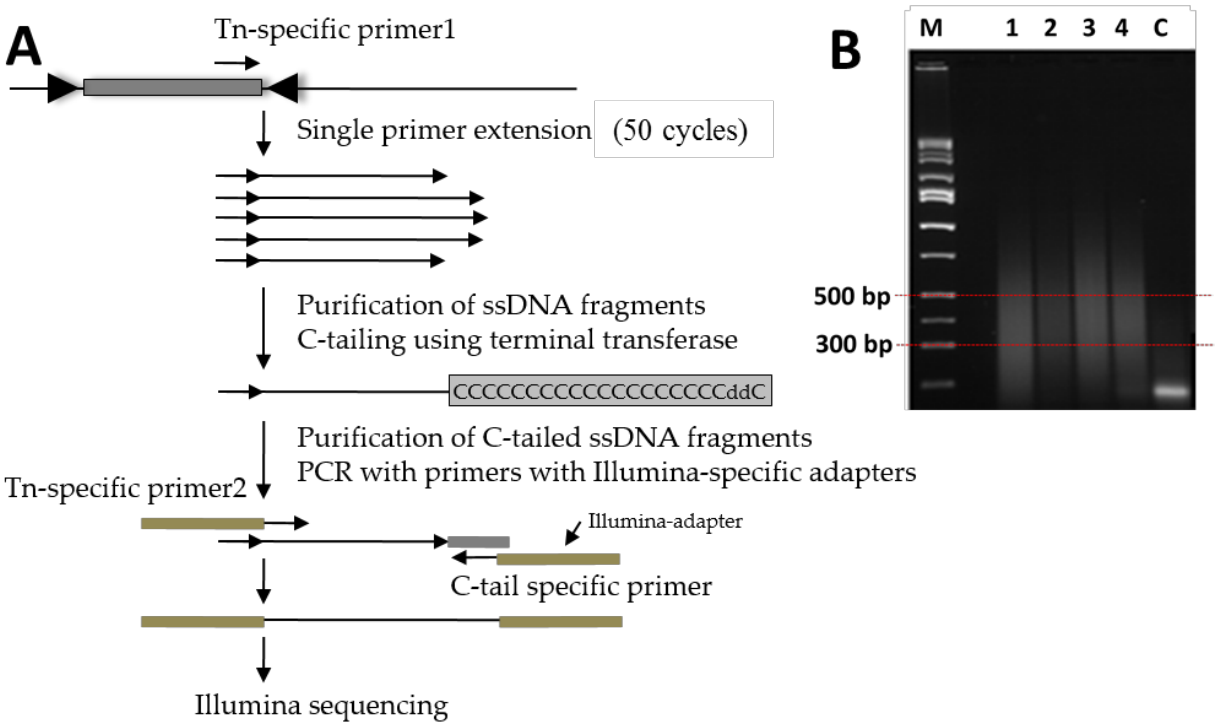


Figure S2. Overlay plot displays global view of genome-wide quantitative distribution of Tn5 insertion read count for all samples. X-axis: Position on the genome; and Y-axis: Number of read count per 100 bp scaled in $\log_{10}$.

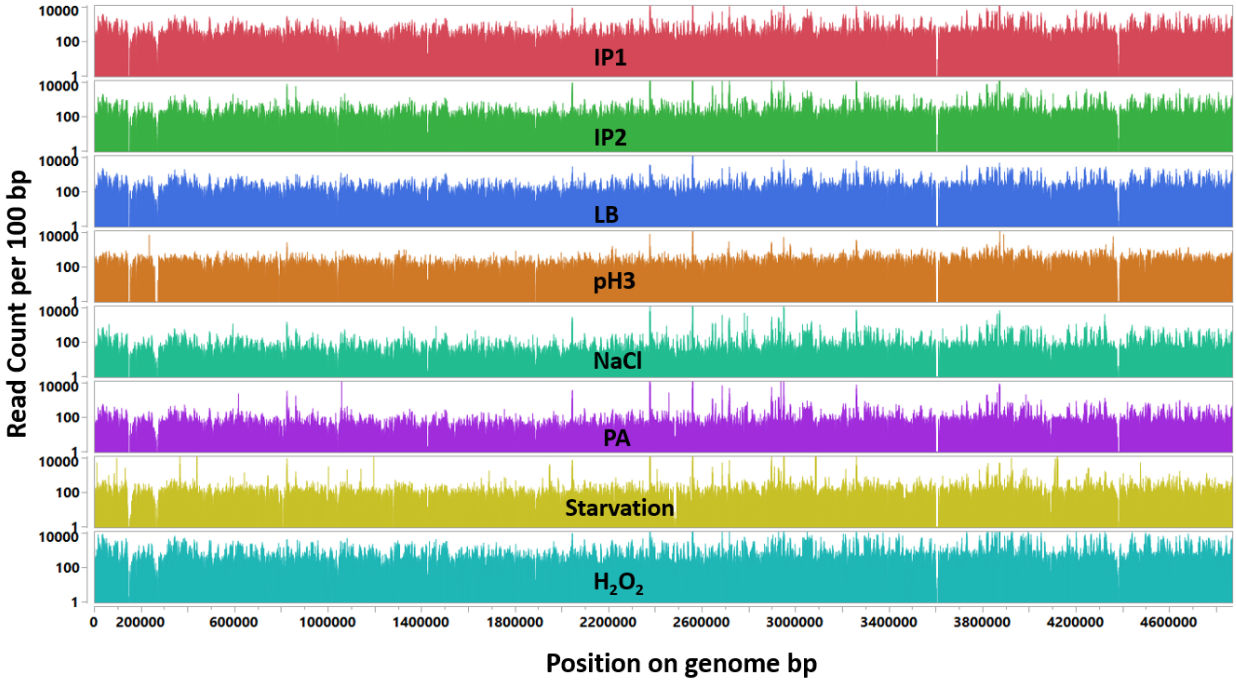


Figure S3. Tn-seq profiles around the selected genomic regions. A) *Salmonella* pathogenicity island 1 (SPI-1) genes encoding type III secretion system (TTSS). Screen shot image produced using Integrative Genomics Viewer (IGV) showing raw read coverage [100-600] in seven conditions. (Blue asterisk: conditionally essential in NaCl and Starvation; and Red asterisk: conditionally essential in Starvation only). B) CpxAR were conditionally essential in starvation. only.

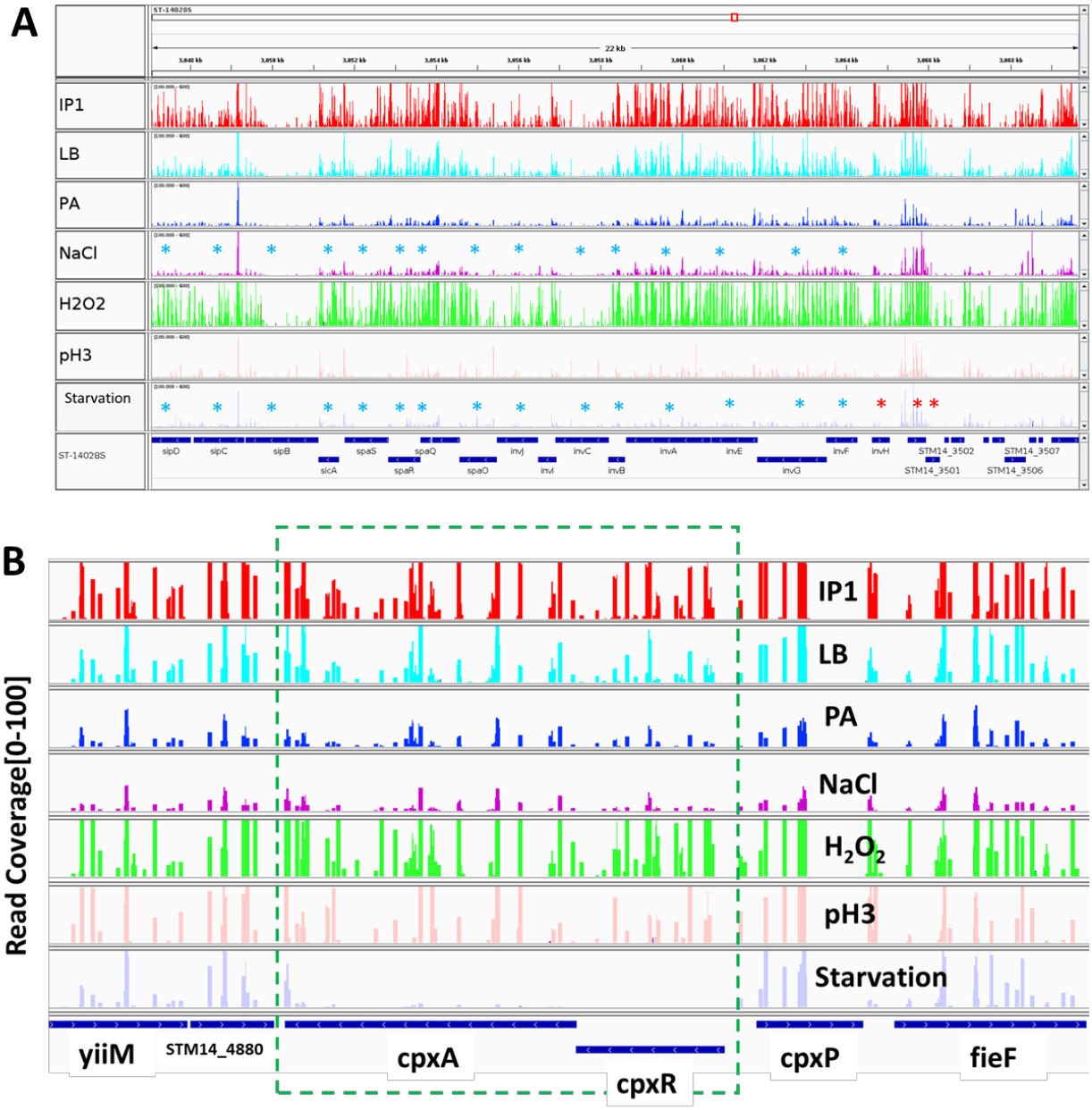


Figure S4. Comparison of the overlapping set of conditionally essential genes of *S.* Typhimurium 14028s (this study) with essential genome of *S.* Typhimurium SL3261 and *S.* Typhimurium LT2.

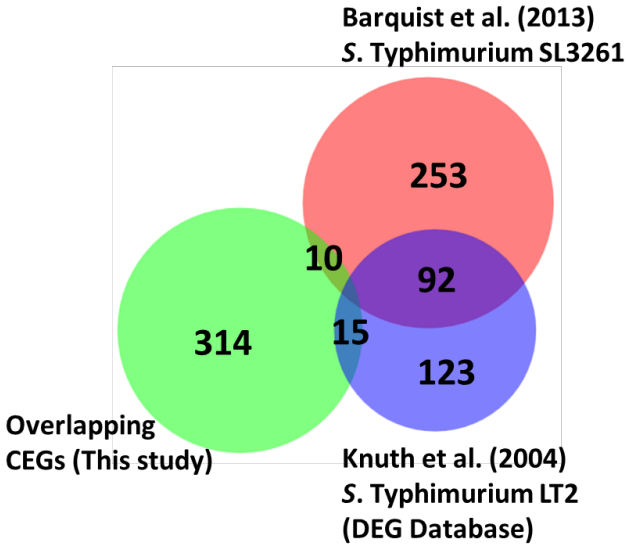
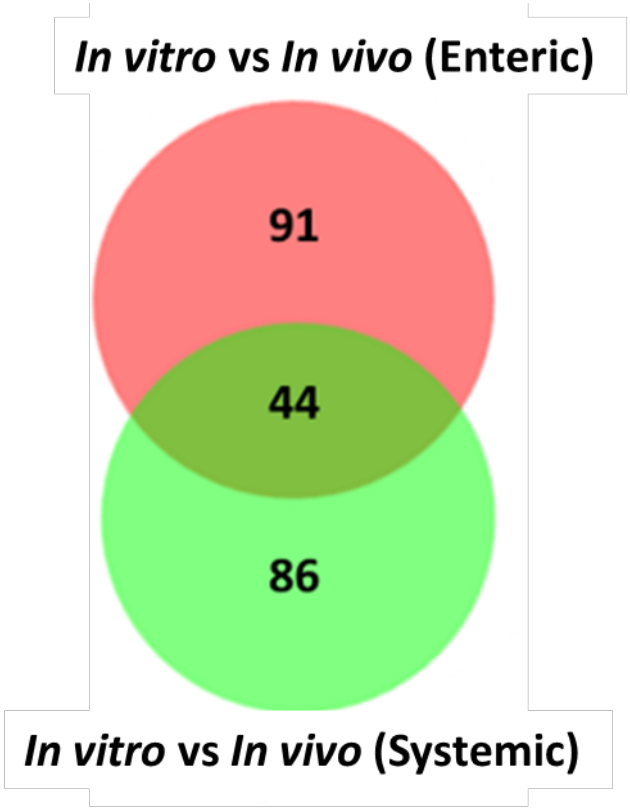


Figure S5. Genes required for enteric, systemic and *in vitro* fitness. Venn diagram shows the number of shared genes between *in vitro* vs *in vivo* (Enteric) (135 CEGs shown in Fig 5) and *in vitro* vs *in vivo* (Systemic) (130 CEGs shown Fig 6). The list of 44 genes required for *in vitro*, enteric and systemic fitness are shown in Table 1.



List of Supplementary Tables

Table S1. Oligonucleotides used in this study.

Table S2. All conditionally essential genes (CEGs) in S. Typhimurium 14028S identified in this study.

Table S3. Comparison of the conditionally essential genes (CEGs) in S. Typhimurium 14028s identified in this study across the 5 stress conditions.

Table S4. The conditionally essential genes (CEGs) in S. Typhimurium 14028s identified in this study that are located in Salmonella Pathogenicity Islands.

Table S5. Comparison of the conditionally essential genes (CEGs) of S. Typhimurium 14028s (this study) with the essential genes of S. Typhimurium identified from previous studies

Table S6. The conditionally essential genes (CEGs) in the presence of the in vitro host stressors (PA, NaCl, pH3, Bile, and LB42) that are also required for enteric infection in farm animals (cattle, pig, and chicken).

Table S7. The conditionally essential genes (CEGs) in the presence of the in vitro host stressors (H2O2, NaCl, pH3, Starvation, and dLB) that are also required for systemic infection (MΦ; infections (MΦ, A-Mice, P-Mice, Sp-Liv).