

SCIENTIFIC REPORTS



OPEN

A Comprehensive Assessment of the Genetic Determinants in *Salmonella* Typhimurium for Resistance to Hydrogen Peroxide Using Proteogenomics

Sardar Karash¹, Rohana Liyanage^{2,3}, Abdullah Qassab¹, Jackson O. Lay Jr.^{2,3} & Young Min Kwon^{1,4}

Salmonella is an intracellular pathogen infecting a wide range of hosts and can survive in macrophages. An essential mechanism used by macrophages to eradicate *Salmonella* is production of reactive oxygen species. Here, we used proteogenomics to determine the candidate genes and proteins that have a role in resistance of *S. Typhimurium* to H₂O₂. For Tn-seq, a saturated Tn5 insertion library was grown *in vitro* under either 2.5 (H₂O₂L) or 3.5 mM H₂O₂ (H₂O₂H). We identified two sets of overlapping genes required for resistance of *S. Typhimurium* to H₂O₂L and H₂O₂H, and the results were validated via phenotypic evaluation of 50 selected mutants. The enriched pathways for H₂O₂ resistance included DNA repair, aromatic amino acid biosynthesis (*aroBK*), Fe-S cluster biosynthesis, iron homeostasis and a putative iron transporter system (*ybbKLM*), and H₂O₂ scavenging enzymes. Proteomics revealed that the majority of essential proteins, including ribosomal proteins, were downregulated upon exposure to H₂O₂. On the contrary, a subset of conditionally essential proteins identified by Tn-seq were analyzed by targeted proteomics, and 70% of them were upregulated by H₂O₂. The identified genes will deepen our understanding on *S. Typhimurium* survival mechanisms in macrophages, and can be exploited to develop new antimicrobial drugs.

Salmonella is a Gram-negative bacterium that infects humans and animals. *Salmonella enterica* has numerous serovars, which include typhoidal and non-typhoidal strains. In contrast to the typhoidal salmonellae which are human restricted pathogens, the non-typhoidal salmonellae (NTS), serovar Enteritidis and Typhimurium, are able to infect a wide range of hosts, causing gastroenteritis¹. The NTS strains, including *Salmonella enterica* serovar Typhimurium, account for 11% (1.2 million cases) of the total foodborne illnesses caused by different pathogens in the United States². It has been estimated that *Salmonella* is responsible for 93.8 million cases of gastroenteritis, leading to 155,000 deaths worldwide annually³. The pathogen remains a continuous threat to the food safety, and public health.

To initiate an infection and survive inside the host, *Salmonella* needs to overcome a myriad of host defense mechanisms. As *Salmonella* reaches the intestine and breaches the epithelial tissue, it enters the macrophages and activates different virulence strategies in order to survive and replicate in them⁴. An essential mechanism uses by the phagocytes to kill and eradicate *Salmonella* is production of reactive oxygen species (ROS). Hydrogen peroxide (H₂O₂), superoxide anion (O₂⁻), and the hydroxyl radical (HO) are derivatives of ROS. The short-lived O₂⁻, produced by the NADPH-dependent phagocytic oxidase, quickly dismutates into H₂O₂, which diffuses across semipermeable bacterial cell membranes. Eventually, Fe²⁺ reduces H₂O₂ to HO via the so called Fenton Reaction⁵⁻⁷. The ROS, including H₂O₂, can damage DNA, iron-sulfur cluster-containing proteins, and other biological molecules in the bacterial cells⁸⁻¹⁰.

¹Cell and Molecular Biology Program, University of Arkansas, Fayetteville, AR, 72701, USA. ²Department of Chemistry, University of Arkansas, Fayetteville, AR, 72701, USA. ³Statewide Mass Spectrometry Facility, Fayetteville, AR, 72701, USA. ⁴Department of Poultry Science, University of Arkansas, Fayetteville, AR, 72701, USA. Correspondence and requests for materials should be addressed to Y.M.K. (email: ykwon@uark.edu)

Numerous genetic factors and proteins that are important for resistance of *S. Typhimurium* to H_2O_2 have been discovered and the underlying mechanisms have been explored^{11,12}. Various approaches and techniques have been employed to study global response of *Salmonella* or related bacteria to H_2O_2 *in vitro* as a model system to simulate the bacterium's response to ROS in phagocytic cells: (i) Two-dimensional gel electrophoresis identified H_2O_2 -induced proteins in *Salmonella*¹³, (ii) DNA microarray identified H_2O_2 induced genes in *E. coli*¹⁴, and (iii) RNA-seq identified H_2O_2 induced genes in *Salmonella*¹⁵. Yet, the factors required for fitness under the given condition cannot be identified with high confidence based on the analysis of transcriptomics or proteomics data¹⁶. Microarray-based tracking of random transposon insertions was used to identify numerous genes in *Salmonella* that are required for survival in mice and macrophages^{17,18}. However, the genetic factors responsible for resistance to ROS cannot be sorted out among all of the genetic factors identified in the study that are required for fitness in the presence of multiple host stressors.

To shed more insights into the underlying mechanisms of *Salmonella* resistance to H_2O_2 , more direct approach linking the gene-phenotype relationships in a genome-wide scale would be necessary. Tn-seq is a powerful approach to allow direct and accurate assessment of the fitness requirement of each gene on the entire genome of a prokaryotic organism¹⁹. In Tn-seq method, a saturated transposon insertion library (input) is exposed to a selective condition, and the mutant population altered through the selection (output) is recovered. Then, the genomic junctions of the transposon insertions are specifically amplified and sequenced from both input and output pools by high-throughput sequencing. The gene fitness can be obtained by calculating the change in relative abundance of the sequence reads corresponding to each gene in the entire genome between the two pools. Tn-seq has been employed to assign gene functions to *Salmonella* genomes in numerous studies: (i) Previously, our lab identified conditionally essential genes that are required for growth in the presence of bile, limited nutrients, and high temperature²⁰, (ii) The genes required for intestinal colonization were identified in chickens, pigs, and cattle²¹, (iii) Candidate essential genes and genes contributing toward bile resistance were identified²², (iv) Core conserved genes for growth in rich media were identified in serovars Typhi and Typhimurium²³. In addition to Tn-seq, electrospray ionization liquid chromatography tandem mass spectrometry (ESI-LC-MS/MS) is a powerful approach for identifying and quantifying proteins in a large scale. The system-wide protein regulation can be determined using mass spectrometry signal intensities of tryptic peptides obtained from two different culture conditions²⁴. The post-translational modification in proteins can be revealed by using proteomic analysis²⁵. Many studies took advantage of proteomic analysis of *Salmonella*. However, to the best of our knowledge, this study is the first to investigate proteogenomics of a bacterium by combining Tn-seq and proteome analysis simultaneously to the same stressor.

In this work, we used Tn-seq method and proteomic analysis in combination to determine system-wide responses of *S. Typhimurium* to two different concentrations of H_2O_2 (H_2O_2L and H_2O_2H). We obtained a comprehensive list of 137 genes that are putatively required for the resistance of *S. Typhimurium* 14028 to H_2O_2 . The roles of 50 selected genes in resistance to H_2O_2 were determined by phenotypic evaluation of the individual deletion mutants. Also, we identified a set of 246 proteins that are differentially expressed in response to H_2O_2 , using data-dependent acquisition (DDA) proteomics, which are largely overlapped with the genes identified by Tn-seq; targeted proteomics showed 70% of the proteins identified by Tn-seq were upregulated by H_2O_2 . In addition to the genes of *S. Typhimurium* previously known to be important for resistance to H_2O_2 , we identified approximately 80 genes that have not been previously associated with resistance to oxidative stress. The results of this study highlighted that the genes in aromatic amino acid biosynthesis, *aroB* and *aroK*, and iron homeostasis, *ybbK*, *ybbL*, and *ybbM*, are crucially important for growth fitness under H_2O_2 stress. The identified candidate genes will expand our understanding on the molecular mechanisms of *Salmonella* survival in macrophages, and serve as new antimicrobial drug targets.

Results and Discussion

The H_2O_2 concentrations and the selections of Tn5 library. First, we sought to determine the growth response of wild type *S. Typhimurium* 14028 cells in LB media containing varying concentrations of H_2O_2 . The wild type cells were grown in LB media that contain different concentrations of H_2O_2 in 96-well plates. After evaluating the growth rates for the cultures, 2.5 and 3.5 mM H_2O_2 were chosen for Tn-seq selections in our study, and termed H_2O_2L and H_2O_2H , respectively. In comparison to *Salmonella* grown in LB media with no H_2O_2 , H_2O_2L and H_2O_2H reduced the growth rates by 10% and 28%, respectively (Fig. 1A). The lag time increased by a 5.7-fold (0.5 vs. 2.9 hr), and an 11-fold (0.5 vs. 5.6 hr) in H_2O_2L and H_2O_2H , respectively. The maximum OD₆₀₀ decreased by only 1% for the H_2O_2L and 2% for the H_2O_2H in comparison to LB media (Fig. 1A).

For the selection of Tn5 library, 20 ml cultures in 300 ml Erlenmeyer flasks containing LB, H_2O_2L , or H_2O_2H were inoculated with the same Tn5 library at the seeding CFUs of the library at 3.5×10^6 . This seeding level provided ~10 CFUs for each Tn5 insertion mutant in the library. The cultures were grown until the mid-exponential phase, in which the CFUs reached 1.17×10^8 (SE 0.01×10^8). It required 7.5 and 9.2 h to reach the cell density as measured by optical density for H_2O_2L and H_2O_2H , respectively, in contrast to 5 h for LB medium (Fig. 1B). We observed some differences in growth responses between the cultures in a 96-well plate and in a 300-ml Erlenmeyer flask. The optical density readings by the plate reader was different in comparison to those by Bio-photometer that we used to measure optical density of the culture in the flask. As a result, the growth curve in Fig. 1A which was based on 96-well plate reader, dose not match exactly with the time required for the Tn5 library to reach the target mid-exponential phase in the flask cultures. In addition, we observed that the H_2O_2 is stable in LB media free of *Salmonella* during the window of time used for the library selection process (Fig. S1), which was also supported by Bogomolnaya *et al.*²⁶.

Preparation of Tn5-seq amplicon library. The *Salmonella* mutants were generated by using the delivery plasmid pBAM1 via conjugation. A total of 325,000 mutant colonies were recovered from 50 plates. Each

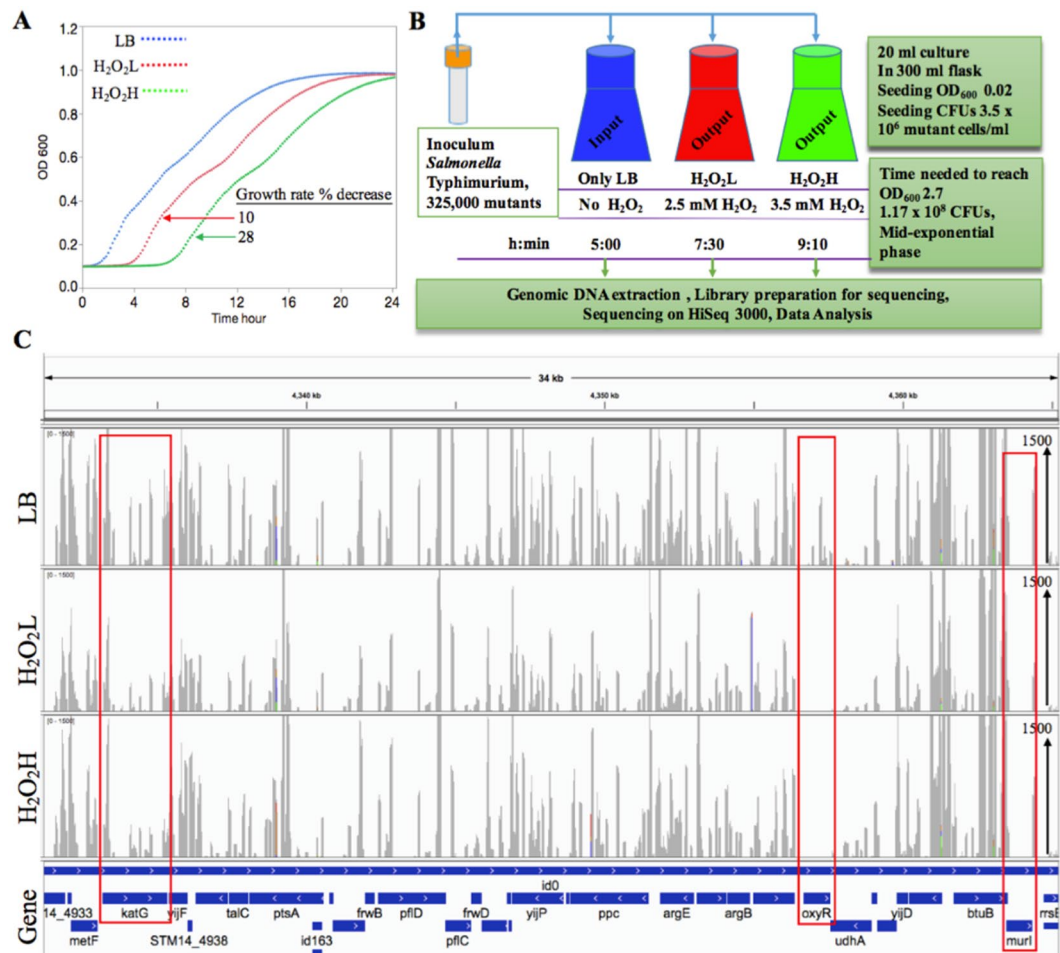


Figure 1. Study design and identification of the genes required for H_2O_2 resistance using Tn-seq. (A) The effect of H_2O_2 on the growth rate of wild type *S. Typhimurium*. An overnight culture of bacteria was diluted 1:200 in the LB medium containing either 0 mM (control: blue), 2.5 mM H_2O_2 (H_2O_2L : red), or 3.5 mM H_2O_2 (H_2O_2H : green). The cultures were incubated at 37 °C for 24 h in a 96-well plate. The reduced growth rates in the presence of H_2O_2 were in reference to LB control. (B) Schematic representation of the Tn-seq experiment design. The *Salmonella* transposon mutant library was inoculated into LB, H_2O_2L and H_2O_2H . The three cultures were grown until they reached mid-exponential phase. Genomic DNA was extracted from each culture and subjected to library preparation, Illumina sequencing, and data analysis. (C) The Tn-seq profiles of the three conditions for the 34 kb genomic region starting from with *metF* gene and ending with *rslE*. The vertical axis represents number of sequencing reads with the maximum of 1500 reads. The genes highlighted in red are *katG*, *oxyR*, and *murI* as examples of non-essential gene, conditionally essential gene, and essential gene, respectively.

mutant contained a single random insertion of Tn5 transposon in the chromosome or plasmid according to DNA sequencing of Tn5-junction sequences for a small set ($n = 71$) of randomly selected Tn5 mutants. We found a significant portion (~20%) of the mutants in the library that were not genuine Tn5 insertions, but the mutants generated as a result of pBAM1 integration into chromosome as determined by their ability to grow in the presence of ampicillin. To prevent the Illumina sequencing reads from being wasted on sequencing Tn5 junctions from these cointegrants, we digested genomic DNA of the input and output libraries with PvuII, which digests immediately outside the inverted repeats on both sides of Tn5. The digested DNA was then used to prepare Tn-seq amplicon library as described in Materials and Methods. Our Tn-seq data analysis indicated that our strategy of removing the DNA sequences originating from cointegrants was effective because only 0.55% of the total HiSeq reads corresponding to Tn5-junctions matched to pBAM1. It should be possible to remove them completely by ensuring complete digestion of genomic DNA with PvuII (Table S1). The method for Tn-seq amplicon library we developed and used in this study has multiple advantages over other Tn-seq protocols, because our method requires only 100 ng of the genomic DNA, and the whole process can be completed in a day²⁷. Additionally, the average length of genomic junction sequences is much longer (Table S1; 92–94 bp) as compared to other Tn-seq methods, which allows more accurate mapping of the insertion sites. When the extension step in the protocol was performed using a conventional 20 nucleotide primer, and the final products of exponential PCR were separated on agarose gel electrophoresis, even the negative controls (the wild type genomic DNA or mutant library genomic DNA without linear extension) showed smear patterns of nonspecific background amplification. However, when

dual priming oligonucleotide (DPO) primer was used in place of the conventional primer for linear extension, non-specific background amplification was completely disappeared. Therefore, we adopted the DPO primer in linear extension step for all library samples in this study. Then, the single-stranded extension products were C-tailed, and used as templates for the exponential PCR step using nested primer specific to Tn5 and poly G primer that contain Illumina adapter sequences along with sample index sequences (Fig. S2). The final PCR products were separated on an agarose gel, and the fragments within the range of 325–625 bp were gel-purified. After pooling of multiple samples, the combined library was sequenced on a HiSeq 3000 run.

Summary of Tn-seq DNA analysis. After de-multiplexing and C-tail trimming of all sequence reads, ~72 million reads of Tn5- junctions with mean read length of 94 bp were obtained. The number of the reads mapped to the complete genome of *S. Typhimurium* 14028 were ~25, 15, and 19 million for LB, H₂O₂L, and H₂O₂H, respectively. The number of unique insertions on the chromosome were 125,449 in the input library, excluding the plasmid (Table S1). On average, Tn5 was inserted in every 39 bp. Number of raw reads per open reading frame (ORF) for H₂O₂L was plotted over the corresponding number of H₂O₂H, which yielded an R² of 0.91, indicating the mutants in the input library quantitatively responded in a similar way for both H₂O₂L and H₂O₂H as expected (Fig. S3). The insertions were mapped to 5,428 genes or 8,022 genes/intergenic regions. Interestingly, the ORF STM14_5121, which is 16.7 kbp long, had the highest number of insertions (~700 insertions) and reads (0.25 M).

Comparison of various bioinformatics pipelines for Tn-seq data analysis. We used 3 different Tn-seq analysis tools to identify the genes and compare the results across the methods with the goal of comprehensive identification of “all” genes required for resistance to H₂O₂. The first tool, ARTIST²⁸, created small non-overlapping genomic windows of 100 bp and the reads from each window were arbitrarily assigned into the middle of the window. The default normalization script of the tool was used. Then, the relative proportions of insertion sites in the output library versus the input were tabulated. Mann-Whiney *U* (MWU) test was used to assess the essentiality of the locus. To consider a gene/intergenic region conditionally essential for growth in the presence of H₂O₂, *p* value had to be ≤0.05 in 90 of the 100 conducted MWU tests. Subsequently, 20 genes and 1 intergenic region were identified for H₂O₂L and 4 genes for H₂O₂H (Table S2). The fact that a smaller number of genes were identified with the higher concentration of H₂O₂ is counterintuitive. It is unclear if our result accurately reflected the number of the genes important for two different concentrations of H₂O₂ or it was due to the shortcoming in the algorithm we used for gene identification.

The second tool, Tn-seq Explorer²⁹, counted insertions in overlapping windows of a fixed size. Using a 550 bp window size, each annotated gene was assigned an essentiality index (EI) which is determined mainly based on the insertion count in a window in this gene. The bimodal distribution of insertion counts per window divided the essential genes to the left and the non-essential genes to the right. To find conditional essential genes, the EI of the output was subtracted from the EI of the input. The genes with negative ΔEI were ranked based on the change in read fold change (Log₂ (H₂O₂L or H₂O₂H/Input)). We found 114 consensus genes between H₂O₂L and H₂O₂H that had at least four-fold reduction in H₂O₂H read counts as compared to the input. The four-fold reduction (Log₂FC = -2) threshold was chosen because some previously known genes that are important for oxidative stress such as *sufS* or *uvrD* were excluded with the more stringent cutoff such as Log₂FC = -3 (Table S2).

The third tool, TRANSIT³⁰, determined read counts of genes in the input and output library. The differences of total read counts between the input and outputs were obtained. The insertion sites were permuted for a number that is specified by the user (we used 10,000 sample). This sampling for each gene gave difference in read counts. The *p* value was calculated from the null distribution of the difference in read counts. We identified 8 and 21 genes for the H₂O₂L and H₂O₂H, respectively, using a *p* value ≤ 0.05 (Table S2).

The combined list of the genes identified by the 3 Tn-seq analysis tools for both H₂O₂L and H₂O₂H included 137 genes (Table S2). All of the genes on this list are expected to have a role in conferring resistance to H₂O₂ and allow *Salmonella* to survive and replicate in the presence of H₂O₂ *in vitro*. Utilizing DAVID and KEGG bioinformatics resources for pathway analysis^{31,32}, the 137 genes were categorized into 15 known pathways with the details available in supplementary information. Of the 21 genes identified by TRANSIT, 19 of these genes were also identified by Tn-seq Explorer, but only 3 out of this 21 were identified by ARTIST. The 19 genes were *hscA*, *rhsR*, *fepD*, *efp*, *oxyR*, *polA*, *ybaD*, *aroD*, *ruvA*, *xthA*, *dps*, *aroB*, *uvrD*, *tonB*, *uvrA*, *aroK*, *ybbM*, *lon*, and *proC*. Two genes, *fepD* and *xthA*, were identified by the all 3 methods and for both conditions (Fig. S4).

The 3 Tn-seq analysis tools are very valuable for Tn-seq data analysis, but each tool has its own advantages and disadvantages. For ARTIST, (i) the user must know how to run scripts in Matlab software, (ii) the analysis is very slow on a personal computer with the HiSeq data, (iii) it has only one method for normalization, but (iv) it can search for essentiality in the intergenic regions. For Tn-seq Explorer, (i) there is no data normalization, and (ii) prediction on small genes is prone to be inaccurate, but (iii) it is very user-friendly and runs fast. For TRANSIT, (i) the user should have some knowledge on running scripts on terminal, (ii) it may need some modification in its Python script according to the way the library was prepared for sequencing, and (iii) a few software packages should be installed on the computer as TRANSIT pre-requisites, but (iv) it does have 6 different methods for data normalization and it runs very fast on a personal computer. We initially aimed at finding the most suitable analytical method to use for analysis of all Tn-seq data in this study. However, unlike our initial expectation, we found the results of the 3 methods did not overlap much (Fig. S4), making it difficult to choose one best method for our Tn-seq data. The majority of the genes uniquely identified by one method were those identified by Tn-seq Explorer. However, it is not surprising considering the distinctive conceptual basis employed by Tn-seq Explorer. We argue that these differences in the 3 methods reflect that the methods are complementary to each other, and we chose to combine the results of the 3 methods to have a comprehensive gene list. Although ARTIST and Tn-seq Explorer are very useful tools for Tn-seq data analysis, we prefer using TRANSIT in our future data

analysis for conditionally essential genes. In the following sections, we continued the downstream analysis mainly based on the 137 genes that include all of the genes identified by all 3 methods.

Validation of Tn-seq results using individual mutants. For the selected 50 genes among the 137 genes identified by Tn-seq, the growth phenotype was determined using individual single deletion mutants in LB, H₂O₂L, and H₂O₂H. The genes were considered to play a role in resistance to H₂O₂, if (i) lag phase time increased, (ii) growth rate reduced or (iii) maximum OD₆₀₀ decreased in the presence of H₂O₂ in comparison to the wild type strain grown in the same conditions. Of the 50 single deletion mutants, 42 mutants were shown to have a role in resistance to H₂O₂ (Fig. 2 and Table S4). One gene, *yhaD*, was identified by all 3 analysis tools, but it did not show the expected phenotype. The *fliD* was also identified by ARTIST, but did not show any phenotype distinguishable from the wild type. The remaining 6 genes that did not show the phenotype was identified by Tn-seq Explorer. Based on the results of the individual mutant assay, we conclude that 84% (42/50) of the genes identified by the Tn-seq analysis and tested using single deletion mutants have a role for resistance to H₂O₂. These results indicate that our Tn-seq analysis identified the genes in *S. Typhimurium* that are required for the wild type level resistance to H₂O₂ with high accuracy.

Proteomics of H₂O₂ response. With ESI-LC-MS/MS in data-dependent acquisition (DDA) mode, the protein regulation was determined using MS1 filtering technique that skyline software offers³³. It uses signal intensities of tryptic peptides derived from the proteins of wild type strain grown in the presence of H₂O₂ in comparison to the control (LB). As described in Materials and Methods section, trypsin digestion of the protein extracts under different conditions generates tryptic peptides that are uniquely related to individual proteins. Tryptic peptides separated by liquid chromatography from the complex samples were first subjected to simple mass measurement (MS1) followed by intensity dependent fragmentation of these peptide ions to produce sequence specific fragment ions by collision-induced dissociation (MS/MS). Tryptic peptides were then identified using these sequence specific fragment ions via MASCOT database search software³⁴, where the sequence specific fragment ions were matched to the proteins in *S. Typhimurium* 14028 reference proteome database^{24,35}. This method of protein analysis is normally referred to as data dependent analysis (DDA). At the beginning of data analysis, the H₂O₂L and H₂O₂H data were compared to LB separately, however it turned out that comparison was not sensitive enough to differentiate between H₂O₂L and H₂O₂H conditions. Hence, the data of H₂O₂L and H₂O₂H were combined for analysis in comparison to LB. We identified 1,104 proteins of *Salmonella* for the 3 conditions (Table S5); of these, 246 proteins were differentially expressed in response to H₂O₂ with *p* values ≤ 0.05 and 90% CI. Proteomics analysis showed that 121 and 125 proteins were upregulated and downregulated in response to stress by H₂O₂, respectively. Since Tn-seq revealed genetic requirements for fitness under the selection conditions, the identified genes are expected to express corresponding proteins under the conditions to perform their cellular functions. Often the proteins required for fitness under a given condition are overexpressed under the condition, but it may not be the case for some proteins. In this study, we had a unique opportunity to comparatively analyze both Tn-seq and the MS data to understand the relationship between genetic requirements and changes in protein expression level under the condition of interest, which was H₂O₂ in this study. We also obtained the list of essential genes based on our Tn-seq data, which could not tolerate insertions by definition, and if we were not certain about essentiality of a gene from our Tn-seq data, the gene was searched for essentiality in the previously reported list of *Salmonella* essential genes²². The comprehensive list of essential genes allowed us to study any correlation between the essentiality and the changes in protein expression. Among the 246 proteins, there were 78 essential and 168 non-essential proteins. Among the 78 essential proteins, 25 were upregulated whereas 53 were downregulated. On the contrary, the majority (n = 96) of the detected non-essential proteins were upregulated, while 72 non-essential proteins were downregulated. To further examine the quantitative relationships closely, 64 genes/proteins identified by both methods (Table S5) were focused on. Among the 64 genes/proteins, 57 genes showed negative Log₂FC based on Tn-seq data, and 41 proteins among the 57 were upregulated at protein level. However, only 12 upregulated proteins (AhpC, ArcA, Crr, DksA, FliC, IcdA, OxyR, Pgm, RecA, RpoS, SlpA, and WecE) had *p* values of ≤ 0.05. The Venn diagram in Fig. S5A shows the common genes/proteins that were identified by both approaches.

Using KEGG pathway analysis^{31,32}, 150 proteins among the 246 were enriched in 21 pathways (Table S6). Interestingly, of the all 59 30S and 50S ribosomal proteins in *S. Typhimurium*, 37 of these proteins (63%) were downregulated in response to H₂O₂. Moreover, of the 8 identified proteins in TCA cycle, 6 proteins were downregulated, including 2 essential proteins.

Although DDA method can be used to search for all proteins in a complex sample, it is prone to miss identification of important proteins due to the fact that fragmentation of tryptic peptides from these proteins may not be triggered as a result of lower peptide ion intensities compared to the threshold set. To quantify proteins expressed for the genes identified by Tn-seq more accurately, we used targeted-proteomic approach by employing liquid chromatography coupled with triple quadrupole mass spectrometry (LC-QQQ-ESI-MS). Here, tryptic peptides of the protein were targeted for fragmentation (MS/MS) independent of their intensities, as described in Materials and Methods, and the observed sequence specific fragment ion intensities from three unique tryptic peptides were utilized for protein quantitation. Of the 137 Tn-seq identified genes, we selected 33 genes to quantify their proteins in response to H₂O₂ by using targeted-proteomics based on one of the multiple criteria (Table S5): (1) the protein expression was not supported by DDA, (2) the protein is a part of the pathways encoded by *aro-ABK* or *ybbKLM*, (3) the protein whose role was validated by mutant assay, or (4) the protein was identified by both Tn-seq and DDA (Fig. S5B). Interestingly, 23 (70%) of the 33 tested proteins were upregulated in response to H₂O₂. This shows a good agreement between the results of the Tn-seq and the targeted-proteomics. These genes that are conditionally required for fitness in the presence of H₂O₂ are usually expected to show upregulated

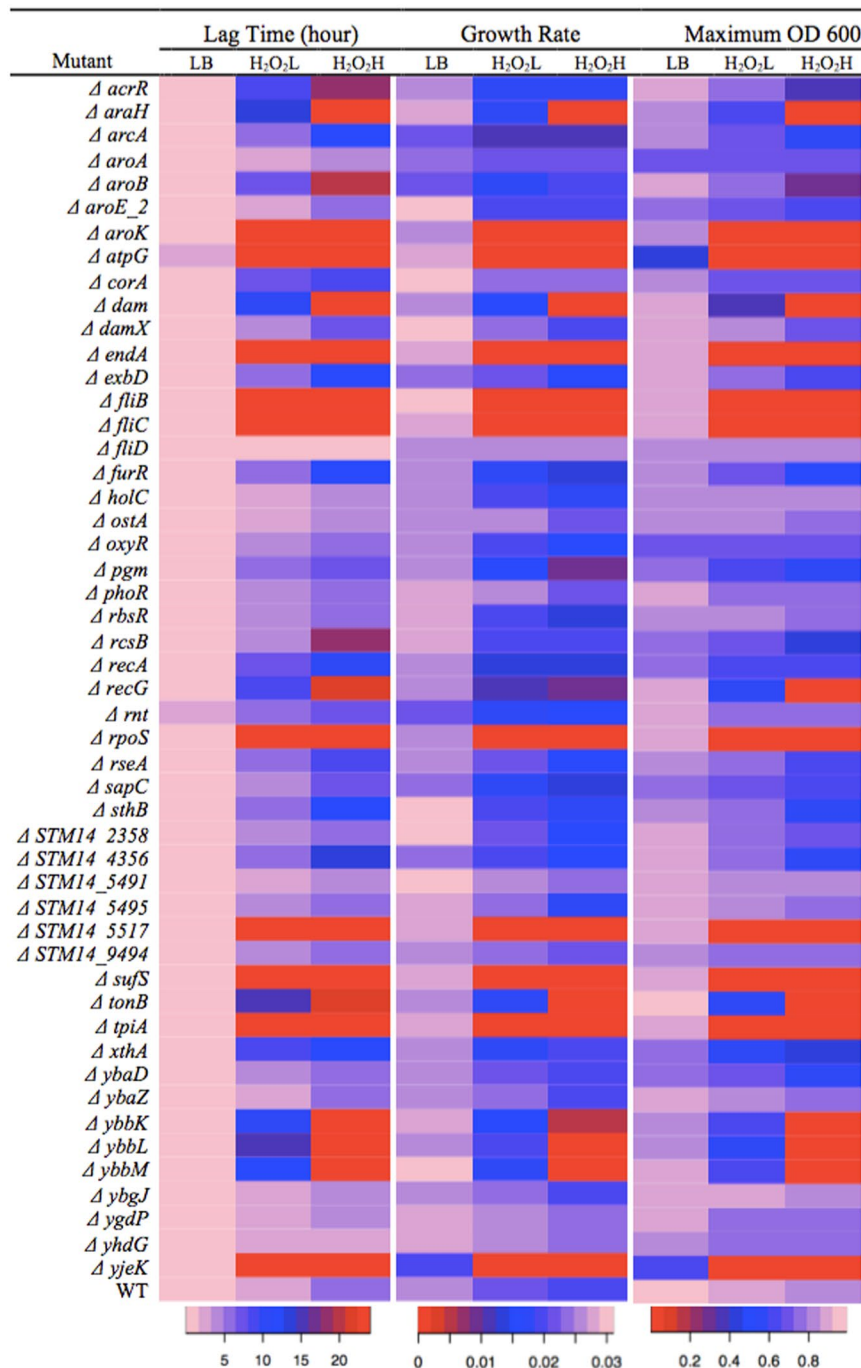


Figure 2. Growth curve of 50 mutants and a wild type *S. Typhimurium* in LB, H₂O₂L and H₂O₂H. The lag phase time, growth rate, and maximum OD₆₀₀ of the individual *S. Typhimurium* mutants and the wild type were measured in LB media supplemented with no H₂O₂ (LB), 2.5 mM H₂O₂ (H₂O₂L) or 3.5 mM H₂O₂ (H₂O₂H). Overnight culture of each of the mutants and wild type was diluted 1:200 in LB, H₂O₂L and H₂O₂H. The cultures were incubated at 37 °C for 24 h in a 96-well plate and the OD₆₀₀ was recorded every 10 min. The lag phase time, growth rate, and maximum OD₆₀₀ were calculated and shown here as a graphical representation. The pale pink color indicates a short lag phase time, a high growth rate, and a high OD₆₀₀. The red color indicates that the bacteria stayed in a lag phase, growth rate was close to a zero, and OD₆₀₀ of the culture did not increase during the 24 h time of the assays. The data of this figure can be found in Table S4.

protein expression in response to the same stimuli. However, it is also possible a protein is conditionally required while the expression is constitutive, and not regulated by H₂O₂.

Aromatic amino acid biosynthesis genes are required for H₂O₂ resistance. Interestingly, our Tn-seq data revealed that the aromatic amino acid biosynthesis and metabolism pathway play a role in conferring

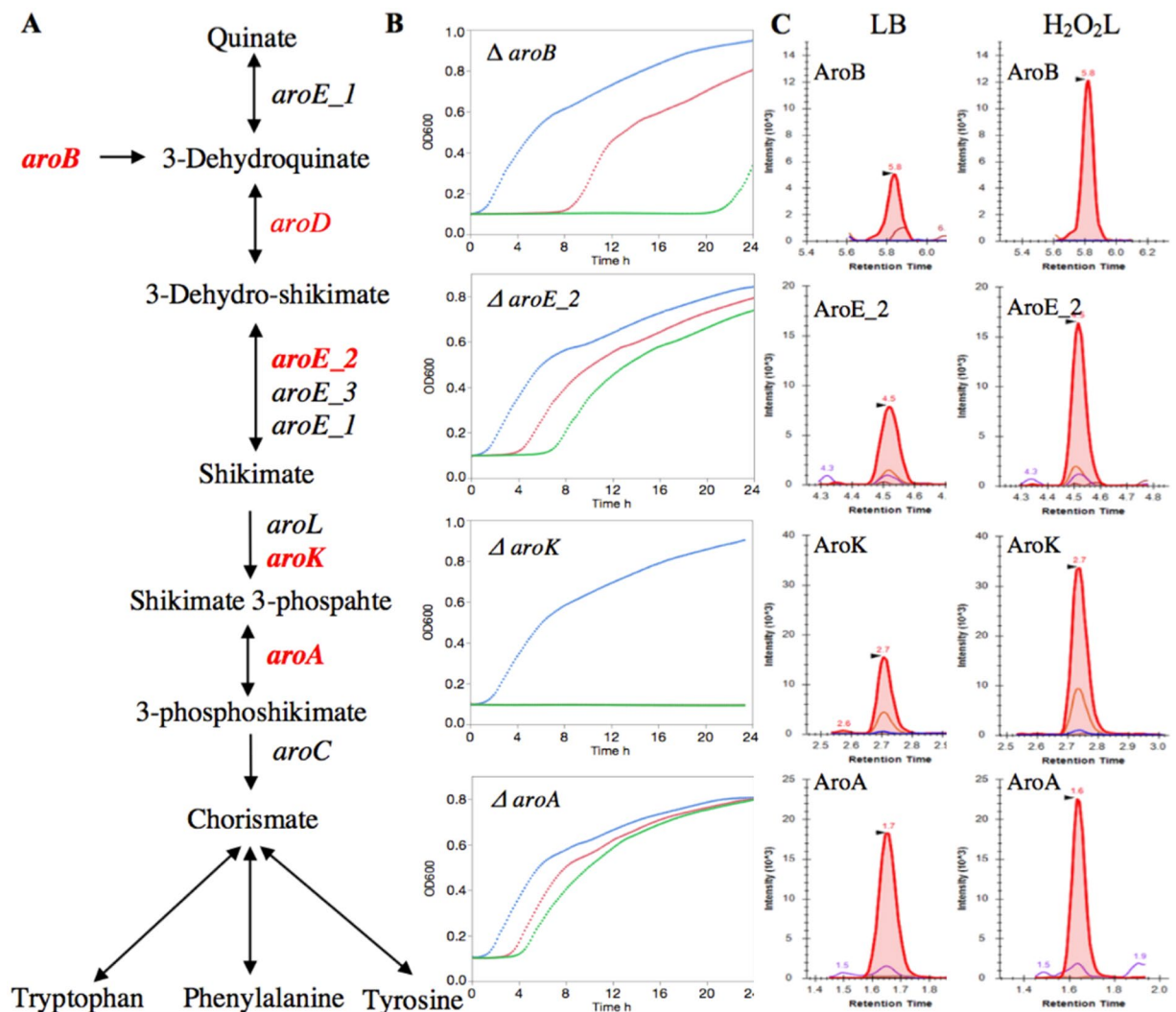


Figure 3. The role of aromatic amino acid biosynthesis genes in resistance to H_2O_2 . **(A)** Schematic representation of aromatic amino acid biosynthesis pathway, adapted from the KEGG pathway database. The genes in red color were identified by the Tn-seq analysis as conditionally essential for H_2O_2 resistance in *Salmonella*. The genes in bold type indicated that the mutant phenotype was also validated by individual mutant assay. **(B)** The overnight cultures of the individual mutants were diluted 1:200 in LB (no H_2O_2) and LB containing either 2.5 mM H_2O_2 (H_2O_2L) or 3.5 mM H_2O_2 (H_2O_2H). The cultures were incubated at 37 °C for 24 h in a 96-well plate. The growth curves are shown LB (blue), H_2O_2L (red), and H_2O_2H (green). In the growth curve of $\Delta aroK$, the red color is overlapped with the green color. **(C)** Upregulation of *Salmonella* proteins in response to H_2O_2L in comparison to LB. Wild type *S. Typhimurium* strain was grown in LB, H_2O_2L , and H_2O_2H until mid-exponential phase. Targeted-proteomics was used to quantify AroB, AroE_2, AroK, and AroA protein expressions in response to H_2O_2L .

resistance in *Salmonella* to H_2O_2 (Fig. 3A and B). Five genes, *aroB*, *aroD*, *aroE_2*, *aroK*, and *aroA* in the aromatic amino acid biosynthesis pathway were identified by Tn-seq, and the fitness of the mutants were significantly reduced in the presence of H_2O_2 . To confirm this, 4 of these genes were evaluated using individual mutant assays. The *Salmonella aroK* mutant showed the strongest phenotype, because it failed to grow in the presence of H_2O_2L or H_2O_2H during 24 h incubation time. Also, the *aroB* mutant exhibited a strong phenotype, significantly extending lag phase for both H_2O_2 conditions. The *aroE_2* mutant also exhibited an extended lag time, but the *aroA* mutant did not show any difference in growth phenotype in the presence of H_2O_2 . In addition, targeted-proteomics also showed that all these 5 proteins were upregulated in response to H_2O_2 (Fig. 3C and Table S5).

ROS damages a variety of biomolecules via Fenton reaction, which consequently lead to metabolic defects, specifically auxotrophy for some aromatic amino acids¹⁰. *E. coli* mutants that lack superoxide dismutase enzymes are unable to grow *in vitro* unless the medium are supplemented with aromatic (Phe, Trp, Tyr), branched-chain (Ile, Leu, Val), and sulfur-containing (Cys, Met) amino acids³⁶. We identified the genes in the aromatic amino acid biosynthesis pathway that are critically important for resistance to H_2O_2 . In this pathway, *aroK* catalyzes the

production of shikimate 3-phosphate from shikimate, which consequently leads to the production of tryptophan, phenylalanine, tyrosine and some metabolites from the chorismate precursor in *E. coli*. Further, *aroK* mutant in *E. coli* displays increased susceptibility to protamine, a model cationic antimicrobial peptide. It has been suggested that resistance to protamine is probably due to the aromatic metabolites and product of *aroK* gene, which act as a signal molecule to simulate the CpxR/CpxA system and Mar regulators³⁷. In our Tn-seq data, *cpxR/cpxA* and *mar-BCRT* were in the list of non-required genes, but the proteomics data indicated that CpxR was upregulated. Also, *aroK* mutant in *E. coli* is resistance to mecillinam, a beta-lactam antibiotic specific to penicillin-binding protein 2. It has been concluded that the AroK has a secondary activity in addition to the aromatic amino acid biosynthesis, probably related to cell division³⁸. In addition, *aroK* gene presents a promising target to develop a non-toxic drug in *Mycobacterium tuberculosis* because *aroK* is the only *in vitro* essential gene among the aromatic amino acid pathway genes and blocking *aroK* kills the bacterium *in vivo*³⁹. Moreover, *aroK* gene plays a general role in *S. Typhimurium* persistence in pigs⁴⁰. The *aroB* is another gene in the pathway that was identified by Tn-seq, which encodes 3-dehydroquinate synthase in the Shikimate pathway, aromatic amino acid biosynthesis pathway. *Salmonella* lacking *aroB* showed a strong growth defect in the presence of H₂O₂. When this mutant was grown in the presence of H₂O₂, it increased the lag phase time by a 114-fold for the H₂O₂L and a 347-fold for the H₂O₂H as compared to the mutant grown in the absence of H₂O₂. *S. Typhimurium* mutant lacking the *aroB* gene is attenuated in BALB/c mice⁴¹. In addition to *aroK* and *aroB*, *aroE_2* was also shown to be important for resistance to H₂O₂, because deletion of the *aroE_2* reduced the growth rate by 35% in the presence of H₂O₂ and increased the lag phase time, too. All these 3 genes in this pathway are required for systemic infection of *Salmonella* in BALB/c mice in a more recent study¹⁸. We observed that there was a strong correlation between the fitness based on Tn-seq data, growth rates measured by individual mutant assays, and upregulation of their proteins quantified via targeted proteomics. This demonstrates the power of proteogenomic approach in discovering and characterizing the genes that are required for growth under a specific condition.

The *ybbM*, *ybbK*, and *ybbL* have a role in H₂O₂ resistance. The mutants with single deletion in each of *ybbK*, *ybbL*, and *ybbM* genes on the same pathway showed a strong phenotype against the activity of H₂O₂ in a dose-dependent manner. Based on Tn-seq data, the fitness of *ybbM* was -1.16 and -1.79 for H₂O₂L and H₂O₂H, respectively (Fig. 4A). The *ybbM* mutant demonstrated decreased growth rate by 38% for H₂O₂L and 100% for the H₂O₂H as compared to the mutant grown in the absence of H₂O₂. This mutant also increased the lag time by a 126-fold and a 267-fold for H₂O₂L and H₂O₂H, respectively (Fig. 4B). Also, the fitness score of *ybbK* was -0.92 for H₂O₂L and -1.81 for H₂O₂H. The *ybbK* mutant showed decrease of growth rate for H₂O₂L and H₂O₂H by 85% and 95%, respectively. The deletion increased the lag phase by a 46-fold and a 114-fold in the presence of H₂O₂L and H₂O₂H, respectively (Fig. 4B). Moreover, the fitness of *ybbL* mutant was -1.05 and -1.73 for H₂O₂L and H₂O₂H, respectively. Deleting the *ybbL* in *Salmonella* led to decrease in growth rate by 27% for H₂O₂L and 92% for H₂O₂H. The lag phase time for this mutant increased by a 22-fold and a 33-fold for H₂O₂L and H₂O₂H, respectively. In addition, YbbM, YbbL, and YbbK proteins were upregulated in response to H₂O₂; YbbM was the most upregulated protein among the 3 proteins (1.46-fold), followed by YbbL (1.29-fold), and YbbK (1.25-fold) (Fig. 4C and Table S5). The fitness scores of the Tn-seq of these 3 genes are correlated strongly with the growth rate, lag time of their respective mutants, and upregulation of their proteins. As the number of reads depletes after the selection for a mutant, (i) there was more reduction in growth rate, (ii) the mutant stays longer in the lag phase, and (iii) the protein expression elevates. These observations clearly point to their role in conferring resistance to the H₂O₂-mediated stress. These genes were described in the *Salmonella* reference genome as follows: *ybbM*, putative YbbM family transport protein, metal resistance protein; *ybbK*, putative inner membrane proteins; *ybbL*, putative ABC transporter, ATP-binding protein YbbL. To the best of our knowledge, there is only one published study on the *ybbM* and *ybbL*⁴². Based on their findings, YbbL and YbbM have a role in iron homeostasis in *E. coli* and are important for survival when the bacterium was challenged with 10 mM H₂O₂ for 30 min; this putative ABC transporter transports iron and lessens ROS species formation that generates via H₂O₂. In this study, we identified an additional gene, *ybbK*, in the same pathway as the gene required for resistance to H₂O₂, strongly establishing the role of these 3 genes in resistance to H₂O₂.

The H₂O₂ scavenging and degrading genes. *Salmonella* employs redundant enzymes to degrade or scavenge ROS. The *katE*, *katG*, and *katN* genes encode catalases, which are involved in H₂O₂ degradation. The *ahpCF*, *tsaA*, and *tpx* genes encode peroxidases, which scavenge H₂O₂. The *sodA*, *sodB*, *sodCI*, and *sodCII* genes encode superoxide dismutases and these enzymes specifically scavenge O₂^{11,12,43–45}. However, none of these were present in the list of genes identified by Tn-Seq. Even though *katE*, *katG*, *ahpC*, *sodA*, *sodCI*, and *sodCII* showed reduced fitness, they did not meet the statistical threshold. However, the proteomics data indicated that AhpC (1.48-fold), SodB (1.46-fold), and TpX (1.39-fold) were upregulated in the presence of H₂O₂ (Table S5) and KatG was also upregulated, but its *p* value was 0.054. This reveals that these 4 proteins were important enzymes for H₂O₂ resistance under our experimental conditions. *Salmonella* containing an *ahpC* promoter-*gfp* fusion shows that expression of the *ahpC* is regulated by ROS that is generated from macrophages or exogenous H₂O₂ and the response to H₂O₂ is in a dose-dependent manner⁴⁶. *Salmonella* mutant that lacks *katE*, *katG*, or *ahpCF* can degrade micromolar concentrations of H₂O₂. However, *Salmonella* mutant that has deletions in the all 5 genes, *katE*, *katG*, *katN*, *ahpCF* and *tsaA* (HpxF), cannot degrade H₂O₂, is unable to proliferate in macrophages, and show reduced virulence in mice¹¹. This emphasizes that *ahpC*, *sodB*, and *tpx* may be the primary players in scavenging and degrading H₂O₂ in our experiment. Why Tn-seq did not detect any of these genes, while proteomics detected only these 3 proteins among others? It may reflect the functional redundancy in the genetic network that prevented single deletions in one of these genes from exhibiting fitness defect. Alternatively, when these mutants were grown together with all other mutants in the library, the functional protein lacking in one mutant due to Tn5 insertion could have been compensated by the other mutants in the library.

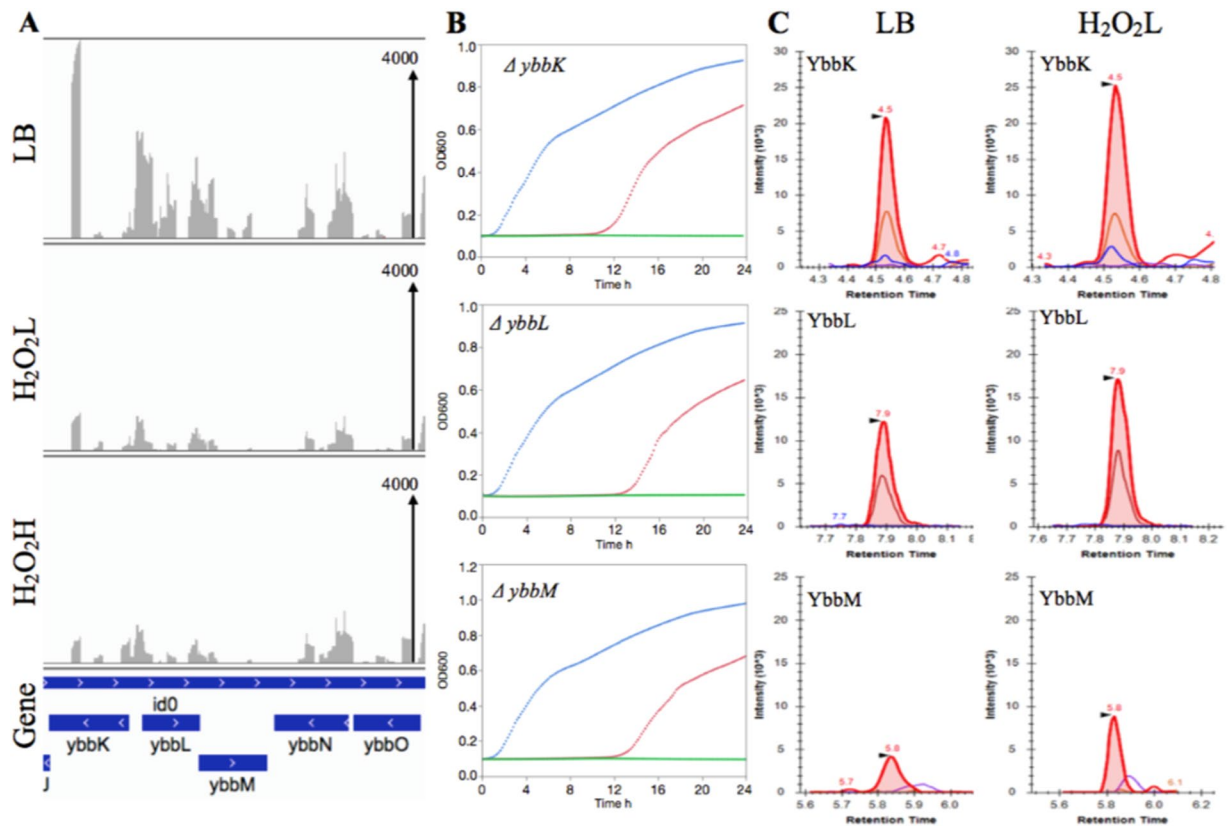


Figure 4. The *ybbK*, *ybbL*, and *ybbM* have a role in resistance to H₂O₂. (A) The Tn-seq profiles of LB (no H₂O₂), H₂O₂L (2.5 mM), and H₂O₂H (3.5 mM) corresponding to the ~6 kb genomic region ranging from *ybbK* and *ybbO*. The read scale on vertical axis is 4000. (B) The growth curve of $\Delta ybbK$, $\Delta ybbL$, and $\Delta ybbM$. The overnight cultures of these three mutants were diluted 1:200 in LB, H₂O₂L, and H₂O₂H. The cultures were incubated at 37 °C for 24 h in a 96-well plate. The growth curves are shown for LB (blue), H₂O₂L (red) and H₂O₂H (green). (C) The wild type *S. Typhimurium* strain was grown in LB, H₂O₂L, and H₂O₂H until mid-exponential phase. Targeted-proteomics was used to quantify YbbK, YbbL, and YbbM protein expressions in response to H₂O₂L.

In addition to these genes, *oxyR* was detected by Tn-seq (Fig. 1C) and DDA proteomics. The *oxyR* was identified by all 3 analysis methods of Tn-seq data and it was on the top of the list, indicating a severe fitness defect of the mutant. The *oxyR* gene encodes H₂O₂ sensor and transcription factor, which mediates protection against ROS. The *katG* and *ahpCF* are regulated by OxyR, peroxide response regulator^{13,14}. Although *Salmonella* OxyR regulon is induced in the *Salmonella*-containing vacuole in macrophage, the *oxyR* mutant was virulent in a BALB/c mouse and can grow well in human neutrophils *in vitro*^{47,48}. The fitness of *oxyR* mutant was reduced in both H₂O₂L and H₂O₂H with the respective fitness score of -4.96 and -5.94 . *Salmonella oxyR* mutant exhibited a growth rate reduction by 24% and 40% for H₂O₂L and H₂O₂H, respectively. Comparing this reduction in growth rate to the other mutants such as *rpoS* or *aroK*, we observed that the *oxyR* mutant did not show severe phenotype and the mutant escaped from the lag phase easily. Moreover, our targeted proteomics indicated that the OxyR was not upregulated significantly. Further studies are needed to uncover the exact role of OxyR in response to ROS. However, previous studies implied that OxyR plays an essential role in resistance to H₂O₂ by regulating other proteins. OxyR induces Dps in *E. coli*, a ferritin-like protein that sequesters iron⁴⁹. Sequestering of iron impairs the Fenton reaction, which consequently provides protection against ROS and reduces the damage of biomolecules. The *dps* gene was identified by the Tn-seq and its fitness score was -2.48 . However, the Dps protein was downregulated based on the DDA proteomic analysis. To confirm this unexpected finding, we conducted the proteomic assay twice and each time with at least 4 technical replicates, but the Dps protein was significantly downregulated with $p = 0.001$. Further, the targeted-proteomics demonstrated the same result, pointing to the downregulation of Dps in response to H₂O₂. This is contrary to the previously reported works on Dps in *Salmonella* and the reason for the discrepancy is unclear.

DNA repair system is important for H₂O₂ resistance. The imposed exogenous H₂O₂ activates DNA repair system in *Salmonella* in order to repair or eliminate the damage that occurred on the nucleotides. The *E. coli* RecA protein repairs double-strand DNA lesions through recombination⁵⁰. In our Tn-seq analysis, the fitness score of this mutant was -5.36 for both concentrations, and in proteomics, the RecA was upregulated (1.79-fold). *Salmonella recA* mutant decreased the maximum OD₆₀₀ by 16% for H₂O₂L and 22% for H₂O₂H as compared to the same mutant grown in LB. *Salmonella recA* mutant was also sensitive to exogenous H₂O₂ in

aerated rich medium²⁶. Moreover, *recG*, recombination and DNA repair gene⁵¹, showed a stronger phenotype than *recA* mutant. The *regG* deletion in *Salmonella* caused the growth rate reduction by 52% for H₂O₂L and 60% for H₂O₂H. This disruption in *recG* also caused the cells to stay in lag phase for a longer time in the presence of H₂O₂ as compared to LB; the lag time increased by a 62-fold and a 159-fold for H₂O₂L and H₂O₂H, respectively. In the blood of patients with *Salmonella* Typhi bacteremia, the proteins encoded by *recA*, *recG*, and *xthA* genes were detected, suggesting these proteins are actively expressed in the blood environment⁵². The XthA protein is another enzyme that participates in DNA repair mechanism induced by H₂O₂ and iron-mediated Fenton reaction. The *xthA* encodes exonuclease III, which repairs the damaged DNA. We found that the *xthA* gene was required based on the Tn-seq assay and its mutant had a reduced fitness score of -3.06 for H₂O₂L and a -4.38 for the H₂O₂H. Further, *Salmonella* lacking the *xthA* increased the lag time by 8-fold and a 12-fold for H₂O₂L and H₂O₂H, respectively. Targeted-proteomics showed upregulation of XthA (1.64-fold) in response to H₂O₂. *Salmonella enterica* serovar Enteritidis defective in *xthA* is susceptible to egg albumin⁵³. *E. coli xthA* mutant is hypersensitive to H₂O₂⁵⁴. The *xthA* is also required for *Mycobacterium tuberculosis* to infect C57BL/6J mice⁵⁵. In addition to the aforementioned genes involved in DNA repair system, *uvrA* encoding Holliday junction DNA helicase motor protein, *uvrC* encoding exonuclease ABC subunit A, *uvrD* encoding DNA-dependent helicase II, and *polA* encoding DNA polymerase I were among the top of the list of the genes identified by Tn-seq as required for resistance to H₂O₂. Collectively, DNA repair system is crucial for the survival of the *Salmonella* in a niche that contains H₂O₂.

Flagellar genes have a role for H₂O₂ resistance. Some flagellar genes, *fliC* and *fliB*, were shown to be important for resistance to H₂O₂. These two genes were identified by Tn-seq and their proteins were shown to be upregulated. *Salmonella* lacking either of these genes exhibited a strong phenotype in the presence of H₂O₂. During 24 h of incubation, *fliC* and *fliB* mutants could not grow in both H₂O₂ conditions. However, growth of *fliD* mutant was not affected by H₂O₂. The *fliC* was shown to have a role in *Salmonella* Typhi interaction with human macrophages and *Salmonella* Typhimurium *fliB* mutant was defective in swarming motility^{56,57}. Currently it remains unclear how flagella genes can be involved in the resistance of *Salmonella* to oxidative stress, which warrants future study into this direction.

Fe-S cluster biogenesis system is required for H₂O₂ resistance. *Salmonella* requires the genes from Fe-S cluster biogenesis system in order to resist H₂O₂. Our Tn-seq analysis identified 5 genes in this system that are required for the resistance. In *isc* operon (Fe-S cluster), *iscA*, *hscB*, and *hscA* were among the genes required to resist H₂O₂. Particularly, the *hscA* is on the top of the gene list identified by Tn-seq. In *E. coli*, this operon is regulated by IscR, iron sulfur cluster regulator⁵⁸; in *Salmonella* the gene *iscR* encoding this transcription regulator is named *yfhP*. The HscB and HscA chaperones are believed to be involved in the maturation of Fe-S proteins^{59,60}. The second operon that is involved in Fe-S protein biogenesis is the *suf*, sulfur mobilization operon. Tn-seq found that two genes in this operon were required for *Salmonella* to resist H₂O₂; *sufS* and *sufC*. *Salmonella* mutant lacking *sufS* exhibited a strong phenotype in the presence of H₂O₂ and could not grow during the 24 h of incubation. The SufS with SufE in *E. coli* form a heterodimeric cysteine desulphurase and SufB, SufC, and SufD form a pseudo-ABC-transporter that could act as a scaffold⁶⁰; this operon is regulated by OxyR¹⁴. The other known genes in these two operons that are present in *Salmonella* are *iscA*, *sufA*, *sufB*, and *sufD*; they showed a reduced fitness, while their *p* values were >0.05. The damage of Fe-S clusters is not only problem for the defective proteins, but also it fuels the Fenton reaction via the released iron and H₂O₂¹⁰. Thus, *Salmonella* uses Fe-S cluster repair system as an arsenal to overcome the damage imposed by H₂O₂.

DNA adenine methylase is important for H₂O₂ resistance. DNA adenine methylase genes, *dam* and *damX*, are involved in *Salmonella* resistance against H₂O₂. Our Tn-seq data showed that fitness of *dam* and *damX* mutants were reduced in the presence of H₂O₂. To confirm this, *Salmonella dam* mutant was grown in both conditions. Under H₂O₂L, the growth rate was reduced by 42% as compared to the mutant in LB and the mutant could not grow under H₂O₂H during the 24 h of incubation. In addition, the lag time of the *Salmonella dam* mutant was extended by a 19-fold for H₂O₂L. While the *Salmonella damX* mutant displayed a moderate phenotype as compared to the *dam* mutant, the *damX* mutant also showed that the growth rate decreased by 23% and 33% for H₂O₂L and H₂O₂H, respectively. The lag time was extended for this mutant by a 25-fold and a 49-fold for H₂O₂L and H₂O₂H, respectively. The *dam* regulates virulence gene expression in *S. Typhimurium*⁶¹. The different levels of Dam affects virulence gene expression, motility, flagellar synthesis, and bile resistance in the pathogenic *S. Typhimurium* 14028⁶². Dam methylation activates the gene that are involved in lipopolysaccharide synthesis⁶³. Moreover, *Salmonella* defective in *damX* is very sensitive to bile⁶⁴. Collectively, our study demonstrates the critical role of DNA adenine methylase in *Salmonella* resistance against H₂O₂.

Other genes for H₂O₂ resistance. Beside the important pathways described above, there were many additional genes also important for resistance to H₂O₂. Among those, the 3 unrelated genes, *rpoS*, *pgm*, and *tonB*, are important ones that deserve more attention. The *rpoS* mutant showed reduced fitness and its protein was upregulated in the presence of H₂O₂. *Salmonella* mutant defective in *rpoS* showed a strong phenotype when grown in the presence of H₂O₂. The *rpoS* encodes the alternative sigma factor σ^s , a subunit of RNA polymerase; it is the master regulator of stress response⁶⁵. This implies that *rpoS* is an important component of the genetic regulatory network that *Salmonella* employs in order to resist H₂O₂. Furthermore, the fitness of *pgm* mutant was reduced and its protein was upregulated in the presence of H₂O₂. Knock out of *pgm* in *Salmonella* caused a decrease in growth rate, increase the lag phase time, and reduce the maximum OD₆₀₀ in the presence of H₂O₂. The *pgm* encodes phosphoglucosyltransferase which are required for catalysis of the interconversion of glucose 1-phosphate and glucose 6-phosphate⁶⁶. This gene contributes to resistance against antimicrobial peptides, is required for *in vivo* fitness in the mouse model, and participates in LPS biosynthesis⁶⁷. Lastly, the fitness of *tonB* mutant was also reduced.

Salmonella lacking *tonB* exhibited a strong phenotype in the presence of H₂O₂ as compared to the mutant grown in LB. The gene mediates iron uptake in the *Salmonella*⁴⁵. In addition, seven of the genes identified in our study (*proC*, *arcA*, *barA*, *exbD*, *flhD*, *fliC*, and *fliD*) were previously shown to be important for interaction of *Salmonella* Typhi with human macrophages⁵⁶.

In summary, we applied Tn-seq and proteomic analysis to find the genes and proteins that are required in *S. Typhimurium* to resist H₂O₂ *in vitro*. As the concentration of H₂O₂ increased, the growth rate reduced, the lag time extended, the fitness of mutants decreased, and some proteins were differentially expressed. Validation of Tn-seq results with individual mutant assays indicated the accuracy of the identified genes in response to the two H₂O₂ concentrations. The targeted-proteomics had a good agreement with Tn-seq. We found many genes that have not been associated to resistance to H₂O₂ previously and these genes will be focus of our future research. *Salmonella* employs multiple pathways to resist H₂O₂ and the most important ones are ROS detoxifying enzymes, amino acid biosynthesis (*aroK* and *aroB*), putative iron transporters (*ybbK*, *ybbL*, *ybbM*), iron homeostasis, Fe-S cluster repair, DNA repair, flagellar and DNA adenine methylase genes. Our study is focused on single gene knockout mutants and their phenotypes, and therefore was not designed nor expected to reveal any synthetically essential genes in the presence of H₂O₂. However, Tn-seq can be also used for genetic interaction mapping as previously demonstrated¹⁹. We expect that genetic interaction mapping study in the future based on Tn-seq would reveal more complete picture of H₂O₂ resistance. The genes identified in this study will broaden our understanding on the mechanisms used by *Salmonella* to survive and persist against ROS in macrophages.

Our unbiased system-wide approach, Tn-seq, was successful in identifying novel genetic determinants that have not been implicated previously in *Salmonella* resistance to oxidative stress. Based on our literature review, the number of genes that are involved in resistance to oxidative stress is over 60 across different bacterial species, and our findings have doubled this number (n = 137) of conditionally required genes for H₂O₂ resistance. Furthermore, the combined use of quantitative proteomic approach has provided additional insights on the function or mode of action of the identified genetic determinants in resisting oxidative stress. As expected, the majority of the proteins important for resistance to H₂O₂ were upregulated in response to the same stressor. However, the expression level did not increase for some proteins, in spite of their known roles in resistance to H₂O₂. Interestingly, the downregulation of Dps and other proteins was counterintuitive to the common mode of protein regulation and function, yet it may point to some unknown aspects of how *Salmonella* regulates the expression of those proteins to better cope with the oxidative stress during infection in macrophage. The genes identified in this study will broaden our understanding on the mechanisms used by *Salmonella* to survive and persist against ROS in macrophages.

Methods

Bacterial strain and construction of Tn5 mutant library. *Salmonella* enterica subsp. enterica serovar Typhimurium ATCC 14028 s (with spontaneous mutation conferring resistance to nalidixic acid (NA) was used in this study. All procedures involving this pathogen (Biosafety level 2) were conducted according to the protocol approved by Institutional Biosafety Committee (IBC). We mutagenized *S. Typhimurium* ATCC 14028s by biparental mating using *Escherichia coli* SM10 λ pir carrying a transposon-delivery plasmid vector pBAM1 (Ampicillin (Amp) resistance) as a donor strain⁶⁸. The plasmid pBAM1 was generously provided by Victor de Lorenzo (Systems and Synthetic Biology Program, Centro Nacional de Biotecnología, Madrid, Spain). The donor strain, *E. coli* Sm10 λ pir (pBAM1), was grown overnight in LB with 50 μ g/ml Amp and recipient strain was grown in LB with 50 μ g/ml NA at 37 °C. Equal volumes (1 ml) of donor and recipient were mixed, centrifuged, washed in 10 mM MgSO₄, and re-suspended in 2 ml PBS (pH 7.4). Then, the mating mixture was concentrated and laid on a 0.45- μ m nitrocellulose filters (Millipore). The filter was incubated at 37 °C on the surface of LB agar plate. After 5 h of conjugation, the cells on the filter was washed in 10 mM MgSO₄, and resuspended in 1 ml MgSO₄. The conjugation mixture was plated on LB agar containing 50 μ g/ml NA and 50 μ g/ml kanamycin (Km). After approximately 24 h at 37 °C, we scraped the colonies into LB broth containing 50 μ g/ml Km and 7% DMSO. The yield was approximately 68,000 individual colonies from each conjugation. Five independent conjugations were conducted to yield approximately 325,000 mutants. The library was stored at -80 °C in aliquots. To determine the frequency of the mutants that have been produced by integration of the entire delivery plasmid, the colonies were picked randomly and streaked on LB plates (Km) and LB plates (Km and Amp). It was shown that ~20% of the cells in the library were resistant to Amp, indicating a significant portion of the Km-resistant colonies was not from authentic transposition events.

Measuring growth responses of *S. Typhimurium* to H₂O₂. To determine the effect of H₂O₂ concentrations on growth parameters, overnight culture of the wild type *S. Typhimurium* 14028 s was inoculated into fresh LB broth media (1:200 dilution) to give a seeding concentration corresponding to OD₆₀₀ of ~0.1. The LB broth contained freshly prepared H₂O₂ to give the final concentrations ranging from 0.05 to 10 mM. The cultures were directly added into 96-well microplate (200 μ l/well). The microplate was incubated in a Tecan Infinite 200 microplate reader at 37 °C, with shaking duration 5 s, shaking amplitude 1.5 mm, and reading OD₆₀₀ every 10 min. The number of replicates were at least three. The lag time, growth rate, and maximum OD₆₀₀ were calculated using GrowthRates script⁶⁹. Growth Rate % decrease was calculated as follows⁷⁰: Growth Rate % decrease = ((μ_{PC} - μ_S) / μ_{PC}) \times 100; where μ = the maximum slope (growth rate), μ_{PC} = growth rate of positive control (without H₂O₂), μ_S = growth rate in the presence of H₂O₂.

Selection of the mutant library for Tn-seq analysis. The transposon library was thawed at room temperature and diluted 10⁻¹ in fresh LB broth. To activate the library, the diluted library was incubated at 37 °C with shaking at 225 rpm for an hour. Then, the culture was washed twice with PBS and resuspended in LB broth medium. The library was inoculated to 20 ml LB broth and LB broth supplemented with either 2.5 or 3.5 mM

H₂O₂ (H₂O₂L and H₂O₂H, respectively), seeding CFU was 3.5×10^6 per ml. Then, when the cultures reached mid-exponential phase, OD₆₀₀ of 2.7 ($\sim 1.17 \times 10^8$ CFU/ml), the incubation was stopped, and the culture was immediately harvested by centrifugation, and stored at -20°C .

Preparation of Tn-seq amplicon libraries. Genomic DNA was extracted from the harvested cells using DNeasy Blood & Tissue kit (Qiagen), and quantified using Qubit dsDNA RB Assay kit (Invitrogen). As described above, 20% of the mutants in the library were the result of the integration of pBAM1 into chromosome. To remove the Tn5-junction sequences originated from the plasmid in the Tn-seq amplicon libraries, genomic DNA was digested with PvuII-HF (New England Biolabs), which digests immediately outside the inverted repeats on both sides of Tn5 in pBAM1, and purified with DNA Clean & Concentrator-5 kit (Zymo Research). Then, a linear PCR extension was performed using a Tn5-specific primer in order to produce single stranded DNA corresponding to Tn5-junction sequences. To increase the specificity in extending into Tn5-junction sequences, the linear PCR was conducted with a dual priming oligonucleotide Tn5-DPO (5'-AAGCTTGCATGCCTGCAGGTIIIICTAGAGGATC-3') that is specific to Tn5 end⁷¹. The PCR reaction contained 25 μl Go Taq Colorless Master Mix (Promega), 20 μM Tn5-DPO primer, 100 ng gDNA, and 50 μl MQ-H₂O. The PCR cycle consisted of the initial denaturation at 95°C for 2 min, followed by 50 cycles at 95°C for 30 sec, 62°C for 45 sec, and 72°C for 10 sec. The PCR product was purified with DNA Clean & Concentrator-5 kit and eluted in 13 μl TE buffer. After that, C-tail was added to the 3' end of the single-stranded DNA. The C-tailing reaction was consisted of 2 μl terminal transferase (TdT) buffer (New England Biolabs), 2 μl CoCl₂, 2.4 μl 10 mM dCTP, 1 μl 1 mM ddCTP, 0.5 μl TdT and 13 μl purified linear PCR product. The reaction was performed at 37°C for 1 h and the enzyme was inactivated by incubation at 70°C for 10 min. The C-tailed product was purified with DNA Clean & Concentrator-5 kit and eluted in 12 μl TE. Next, the exponential PCR was performed with forward primer, P5-BRX-TN5-MEO, 5'-AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTNNNNAG-6 nt barcode-CCTAGCGGCCTTAATTAAAGATGTGTATAAGAG-3' and reverse primer, P7-16G, 5'-CAAGCAGAAGACGGCATAACGAGCTCTCCGATCTGGGGGGGGGGGGGGGGGG-3' to attach Illumina adapter sequences along with the sample barcodes. The PCR reaction contained 25 μl Go Taq Green Master Mix, 10 μM P5-BRX-TN5-MEO primer, 10 μM P7-16G primer, 1 μl purified C-tailed genomic junctions, and MQ-H₂O to 50 μl ; the PCR condition started with initial denaturation at 95°C for 2 min, followed by 36 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 20 sec, with the final extension at 72°C for 5 min. Then, the size selection of the DNA was performed using agarose gel electrophoresis. The 50 μl PCR products were incubated at 60°C for 15 min and incubated on ice for 5 min, and immediately loaded on the 1% agarose gel in 0.5% TAE buffer. After running the gel, the DNA fragment of size 325–625 bp was cut and put in a microtube for each sample. The DNA was extracted from the gel using Zymoclean Gel DNA Recovery kit (Zymo Research). The prepared DNA libraries were quantified using Qubit dsDNA RB Assay kit. Since each library has its own barcode, the libraries were combined and sequenced on a flow cell of HiSeq. 3000 using single end read and 151 cycles (Illumina) at the Center for Genome Research & Biocomputing in Oregon State University.

Tn-seq data analysis. The preliminary data analysis was conducted by using a super computer in the High Performance Computing Center (AHPCC) at the University of Arkansas. The libraries that were multiplexed for sequencing were de-multiplexed using a custom Python script. The script searched for the six-nucleotide barcode for each library for perfect matches. In order to extract the transposon genomic junctions, we used Tn-Seq Pre-Processor (TPP) tool³⁰ with some modifications in the script. The TPP searched for the 19 nucleotide inverted repeat (IR) sequence and identified five nucleotides (GACAG) at the end of the IR sequence, allowing one nucleotide mismatch. The Tn5-junctions that start immediately after GACAG were extracted and the C-tails at the end of junctions were removed. Tn5-junction sequences less than 20 nucleotides were discarded and remaining Tn5-junctions were mapped to the *Salmonella enterica* serovar Typhimurium 14028 s genome and plasmid using BWA-0.7.12⁷². To identify genes that are required for H₂O₂ resistance, the following three Tn-seq analysis tools were used for comparative analysis: (i) ARTIST²⁸: the genomic junctions were mapped to the reference genome using Bowtie 2.2.7⁷³. The number of insertions and reads were determined for genes and intergenic regions. The data were normalized with default script in the ARTIST. Then, the relative abundance of Tn5 insertions in the output library versus the input were calculated. Later, the *p* values were calculated from a 100 independent Mann-Whitney U test (MWU) analysis that were carried out on input and output data for each gene. Finally, the genes were considered conditionally essential if the *p* values were ≤ 0.05 in the 90 of the 100 MWU tests. (ii) Tn-seq Explorer²⁹: The output SAM files from the TTP were used as input to the Tn-seq Explorer. The unique insertions with less than 20 reads were removed from the input and outputs. Using the window size of 550 and excluding 5% of beginning of genes and 20% of the end of genes, Essentiality Index (EI), number of unique insertions, and total number of reads per gene were counted. The EI of more than 10 were removed from the input. Genes with less than 300 nucleotides were removed. Differential EI were calculated from input and outputs ($\Delta\text{IE} = \text{output EI} - \text{input EI}$) and genes with ΔIE more than -1 were removed. Log₂ fold change of reads were calculated from input and output ($\text{Log}_2\text{FC} = \log_2(\text{output reads}/\text{input reads})$) and the genes were ranked based on the Log₂FC value from least value to highest. The genes with Log₂FC value of less than -2 from the H₂O₂H and present in H₂O₂L were considered conditionally essential. (iii) TRANSIT³⁰: The output wig files from the TTP was used as input data file for TRANSIT. The comparative analysis was conducted with Tn5 resampling option. The reads were normalized with trimmed total reads (TTR). Insertions outside the 5% and 10% sequences from 5'- and 3'- ends were removed, respectively. The genes were considered conditionally essential if *p* values ≤ 0.05 .

Phenotypic evaluation of individual deletion mutants. The mutants were obtained from *Salmonella enterica* subsp. *enterica*, 14028 s (Serovar Typhimurium) Single-Gene Deletion Mutant Library through BEI

Resources (www.beiresources.org). The overnight cultures of *S. Typhimurium* mutants were added into fresh LB broth media containing freshly prepared H₂O₂ (2.5, or 3.5 mM/ml) (1:200 dilution) to give seeding OD₆₀₀ of 0.1. The cultures were directly added into 96-well microplates and incubated in Tecan Infinite 200 at 37 °C for 24 h. The lag time, growth rate, and maximum OD₆₀₀ were calculated using GrowthRates⁶⁹.

Sample preparation for proteomics and mass spectrometry analysis. The overnight culture of the wild type *S. Typhimurium* 14028 was diluted 1:200 in 50 ml LB medium, and LB containing either 2.5 or 3.5 mM H₂O₂ in a 300-ml flask. The cultures were grown until mid-exponential phase (OD₆₀₀ of 2.7), and the 50 ml volume of cultures were used for a total protein extraction by using Qproteome Bacterial Protein Prep kit (Qiagen). Disulfide bonds within proteins were reduced with 2-Mercaptoethanol and separated by SDS-PAGE gel electrophoresis. For each condition, there were three lanes with approximately 300 µg of proteins. The gel then was stained via coomassie blue dye. The gel portions of 3 lanes for each condition were cut out and chopped into small pieces, pooled together, washed twice with 50 mM NH₄HCO₃, destained with NH₄HCO₃/ 50% Acetonitrile (ACN), and dried with pure ACN. Then, the proteins were reduced using 10 mM Dithiothreitol in 50 mM NH₄CO₃ and the alkylation was conducted with 10 mg/ml Iodoacetamide Acid in 50 mM NH₄CO₃. After that, the proteins were washed with NH₄HCO₃, and dried with pure ACN. Mass spectrometry grade Trypsin gold from Promega (~20 ng/µl in 50 mM NH₄HCO₃) was added to dried gels, and left it overnight for efficient in-gel digestion of the proteins at 37 °C. During the digestion, tryptic peptides diffused out into the solution. Gel pieces then were extracted three times by 50% CAN/0.1% TFA solution and incubated at 37 °C for 15 minutes. Later, these digests were analyzed by ESI-LC-MS/MS at State Wide Mass Spectrometry Facility, University of Arkansas at Fayetteville. Data dependent analysis (DDA) for the in-gel trypsin digested samples from each condition were performed by using an Agilent 1200 series micro flow HPLC in line with Bruker Amazon-SL quadrupole ion trap ESI mass spectrometer (QIT-ESI-MS). All the ESI-MS analyses were performed in a positive ion mode using Bruker captive electrospray source with a dry nitrogen gas temperature of 200 °C, with nitrogen flow rate of 3 L/minute. LC-MS/MS data were acquired in the Auto MS(n) mode with optimized trapping condition for the ions at m/z 1000. MS scans were performed in the enhanced scanning mode (8100 m/z/second), while the collision-induced dissociation or the MS/MS fragmentation scans were performed automatically for top ten precursor ions with a set threshold for one minute in the UltraScan mode (32,500 m/z/second). Tryptic peptides were separated by reverse-phase high-performance liquid chromatography (RP-HPLC) using a Zorbax SB C18 column, (150 × 0.3 mm, 3.5 µm particle size, 300 Å pore size, Agilent Technologies), with a solvent flow rate of 4 µL/minute, and a gradient of 5%–38% consisting of 0.1% FA (solvent A) and ACN (solvent B) over a time period of 320 minutes. Tryptic peptides were then identified by searching MS/MS data in *S. Typhimurium* 14028 reference proteome database^{24,35} by using MASCOT database search software³⁴. MS1 intensities of the integrated areas of these identified tryptic peptides were compiled and grouped in skyline software according to the replicates/conditions to perform statistical analysis, there were 4 technical replicates. Targeted protein work were performed using Shimadzu UPLC-20A coupled to 8050 triple quadrupole ESI-MS with heated probe. Sequence specific fragment ion intensities from at least three unique tryptic peptides from the protein of interest were used in the protein quantitation. Multiple reaction monitoring (MRM) events corresponding to sequence specific fragment ions derived from the precursor tryptic peptides were targeted to operate at a certain specific retention time intervals predicted by in house retention time library. This library was generated using the correlation of relative hydrophobicity of the tryptic peptides with their retention times (RT) from highly common housekeeping proteins, for the UPLC method used in this analysis as described below. While the RT were correlated well within 99% confidence, sufficient number of sequence specific fragment ions were used as basis for identification of the tryptic peptide by MS/MS alone. Specificity and the confidence was achieved by incorporating RT prediction. In addition to the application of skyline in quantitation, skyline software was also used in predicting RT and optimizing parameters such as collision energies and voltages with the help of Shimadzu Labsolution software. Tryptic peptides were separated by reverse-phase ultra-high-performance liquid chromatography (RP-UPLC) compatible Shimadzu C18, 1.9-micron particle size, 50 × 2.1 mm column (SN # 16041880 T), with a solvent flow rate of 0.3 mL/minute, and a gradient of 5%–90% consisting of 0.1% FA (solvent A) and 0.1% FA in ACN (solvent B) over a time period of 10 minutes.

Data availability. Tn-seq sequencing data are available on NCBI Sequence Read Archive. LB: PRJNA352537; H₂O₂L: PRJNA352862; H₂O₂H: PRJNA352865. Proteomics data are available on <https://panoramaweb.org/lab-key/cMOPwl.url>.

References

- LaRock, D. L., Chaudhary, A. & Miller, S. I. Salmonellae interactions with host processes. *Nat. Rev. Microbiol.* **13**, 191–205 (2015).
- Scallan, E. *et al.* Foodborne illness acquired in the United States – major pathogens. *Emerg. Infect. Dis.* **17** (2011).
- Majowicz, S. E. *et al.* The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin. Infect. Dis.* **50**, 882–889 (2010).
- Haraga, A., Ohlson, M. B. & Miller, S. I. Salmonellae interplay with host cells. *Nat. Rev. Microbiol.* **6**, 53–66 (2008).
- Nauseef, W. Assembly of the phagocyte NADPH oxidase. *Histochem. Cell Biol.* **122**, 277–291 (2004).
- Seaver, L. C. & Imlay, J. A. Hydrogen peroxide fluxes and compartmentalization inside growing. *Escherichia coli*. *J. Bacteriol.* **183**, 7182–7189 (2001).
- Park, S., You, X. & Imlay, J. A. Substantial DNA damage from submicromolar intracellular hydrogen peroxide detected in Hpx-mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **102**, 9317–9322 (2005).
- Imlay, J. A. & Linn, S. DNA damage and oxygen radical toxicity. *Science* **240**, 1302–1309 (1988).
- Imlay, J. A. Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol.* **59**, 1073–1082 (2006).
- Slauch, J. M. How does the oxidative burst of macrophages kill bacteria? Still an open question. *Mol. Microbiol.* **80**, 580–583 (2011).
- Hébrard, M., Julie, P. M. V., Méresse, S., Barras, F. & Aussel, L. Redundant Hydrogen Peroxide Scavengers Contribute to *Salmonella* Virulence and Oxidative Stress Resistance. *J. Bacteriol.* **191**, 4605–4614 (2009).
- Horst, S. A. *et al.* Thiol Peroxidase Protects *Salmonella enterica* from Hydrogen Peroxide Stress *In Vitro* and Facilitates Intracellular Growth. *J. Bacteriol.* **192**, 2929–2932 (2010).

13. Morgan, R. W., Christman, M. F., Jacobson, F. S., Storz, G. & Ames, B. N. Hydrogen Peroxide-Inducible Proteins in *Salmonella* Typhimurium Overlap with Heat Shock and Other Stress Proteins. *Proc. Natl. Acad. Sci. USA* **83**, 8059–8063 (1986).
14. Zheng, M. *et al.* DNA Microarray-Mediated Transcriptional Profiling of the *Escherichia coli* Response to Hydrogen Peroxide. *J. Bacteriol.* **183**, 4562–4570 (2001).
15. Kroger, C. *et al.* An Infection-Relevant Transcriptomic Compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host Microbe* **14**, 683–695 (2013).
16. Turner, K. H., Wessel, A. K., Palmer, G. C., Murray, J. L. & Whiteley, M. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc. Natl. Acad. Sci. USA* **112**, 4110 (2015).
17. Chan, K., Kim, C. C. & Falkow, S. Microarray-Based Detection of *Salmonella enterica* Serovar Typhimurium Transposon Mutants That Cannot Survive in Macrophages and Mice. *Infect. Immun.* **73**, 5438–5449 (2005).
18. Silva-Valenzuela, C. *et al.* Analysis of Two Complementary Single-Gene Deletion Mutant Libraries of *Salmonella* Typhimurium in Intraperitoneal Infection of BALB/c Mice. *Front. Microbiol.* **6** (2016).
19. van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* **6**, 767–772 (2009).
20. Khatiwara, A. *et al.* Genome Scanning for Conditionally Essential Genes in *Salmonella enterica* Serotype Typhimurium. *Appl. Environ. Microbiol.* **78**, 3098–3107 (2012).
21. Chaudhuri, R. *et al.* Comprehensive Assignment of Roles for *Salmonella* Typhimurium Genes in Intestinal Colonization of Food-Producing Animals. *PLoS Genet.* **9**, e1003456 (2013).
22. Langridge, G. C. *et al.* Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* **19**, 2308–2316 (2009).
23. Barquist, L. *et al.* A comparison of dense transposon insertion libraries in the *Salmonella* serovars Typhi and Typhimurium. *Nucleic Acids Res.* **41**, 4549–4564 (2013).
24. Domon, B. & Aebersold, R. Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* **28**, 710–721 (2010).
25. Jensen, O. N. & Mann, M. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**, 255–261 (2003).
26. Bogomolnaya, L. *et al.* The ABC-Type Efflux Pump MacAB Protects *Salmonella enterica* serovar Typhimurium from Oxidative Stress. *Mbio* **4**, e00630 (2013).
27. Kwon, Y. M., Ricke, S. C. & Mandal, R. K. Transposon sequencing: methods and expanding applications. *Appl. Microbiol. Biotechnol.* **100**, 31–43 (2016).
28. Pritchard, J. R. *et al.* ARTIST: high-resolution genome-wide assessment of fitness using transposon-insertion sequencing. *PLoS Genet.* **10**, e1004782–e1004782 (2014).
29. Solaimanpour, S., Sarmiento, F. & Mrazek, J. Tn-Seq Explorer: A Tool for Analysis of High-Throughput Sequencing Data of Transposon Mutant Libraries. *PLoS One* **10**, e0126070 (2015).
30. DeJesus, M. A., Ambadipudi, C., Baker, R., Sasseti, C. & Ioerger, T. R. TRANSIT—A Software Tool for HimarI TnSeq Analysis. *PLoS Comput. Biol.* **11**, e1004401–e1004401 (2015).
31. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
32. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
33. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, (966–968 (2010)).
34. Perkins, D., Pappin, D., Creasy, D. & Cottrell, J. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
35. Cravatt, B. F., Yates, J. R. III & Simon, G. M. The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991–1000 (2007).
36. Carliz, A. & Touati, D. Isolation of superoxide dismutase mutants in *Escherichia coli*: is superoxide dismutase necessary for aerobic life? *EMBO J.* **5**, 623–630 (1986).
37. Weatherspoon-Griffin, N., Yang, D., Kong, W., Hua, Z. & Shi, Y. The CpxR/CpxA Two-component Regulatory System Up-regulates the Multidrug Resistance Cascade to Facilitate *Escherichia coli* Resistance to a Model Antimicrobial Peptide. *J. Biol. Chem.* **289**, 32571–32582 (2014).
38. Vinella, D., Gagny, B., Joseleau-Petit, D., D’Ari, R. & Cashel, M. Mecillinam resistance in *Escherichia coli* is conferred by loss of a second activity of the AroK protein. *J. Bacteriol.* **178**, 3818–3828 (1996).
39. Kaur, P. *et al.* Unravelling the Secrets of Mycobacterial Cidalty through the Lens of Antisense. *PLoS One* **11**, e0154513 (2016).
40. Van Parys, A. *et al.* Tissue-Specific *Salmonella* Typhimurium Gene Expression during Persistence in Pigs. *PLoS One* **6**, e24120 (2011).
41. Günel-Özcan, A., Brown, K. A., Allen, A. G. & Maskell, D. J. *Salmonella* Typhimurium *aroB* mutants are attenuated in BALB/c mice. *Microb. Pathog.* **23**, 311–316 (1997).
42. Nicolaou, S. A., Fast, A. G., Nakamaru-Ogiso, E. & Papoutsakis, E. T. Overexpression of *fetA* (*ybbL*) and *fetB* (*ybbM*), Encoding an Iron Exporter, Enhances Resistance to Oxidative Stress in *Escherichia coli*. *Appl. Environ. Microbiol.* **79**, 7210–7219 (2013).
43. Robbe-Saule, V., Coynault, C., Ibanez-Ruiz, M., Hermant, D. & Norel, F. Identification of a non-haem catalase in *Salmonella* and its regulation by RpoS (σ S). *Mol. Microbiol.* **39**, 1533–1545 (2001).
44. Craig, M. & Schlauch, J. Phagocytic Superoxide Specifically Damages an Extracytoplasmic Target to Inhibit or Kill *Salmonella*. *PLoS One* **4**, e4975 (2009).
45. Tsolis, R. M., Bäuml, A. J., Heffron, F. & Stojiljkovic, I. Contribution of TonB- and Feo-mediated iron uptake to growth of *Salmonella* Typhimurium in the mouse. *Infect. Immun.* **64**, 4549–4556 (1996).
46. Aussel, L. *et al.* *Salmonella* detoxifying enzymes are sufficient to cope with the host oxidative burst. *Mol. Microbiol.* **80**, 628–640 (2011).
47. Taylor, P. D., Inchley, C. J. & Gallagher, M. P. The *Salmonella* Typhimurium AhpC Polypeptide Is Not Essential for Virulence in BALB/c Mice but Is Recognized as an Antigen during Infection. *Infect. Immun.* **66**, 3208–3217 (1998).
48. Papp-Szabó, E., Firtel, M. & Josephy, P. D. Comparison of the sensitivities of *Salmonella* Typhimurium *oxyR* and *katG* mutants to killing by human neutrophils. *Infect. Immun.* **62**, 2662–2668 (1994).
49. Imlay, J. Diagnosing oxidative stress in bacteria: not as easy as you might think. *Curr. Opin. Microbiol.* **24**, 124–131 (2015).
50. Carlsson, J. & Carpenter, V. S. The *recA+* gene product is more important than catalase and superoxide dismutase in protecting *Escherichia coli* against hydrogen peroxide toxicity. *J. Bacteriol.* **142**, 319–321 (1980).
51. Ishioka, K., Iwasaki, H. & Shinagawa, H. Roles of the *recG* gene product of *Escherichia coli* in recombination repair: Effects of the Delta *recG* mutation on cell division and chromosome partition. *Genes Genet. Syst.* **72**, 91–99 (1997).
52. Sheikh, A. *et al.* In Vivo Expression of *Salmonella enterica* Serotype Typhi Genes in the Blood of Patients with Typhoid Fever in Bangladesh. *PLoS Negl. Trop. Dis.* **5**, e1419 (2011).
53. Lu, S., Killoran, P. B. & Riley, L. W. Association of *Salmonella enterica* Serovar Enteritidis YafD with Resistance to Chicken Egg Albumen. *Infect. Immun.* **71**, 6734–6741 (2003).
54. Dempfle, B., Halbros, J. & Linn, S. *Escherichia coli xth* mutants are hypersensitive to hydrogen peroxide. *J. Bacteriol.* **153**, 1079–1082 (1983).

55. Sasseti, C. M. & Rubin, E. J. Genetic Requirements for Mycobacterial Survival during Infection. *Proc. Natl. Acad. Sci. USA* **100**, 12989–12994 (2003).
56. Sabbagh, S., Lepage, C., McClelland, M. & Daigle, F. Selection of *Salmonella enterica* Serovar Typhi Genes Involved during Interaction with Human Macrophages by Screening of a Transposon Mutant Library. *PLoS One* **7**, e36643 (2012).
57. Bogomolnaya, L. *et al.* Identification of Novel Factors Involved in Modulating Motility of *Salmonella enterica* Serotype Typhimurium. *PLoS One* **9**, e111513 (2014).
58. Schwartz, C. J. *et al.* IscR, An Fe-S Cluster-Containing Transcription Factor, Represses Expression of *Escherichia coli* Genes Encoding Fe-S Cluster Assembly Proteins. *Proc. Natl. Acad. Sci. USA* **98**, 14895–14900 (2001).
59. Unciuleac, M. *et al.* Vitro activation of apo-aconitase using a [4Fe-4S] cluster-loaded form of the IscU [Fe-S] cluster scaffolding protein. *Biochemistry (N. Y.)* **46**, 6812–6821 (2007).
60. Py, B. & Barras, F. Building Fe-S proteins: bacterial strategies. *Nat. Rev. Microbiol.* **8**, 436–446 (2010).
61. Balbontin, R. *et al.* DNA Adenine Methylation Regulates Virulence Gene Expression in *Salmonella enterica* Serovar Typhimurium. *J. Bacteriol.* **188**, 8160–8168 (2006).
62. Badie, G., Heithoff, D. M., Sinsheimer, R. L. & Mahan, M. J. Altered Levels of *Salmonella* DNA Adenine Methylase Are Associated with Defects in Gene Expression, Motility, Flagellar Synthesis, and Bile Resistance in the Pathogenic Strain 14028 but Not in the Laboratory Strain LT2. *J. Bacteriol.* **189**, 1556–1564 (2007).
63. Sarnacki, S. H. *et al.* Dam Methylation Controls O-Antigen Chain Length in *Salmonella enterica* Serovar Enteritidis by Regulating the Expression of Wzz Protein. *J. Bacteriol.* **191**, 6694–6700 (2009).
64. Lopez-Garrido, J., Cheng, N., Garcia-Quintanilla, F., Garcia-del Portillo, F. & Casadesus, J. Identification of the *Salmonella enterica damX* Gene Product, an Inner Membrane Protein Involved in Bile Resistance. *J. Bacteriol.* **192**, 893–895 (2010).
65. Hengge-Aronis, R. Signal Transduction and Regulatory Mechanisms Involved in Control of the σ S (RpoS) Subunit of RNA Polymerase. *Microbiol. Mol. Biol. Rev.* **66**, 373–395 (2002).
66. Lu, M. & Kleckner, N. Molecular cloning and characterization of the *pgm* gene encoding phosphoglucomutase of *Escherichia coli*. *J. Bacteriol.* **176**, 5847–5851 (1994).
67. Paterson, G. K., Cone, D. B., Peters, S. E. & Maskell, D. J. The enzyme phosphoglucomutase (*Pgm*) is required by *Salmonella enterica* serovar Typhimurium for O-antigen production, resistance to antimicrobial peptides and *in vivo* fitness. *Microbiology* **155**, 3403–3410 (2009).
68. Martínez-García, E., Calles, B., Arévalo-Rodríguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol.* **11**, 38–38 (2011).
69. Hall, B., Acar, H., Nandipati, A. & Barlow, M. Growth Rates Made Easy. *Mol. Biol. Evol.* **31**, 232–238 (2014).
70. Yao, Q. *et al.* Assembly of nanoions via electrostatic interactions: ion-like behavior of charged noble metal nanoclusters. *Sci. Rep.* **4**, 3848–3848 (2014).
71. Chun, J. *et al.* Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *Nucleic Acids Res.* **35**, e40–e40 (2007).
72. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
73. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25–R25 (2009).

Acknowledgements

We thank Jack Jiang, a high school student, for writing the demultiplexing Python script. We are thankful for Dr. Thomas R. Ioerger and Dr. Michael A. DeJesus at Texas A&M University for the help on the use of TRNASIT. We extend our special thanks to Dr. Jeff F. Pummill and Dr. Pawel Wolinski at Arkansas High Performance Computing Center for their help to use the facility. This work was partially supported by Arkansas Biosciences Institute. This research was also supported by the Arkansas High Performance Computing Center which is funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission. The first author was supported by his parents, Cell and Molecular Biology (CEMB) program at the University of Arkansas, and Human Capacity Development Program-Kurdistan Regional Government (HCDP-KRG).

Author Contributions

Conceived and designed the experiments: Y.K., S.K. Performed the experiments, analyzed the data, wrote the manuscript: S.K. Proteomics work: R.L., S.K., A.Q., J.L. Revised the manuscript: Y.K., S.K. All authors read the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-17149-9>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017