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

Sardar Karash[1], Rohana Liyanage[2,3], Abdullah Qassab[1], Jackson O. Lay, Jr.[2,3], and Young Min Kwon[1,4*]

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

[2]Department of Chemistry, University of Arkansas, Fayetteville, AR 72701,

[3]Statewide Mass Spectrometry Facility, Fayetteville, AR 72701, and

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

*Corresponding author:

Department of Poultry Science

College of Agricultural, Food and Life Sciences

University of Arkansas

Fayetteville, AR 72701

Phone : (479) 575-4935

Email ykwon@uark.edu

ABSTRACT

Salmonella is a Gram-negative bacterium that infects a wide range of hosts. *Salmonella* Typhimurium causes gastroenteritis in human, and can survive and replicate in macrophages. An essential mechanism used by the macrophages to eradicate *Salmonella* is production of reactive oxygen species. Here, we used proteogenomic approaches to determine the candidate genes and proteins that have a role in resistance of *S.* Typhimurium to H_2O_2. For Tn-seq, a highly saturated Tn5 insertion library was grown *in vitro* in Luria-Bertani broth (LB) medium as well as LB containing either 2.5 (H_2O_2L) or 3.5 mM H_2O_2 (H_2O_2H). We identified two sets of overlapping genes that are required for resistance of *S*. Typhimurium to H_2O_2L and H_2O_2H, and the result was validated via phenotypic evaluation of 50 selected mutants. The enriched pathways for resistance to H_2O_2 included DNA repair, aromatic amino acid biosynthesis, Fe-S cluster biosynthesis, iron homeostasis, flagellar genes, H_2O_2 scavenging enzymes, and DNA adenine methylase. Particularly, we identified aromatic amino acid biosynthesis (*aroB*, and *aroK*) and putative iron transporter system (*ybbK*, *ybbL*, and *ybbM*) as novel mechanisms for resistance to H_2O_2. The proteomics revealed that the majority of essential proteins, including ribosomal proteins, were downregulated upon exposure to H_2O_2. A subset of the proteins identified by Tn-seq were analyzed by targeted proteomics, and 70 % of them were upregulated upon exposure to H_2O_2. The identified candidate genes will deepen our understanding on the mechanisms of *S*. Typhimurium survival in macrophages, and can be exploited to develop new antimicrobial drugs.

45 **IMPORTANCE**

46 *Salmonella* infection is frequently caused by consumption of contaminated food or water. The
47 infection may lead to gastroenteritis or typhoid fever, depending on the *Salmonella* serovars. Even
48 though the bacterium encounters the immune defense arsenals of the infected host, including
49 reactive oxygen species in phagocytes, the bacterium can survive and replicate. In this study,
50 proteogenomic approaches were used in order to identify the genes and proteins that have a role
51 in resistance to H_2O_2. In addition to the H_2O_2 scavenging and degrading enzymes, aromatic amino
52 acid biosynthesis and iron homeostasis were identified among the most important pathways for
53 H_2O_2 resistance. These findings will deepen our knowledge on the mechanisms of *Salmonella*
54 survival in phagocytes and other niches with oxidative stress, and also provides novel targets to
55 develop new antimicrobial therapeutics.

56
57
58
59
60
61
62
63
64
65
66
67
68

69 **Introduction**

70 *Salmonella* is a Gram-negative bacterium that infects humans and animals. *Salmonella enterica*
71 has numerous serovars, which include typhoidal and non-typhoidal strains. In contrast to the
72 typhoidal salmonellae which are human restricted pathogens, the non-typhoidal salmonellae
73 (NTS), serovar Enteritidis and Typhimurium, are able to infect a wide range of hosts, causing
74 gastroenteritis (1). The NTS strains, including *Salmonella enterica* serovar Typhimurium, account
75 for 11% (1.2 million cases) of the total foodborne illnesses caused by different pathogens in the
76 United States (2). It has been estimated that *Salmonella* is responsible for 93.8 million cases of
77 gastroenteritis, leading to 155,000 deaths worldwide annually (3). The pathogen remains a
78 continuous threat to the food safety, and public health.

79 To initiate an infection and survive inside the host, *Salmonella* needs to overcome a myriad of host
80 defense mechanisms. As *Salmonella* reaches the intestine and breaches the epithelial tissue, it
81 enters the macrophages and activates different virulence strategies in order to survive and replicate
82 in them (4). An essential mechanism used by the phagocytes to kill and eradicate *Salmonella* is
83 production of reactive oxygen species (ROS). Hydrogen peroxide (H_2O_2), superoxide anion ($O_2^{\cdot -}$
84), and the hydroxyl radical ($HO^{\cdot}$) are derivatives of ROS. The short-lived $O_2^{\cdot -}$, produced by the
85 NADPH-dependent phagocytic oxidase, quickly dismutates into H_2O_2, which diffuses across
86 semipermeable bacterial cell membranes. Eventually, H_2O_2 reduces to $HO^{\cdot}$ by Fe^{2+} via Fenton
87 reaction (5-7). The ROS, including H_2O_2, can damage DNA, iron-sulfur cluster-containing
88 proteins, and other biological molecules in the bacterial cells (8-10).

89 Numerous genetic factors and proteins that are important for resistance of *S.* Typhimurium to H_2O_2
90 have been discovered and the underlying mechanisms have been explored (11,12). Various
91 approaches and techniques have been employed to study global response of *Salmonella* or related

92 bacteria to H_2O_2 *in vitro* as a model system to simulate the bacterium's response to ROS in
93 phagocytic cells: (i) Two-dimensional gel electrophoresis identified H_2O_2-induced proteins in
94 *Salmonella* (13), (ii) DNA microarray identified H_2O_2 induced genes in *E. coli* (14), and (iii) RNA-
95 seq identified H_2O_2 induced genes in *Salmonella* (15). However, the factors required for fitness
96 under the given condition cannot be identified with high confidence based on the analysis of
97 transcriptomics or proteomics data (16). Microarray-based tracking of random transposon
98 insertions was used to identify numerous genes in *Salmonella* that are required for survival in mice
99 and macrophages (17-18). However, the genetic factors responsible for resistance to ROS cannot
100 be sorted out among all of the genetic factors identified in the study that are required for fitness in
101 the presence of multiple host stressors.
102 To shed more insights into the underlying mechanisms of *Salmonella* resistance to H_2O_2, more
103 direct approach linking the gene-phenotype relationships in a genome-wide scale would be
104 necessary. Tn-seq is a powerful approach to allow direct and accurate assessment of the fitness
105 requirement of each gene on the entire genome of a prokaryotic organism (19). In Tn-seq method,
106 a saturated transposon insertion library (input) is exposed to a selective condition, and the mutant
107 population altered through the selection (output) is recovered. Then, the genomic junctions of the
108 transposon insertions are specifically amplified and sequenced from both input and output pools
109 by high-throughput sequencing. The gene fitness can be obtained by calculating the change in
110 relative abundance of the sequence reads corresponding to each gene in the entire genome between
111 the two pools. Tn-seq has been employed to assign gene functions to *Salmonella* genomes in
112 numerous studies: (i) Previously, our lab identified conditionally essential genes that are required
113 for growth in the presence of bile, limited nutrients, and high temperature (20), (ii) The genes
114 required for intestinal colonization were identified in chickens, pigs, and cattle (21), (iii) Candidate

115 essential genes and genes contributing toward bile resistance were identified (22), (iv) Core
116 conserved genes for growth in rich media were identified in serovars Typhi and Typhimurium
117 (23). In addition to Tn-seq, electrospray ionization liquid chromatography tandem mass
118 spectrometry (ESI-LC-MS/MS) is a powerful approach in identifying and quantifying proteins in
119 a large scale. The system-wide protein regulation can be determined using mass spectrometry
120 signal intensities of tryptic peptides obtained from two different culture conditions (24). The post-
121 translational modification in proteins can be revealed by using proteomic analysis (25). Many
122 studies took advantage of proteomic analysis of *Salmonella*. However, to the best of our
123 knowledge, this study is the first to investigate proteogenomics of a bacterium by combining Tn-
124 seq and proteome analysis simultaneously to the same stressor.
125 In this work, we used Tn-seq method and proteomic analysis in combination to determine system-
126 wide responses of *S.* Typhimurium to two different concentrations of H_2O_2 (H_2O_2L and H_2O_2H).
127 We obtained a comprehensive list of 137 genes that are putatively required for the resistance of *S.*
128 Typhimurium 14028 to H_2O_2. The role of 50 selected genes in resistance to H_2O_2 were determined
129 by phenotypic evaluation of the individual deletion mutants. Also, we identified a set of 246
130 proteins that are differentially expressed in response to H_2O_2, using data-dependent acquisition
131 (DDA) proteomics, which are largely overlapped with the genes identified by Tn-seq; targeted
132 proteomics showed 70% of the proteins identified by Tn-seq were upregulated by H_2O_2. In addition
133 to the genes of *S.* Typhimurium previously known to be important for resistance to H_2O_2, we
134 identified approximately 80 genes that have not been previously associated with resistance to
135 oxidative stress. The result of this study highlighted that the genes in aromatic amino acid
136 biosynthesis, *aroB* and *aroK*, and iron homeostasis, *ybbK*, *ybbL*, and *ybbM*, are crucially important
137 for growth fitness under H_2O_2 stress. The identified candidate genes will expand our understanding

138 on the molecular mechanisms of *Salmonella* survival in macrophages, and serve as new
139 antimicrobial drug targets.
140
141 **Results and Discussion**
142 **The H_2O_2 concentrations and the selections of Tn5 library**
143 First, we sought to determine the growth response of the wild type *S.* Typhimurium 14028 cells in
144 LB media containing varying concentrations of H_2O_2. The wild type cells were grown in LB media
145 that contain different concentrations of H_2O_2 in 96-well plates. After evaluating the growth rates
146 for the cultures, 2.5 and 3.5 mM H_2O_2 were chosen for Tn-seq selections in our study, and termed
147 H_2O_2L and H_2O_2H, respectively. In comparison to *Salmonella* grown in LB media with no H_2O_2,
148 H_2O_2L and H_2O_2H reduced the growth rates by 10% and 28%, respectively (Fig. 1A). The lag time
149 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,
150 respectively. The maximum OD_{600} decreased by only 1% for the H_2O_2L and 2% for the H_2O_2H in
151 comparison to LB media (Fig. 1A).
152 For the selection of Tn5 library, 20 ml cultures in 300 ml Erlenmeyer flasks containing LB, H_2O_2L,
153 or H_2O_2H were inoculated with the same Tn5 library at the seeding CFUs of the library at 3.5×10^6.
154 This seeding level provided ~10 CFUs for each Tn5 insertion mutant in the library. The cultures
155 were grown until the mid-exponential phase, in which the CFUs reached 1.17×10^8 (SE 0.01×10^8).
156 It required 7.5 and 9.2 h to reach the cell density as measured by optical density for H_2O_2L and
157 H_2O_2H, respectively, in contrast to 5 h for LB medium (Fig. 1B). We observed some differences
158 in growth responses between the cultures in a 96-well plate and in a 300-ml Erlenmeyer flask. The
159 optical density readings by the plate reader was different in comparison to those by Bio-photometer
160 that we used to measure optical density of the culture in the flask. As a result, the growth curve in

161 Fig.1.A which was based on 96-well plate reader, dose not match exactly with the time required
162 for the Tn5 library to reach the target mid-exponential phase in the flask cultures. In addition, we
163 observed that the H_2O_2 is stable in LB media free of *Salmonella* during the window of time used
164 for the library selection process (Fig.S1), which was also supported by Bogomolnaya et al. (26).

165 **Preparation of Tn5-seq amplicon library**

166 The *Salmonella* mutants were generated by using the delivery plasmid pBAM1 via conjugation. A
167 total of 325,000 mutant colonies were recovered from 50 plates. Each mutant contained a single
168 random insertion of Tn5 transposon in the chromosome or plasmid according to DNA sequencing
169 of Tn5-junction sequences for a small set (n = 71) of randomly selected Tn5 mutants. We found a
170 significant portion (~20%) of the mutants in the library that were not genuine Tn5 insertions, but
171 the mutants generated as a result of pBAM1 integration into chromosome as determined by their
172 ability to grow in the presence of ampicillin. To prevent the Illumina sequencing reads from being
173 wasted on sequencing Tn5 junctions from these cointegrants, we digested genomic DNA of the
174 input and output libraries with PvuII, which digests immediately outside the inverted repeats on
175 both sides of Tn5. The digested DNA was then used to prepare Tn-seq amplicon library as
176 described in Materials and Methods. Our Tn-seq data analysis indicated that our strategy of
177 removing the DNA sequences originating from cointegrants was effective because only 0.55% of
178 the total HiSeq reads corresponding to Tn5-junctions matched to pBAM1. It should be possible to
179 remove them completely by ensuring complete digestion of genomic DNA with PvuII. The method
180 for Tn-seq amplicon library we developed and used in this study has multiple advantages over
181 other Tn-seq protocols, because our method requires only 100 ng of the genomic DNA, and the
182 whole process can be completed in a day (27). When the extension step in the protocol was
183 performed using a conventional 20 nucleotide primer, and the final products of exponential PCR

184 were separated on agarose gel electrophoresis, even the negative controls (the wild type genomic
185 DNA or mutant library genomic DNA without linear extension) showed smear patterns of
186 nonspecific background amplification. However, when dual priming oligonucleotide (DPO)
187 primer was used in place of the conventional primer for linear extension, non-specific background
188 amplification was completely disappeared. Therefore, we adopted the DPO primer in linear
189 extension step for all library samples in this study. Then, the single-stranded extension products
190 were C-tailed, and used as templates for the exponential PCR step using nested primer specific to
191 Tn5 and poly G primer that contain Illumina adapter sequences along with sample index sequences
192 (Fig. S2). The final PCR products were separated on an agarose gel, and the fragments within the
193 range of 325-625 bp were gel-purified. After pooling of multiple samples, the combined library
194 was sequenced on a HiSeq 3000.

195 **Summary of Tn-seq DNA analysis**

196 After de-multiplexing and C-tail trimming of all sequence reads, ~72 million reads of Tn5-
197 junctions with mean read length of 94 bp were obtained. The number of the reads mapped to the
198 complete genome of *S.* Typhimurium 14028 were ~25, 15, and 19 million for LB, H_2O_2L, and
199 H_2O_2H, respectively. The number of unique insertions on the chromosome were 125,449 in the
200 input library, excluding the plasmid (Table S1). On average, Tn5 was inserted in every 39 bp.
201 Number of raw reads per open reading frame (ORF) for H_2O_2L was plotted over the corresponding
202 number of H_2O_2H, which yielded an R^2 of 0.91, indicating the mutants in the input library
203 quantitatively responded in a similar way for both H_2O_2L and H_2O_2H as expected (Fig. S3). The
204 insertions were mapped to 5,428 genes or 8,022 genes/intergenic regions. Interestingly, the ORF
205 STM14_5121, which is 16.7 kbp long, had the highest number of insertions (~700 insertions) and
206 reads (0.25 M).

207 **Comparison of various bioinformatics pipelines for Tn-seq data analysis**

208 We used 3 different Tn-seq analysis tools to identify the genes and compare the results across the
209 methods with the goal of comprehensive identification of "all" genes required for resistance to
210 H_2O_2. The first tool, ARTIST (28), created small non-overlapping genomic windows of 100 bp
211 and the reads from each window were arbitrarily assigned into the middle of the window. The
212 default normalization script of the tool was used. Then, the relative proportions of insertion sites
213 in the output library versus the input were tabulated. Mann-Whiney *U* (MWU) test was used to
214 assess the essentiality of the locus. To consider a gene/intergenic region conditionally essential for
215 growth in the presence of H_2O_2, p value had to be ≤ 0.05 in 90 of the 100 conducted MWU tests.
216 Subsequently, 20 genes and 1 intergenic region were identified for H_2O_2L and 4 genes for H_2O_2H
217 (Data Set S1). We speculate the reason that more genes were identified for H_2O_2L in comparison
218 to H_2O_2H, was partially due to the lower number of total reads of H_2O_2L as compared to H_2O_2H,
219 even though the read numbers of H_2O_2L was normalized to those of the input.

220 The second tool, Tn-seq Explorer (29), counted insertions in overlapping windows of a fixed size.
221 Using a 550 bp window size, each annotated gene was assigned an essentiality index (EI) which
222 is determined mainly based on the insertion count in a window in this gene. The bimodal
223 distribution of insertion counts per window divided the essential genes to the left and the non-
224 essential genes to the right. To find conditional essential genes, the EI of the output was subtracted
225 from the EI of the input. The genes with negative ΔEI were ranked based on the change in read
226 fold change (Log2 (H_2O_2L or H_2O_2H/Input)). We found 114 consensus genes between H_2O_2L and
227 H_2O_2H that had at least four-fold reduction in H_2O_2H read counts as compared to the input. The
228 four-fold reduction (Log2FC = -2) threshold was chosen based on our validation study of Tn-seq
229 data by single mutant assays (Data Set S1).

230 The third tool, TRANSIT (30), determined read counts of genes in the input and output library.
231 The differences of total read counts between the input and outputs were obtained. The insertion
232 sites were permutated for a number that is specified by the user (we used 10,000 sample). This
233 sampling for each gene gave difference in read counts. The *p* value was calculated from the null
234 distribution of the difference in read counts. We identified 8 and 21 genes for the H_2O_2L and
235 H_2O_2H, respectively, using a *p* value ≤ 0.05 (Data Set S1). For normalization, trimmed total reads
236 (TTR) was chosen for this study
237 The combined list of the genes identified by the 3 Tn-seq analysis tools for both H_2O_2L and H_2O_2H
238 included 137 genes (Data Set S1). All of the genes on this list are expected to have a role in
239 conferring resistance to H_2O_2 and allow *Salmonella* to survive and replicate in the presence of
240 H_2O_2 *in vitro*. Of the 21 genes identified by TRANSIT, 19 of these genes were also identified by
241 Tn-seq Explorer, but only 3 out of this 21 were identified by ARTIST. The 19 genes were *hscA*,
242 *rbsR*, *fepD*, *efp*, *oxyR*, *polA*, *ybaD*, *aroD*, *ruvA*, *xthA*, *dps*, *aroB*, *uvrD*, *tonB*, *uvrA*, *aroK*, *ybbM*,
243 *lon*, and *proC*. Two genes, *fepD* and *xthA*, were identified by the all 3 methods and for both
244 conditions.
245 The 3 Tn-seq analysis tools are very valuable for Tn5 data analysis, but each tool has its own
246 advantages and disadvantages. For ARTIST, (i) the user must know how to run scripts in Matlab
247 software, (ii) the analysis is very slow on a personal computer with the HiSeq data, (iii) it has only
248 one method for normalization, but (iv) it can search for essentiality in the intergenic regions. For
249 Tn-seq Explorer, (i) there is no data normalization, and (ii) prediction on small genes is prone to
250 be inaccurate, but (iii) its very user-friendly and runs fast. For TRANSIT, (i) the user should have
251 some knowledge on running scripts on terminal, (ii) it may need some modification in its Python
252 script according to the way the library was prepared for sequencing, and (iii) a few software

253 packages should be installed on the computer as TRANSIT pre-requisites, but (iv) it does have 6
254 different methods for data normalization and it runs very fast on a personal computer. Although
255 ARTIST and Tn-seq Explorer are very useful tools for Tn-seq data analysis, we prefer using
256 TRANSIT in our future data analysis for conditionally essential genes. In the following sections,
257 we continued the downstream analysis mainly based on the 137 genes that include all of the genes
258 identified by all 3 methods.

259 **The enriched pathways for resistance to H_2O_2**

260 In order to categorize the identified genes that are required for *Salmonella* resistance to the H_2O_2,
261 the 137 genes were subjected to pathway enrichment analysis using DAVID Bioinformatics
262 Resources 6.7, NIAID/NIH (31). A total of 15 KEGG pathways (32) were recognized for 69 genes
263 on the list. The enriched pathways include homologous recombination (*ruvC*, *polA*, *ruvA*, *ruvB*,
264 *priB*, *recA*, *recR*, *holC*, *holD*, *recC*, *recG*), nucleotide excision repair (*uvrD*, *polA*, *uvrA*, *uvrC*),
265 mismatch repair (*dam*, *uvrD*, *holC*, *holD*), RNA degradation (*pnp*, *hfq*, *ygdP*), purine and
266 pyrimidine metabolism (*apaH*, *polA*, *pnp*, *arcC*, *spoT*, *holC*, *holD*, *cmk*, *dcd*, *pnp*), phenylalanine,
267 tyrosine and tryptophan biosynthesis (*aroD*, *aroB*, aroA, *aroK*, *aroE_2*), arginine and proline
268 metabolism (*proC*, *arcC*, *arcA*), starch and sucrose metabolism (*otsA*, *crr*, *pgm*), glycolysis and
269 gluconeogenesis (*crr*, *aceE*, *pgm*, *tpiA*), oxidative phosphorylation (*atpG*, *atpA*, *cydA*), DNA
270 replication (*polA*, *holC*, *holD*), flagellar assembly (*fliJ*, *fliD*, *flhD*, *fliC*), ABC transporter (*sapC*,
271 *yejE*, *yjgP*, *fepC*, *fepD*, *cbiM*, *cysP*), and two component system (*rcsB*, *rcsA*, *fliC*, *arcA*, *arcB*,
272 *phoR*, *barA*, *flhD*). Since KEGG was not able to recognize many genes on the list, we used
273 SP_PIR_Keywords of functional categories, which recognized majority of the genes and
274 categorized them into 55 functional categories (Table S2), excluding 15 uncharacterized genes
275 (ORFs). Among these categories were stress response (*rpoE*, *lon*, *dnaJ*, *hfq*, *yaiB*), iron (*dps*, *entD*,

276 *iscA*, *yjeB*, *yhgI*), and transcription regulation (*rcsA*, *oxyR*, *rpoE*, *yjeB*, *arcA*, *argR*, *rbsR*, *rpoS*,
277 *fadR*, *rcsB*, *furR*, *flhD*).

278 **Validation of Tn-seq results using individual mutants**

279 For the selected 50 genes among the 137 genes identified by Tn-seq, the growth phenotype was
280 determined using individual single deletion mutants in LB, H_2O_2L, and H_2O_2H. The genes were
281 considered to play a role in resistance to H_2O_2, if (i) lag phase time increased, (ii) growth rate
282 reduced or (iii) maximum OD_{600} decreased in the presence of H_2O_2 in comparison to the wild type
283 strain grown in the same conditions. Of the 50 single deletion mutants, 42 mutants were shown to
284 have a role in resistance to H_2O_2 (Fig. 2 and Data Set S2). One gene, *yhaD*, was identified by all
285 3 analysis tools, but it did not show the expected phenotype. The *fliD* was also identified by
286 ARTIST, but did not show any phenotype distinguishable from the wild type. The remaining 6
287 genes that did not show the phenotype was identified by Tn-seq Explorer. Based on the results of
288 the individual mutant assay, we conclude that 84% (42/50) of the genes identified by the Tn-seq
289 analysis and tested using single deletion mutants have a role for resistance to H_2O_2. These results
290 indicate that our Tn-seq analysis identified the genes in *S.* Typhimurium that are required for the
291 wild type level resistance to H_2O_2 with high accuracy.

292 **Proteomics of H_2O_2 response**

293 With ESI-LC-MS/MS in data-dependent acquisition (DDA) mode, the protein regulation was
294 determined using MS1 filtering technique that skyline software offers (33). It uses signal
295 intensities of tryptic peptides derived from the proteins of the wild type strain grown in the
296 presence of H_2O_2 in comparison to the control (LB). As described in Materials and Methods
297 section, trypsin digestion of the protein extracts under different conditions generates tryptic
298 peptides that are uniquely related to individual proteins. Tryptic peptides separated by liquid

299 chromatography from the complex samples were first subjected to simple mass measurement
300 (MS1) followed by intensity dependent fragmentation of these peptide ions to produce sequence
301 specific fragment ions by collision-induced dissociation (MS/MS). Tryptic peptides were then
302 identified using these sequence specific fragment ions via MASCOT database search software
303 (34), where the sequence specific fragment ions were matched to the proteins in *S.* Typhimurium
304 14028S reference proteome database (24, 35). This method of protein analysis is normally referred
305 to as data dependent analysis (DDA) method. At the beginning of data analysis, the H_2O_2L and
306 H_2O_2H data were compared to LB separately, however it turned out that comparison was not
307 sensitive enough to differentiate between H_2O_2L and H_2O_2H conditions. Hence, the data of H_2O_2L
308 and H_2O_2H were combined in the analysis in comparison to LB. We identified 1,104 proteins of
309 *Salmonella* for the 3 conditions (Data Set S3); of these, 246 proteins were differentially expressed
310 in response to H_2O_2 with p values ≤ 0.05 and 90% CI. The upregulated proteins were 121 and the
311 downregulated proteins were 125 in response to H_2O_2. Since Tn-seq revealed genetic requirements
312 for fitness under the selection conditions, the identified genes are expected to express their proteins
313 under the conditions to perform their cellular functions. Often the proteins required for fitness
314 under a given condition are overexpressed under the condition, but it may not be the case for some
315 proteins. In this study, we had a unique opportunity to comparatively analyze both Tn-seq and the
316 MS data to understand the relationship between genetic requirements and changes in expression
317 level under the condition of interest, which was H_2O_2 in this study. We also obtained the list of
318 essential genes based on our Tn-seq data, which could not tolerate insertions by definition, and if
319 we were not certain about essentiality of a gene from our Tn-seq data, the gene was searched for
320 essentiality in the previously reported list of *Salmonella* essential genes (22). The comprehensive
321 list of essential genes allowed us to study any correlation between the essentiality and the changes

322 in protein expression. Among the 246 proteins, there were 78 essential and 168 non-essential
323 proteins. Among the 78 essential proteins, 25 were upregulated whereas 53 were downregulated.
324 On the contrary, the majority (n = 96) of the detected non-essential proteins were upregulated,
325 while 72 non-essential proteins were downregulated. To further examine the quantitative
326 relationships closely, 64 genes/proteins identified by both methods (Data Set S3) were focused on.
327 Among the 64 genes/proteins, 57 genes showed negative Log2FC based on Tn-seq data, and 41
328 proteins among the 57 were upregulated at protein level. However, only 12 proteins had *p* values
329 of ≤ 0.05 (AhpC, ArcA, Crr, DksA, FliC, IcdA, OxyR, Pgm, RecA, RpoS, SlpA, and WecE).
330 Using KEGG pathway analysis, 150 proteins among the 246 were enriched in 21 pathways (Table
331 S3). Interestingly, of the all 59 30S and 50S ribosomal proteins in *S.* Typhimurium, 37 of these
332 proteins (63%) were downregulated in response to H_2O_2. Moreover, of the 8 identified proteins in
333 TCA cycle, 6 proteins were downregulated, including 2 essential proteins.
334 Although DDA method can be used to search for all proteins in a complex sample, it is prone to
335 miss the identification of important proteins due to the fact that fragmentation of tryptic peptides
336 from these proteins may not be triggered as a result of lower peptide ion intensities compared to
337 the threshold set. To quantify the proteins of the genes identified by Tn-seq more precisely and
338 accurately, we used targeted-proteomic approach by employing liquid chromatography coupled
339 with triple quadrupole mass spectrometry (LC-QQQ-ESI-MS). Here, tryptic peptides of the protein
340 were targeted for fragmentation (MS/MS) independent of their intensities, as described in
341 Materials and Methods, and the observed sequence specific fragment ion intensities from three
342 unique tryptic peptides were utilized for protein quantitation. Of the 137 Tn-seq identified genes,
343 we selected 33 genes to quantify their proteins in response to H_2O_2 by using targeted proteomics
344 (Dataset S3). Interestingly, 23 (70%) of the 33 tested proteins were upregulated in response to

345 H_2O_2. This shows a good agreement between the results of the Tn-seq and the targeted proteomics.

346 **Aromatic amino acid biosynthesis and H_2O_2**

347 Interestingly, our Tn-seq data revealed that the aromatic amino acid biosynthesis and metabolism
348 pathway play a role in conferring resistance in *Salmonella* to H_2O_2 (Fig. 3A and 3B). Five genes,
349 *aroB*, *aroD*, *aroE_2*, *aroK*, and *aroA* in the aromatic amino acid biosynthesis pathway were
350 identified by Tn-seq, and the fitness of the mutants were significantly reduced in the presence of
351 H_2O_2. To confirm this, 4 of these genes were evaluated using individual mutant assays. The
352 *Salmonella aroK* mutant showed the strongest phenotype, because it failed to grow in the presence
353 of H_2O_2L or H_2O_2H during 24 h incubation time. Also, the *aroB* mutant exhibited a strong
354 phenotype, significantly extending lag phase for both H_2O_2 conditions. The *aroE_2* mutant also
355 exhibited an extended lag time, but the *aroA* mutant did not show any difference in growth
356 phenotype in the presence of H_2O_2. In addition, targeted-proteomics also showed that all these 5
357 proteins were upregulated in response to H_2O_2. The most upregulated protein was *aroK* and this
358 was followed by the *aroE_2*, *aroA*, *aroB*, and *aroD* (Fig. 3C and Data Set S3).

359 The ROS damages a variety of biomolecules via Fenton reaction, which consequently lead to
360 metabolic defects, specifically auxotrophy for some aromatic amino acids (10). The *E. coli* mutants
361 that lack superoxide dismutase enzymes are unable to grow *in vitro* unless the medium are
362 supplemented with aromatic (Phe, Trp, Tyr), branched-chain (Ile, Leu, Val), and sulfur-containing
363 (Cys, Met) amino acids (36). We identified the genes in the aromatic amino acid biosynthesis
364 pathway that are critically important for resistance to H_2O_2. In this pathway, *aroK* catalyzes the
365 production of shikimate 3-phosphate from shikimate, which consequently leads to the production
366 of tryptophan, phenylalanine, tyrosine and some metabolites from the chorismate precursor in *E.*
367 *coli*. Further, the *aroK* mutant in *E.* coli displays increased susceptibility to protamine, a model

368 cationic antimicrobial peptide. It has been suggested that resistance to protamine is probably due
369 to the aromatic metabolites and product of *aroK* gene, which act as a signal molecule to simulate
370 the CpxR/CpxA system and Mar regulators (37). In our Tn-seq data, *cpxR*/*cpxA* and *marBCRT*
371 were in the list of non-required genes, but the proteomics data indicated that CpxR was
372 upregulated. Also, the *aroK* mutant in *E. coli* is resistance to mecillinam, a beta-lactam antibiotic
373 specific to penicillin-binding protein 2. It has been concluded that the AroK has a secondary
374 activity in addition to the aromatic amino acid biosynthesis, probably related to cell division (38).
375 In addition, the *aroK* gene presents a promising target to develop a non-toxic drug in
376 *Mycobacterium tuberculosis* because *aroK* is the only *in vitro* essential gene among the aromatic
377 amino acid pathway genes and blocking *aroK* kills the bacterium *in vivo* (39). Moreover, the *aroK*
378 gene plays a general role in *S.* Typhimurium persistence in pigs (40). The *aroB* is another gene in
379 the pathway that was identified by Tn-seq, which encodes 3-dehydroquinate synthase in the
380 Shikimate pathway, aromatic amino acid biosynthesis pathway. The *Salmonella* lacking *aroB*
381 showed a strong growth defect in the presence of H_2O_2. When this mutant was grown in the
382 presence of H_2O_2, it increased the lag phase time by a 114-fold for the H_2O_2L and a 347-fold for
383 the H_2O_2H as compared to the mutant grown in absence of H_2O_2. *S.* Typhimurium mutant lacking
384 the *aroB* gene is attenuated in BALB/c mice (41). In addition to *aroK* and *aroB*, *aroE_2* was also
385 shown to be important for resistance to H_2O_2, because deletion of the *aroE_2* reduced the growth
386 rate by 35% in the presence of H_2O_2 and increased the lag phase time, too. All these 3 genes in this
387 pathway are required for systemic infection of *Salmonella* in BALB/c mice in a more recent study
388 (18). We observed that there was a strong correlation between the fitness based on Tn-seq data,
389 growth rates measured by individual mutant assays, and upregulation of their proteins quantified
390 via targeted proteomics. This demonstrates the power of proteogenomic approach in discovering

391 and characterizing the genes that are required for growth under a specific condition.

392 **The *ybbM*, *ybbK*, and *ybbL* have a role in H_2O_2 resistance**

393 The mutants with single deletion in each of *ybbK*, *ybbL*, and *ybbM* genes on the same pathway
394 showed a strong phenotype against the activity of H_2O_2 in a dose-dependent manner. Based on Tn-
395 seq data, the fitness of *ybbM* was -1.16 and -1.79 for H_2O_2L and H_2O_2H, respectively (Fig.4A).
396 The *ybbM* mutant demonstrated decreased growth rate by 38% for H_2O_2L and 100% for the H_2O_2H
397 as compared to the mutant grown in the absence of H_2O_2. This mutant also increased the lag time
398 by a 126-fold and a 267-fold for H_2O_2L and H_2O_2H, respectively (Fig.4B). Also, the fitness score
399 of *ybbK* was -0.92 for H_2O_2L and -1.81 for H_2O_2H. The *ybbK* mutant showed decrease of growth
400 rate for both conditions H_2O_2L and H_2O_2H by 85% and 95%, respectively. The deletion increased
401 the lag phase by a 46-fold and a 114-fold in the presence of H_2O_2L and H_2O_2H, respectively
402 (Fig.4B). Moreover, the fitness of *ybbL* mutant was -1.05 and -1.73 for H_2O_2L and H_2O_2H,
403 respectively. Deleting the *ybbL* in *Salmonella* led to decrease in growth rate by 27% for H_2O_2L
404 and 92% for H_2O_2H. The lag phase time for this mutant increased by a 22-fold and a 33-fold for
405 H_2O_2L and H_2O_2H, respectively. In addition, YbbM, YbbL, and YbbK proteins were upregulated
406 in response to H_2O_2; YbbM was the most upregulated protein among the 3 proteins, followed by
407 YbbL, and YbbK (Fig. 4C and Data Set 3). The fitness scores of the Tn-seq of these 3 genes are
408 correlated strongly with the growth rate, lag time of their respective mutants, and upregulation of
409 their proteins. As the number of reads depletes after the selection for a mutant, (i) there was more
410 reduction in growth rate, (ii) the mutant stays longer in the lag phase, and (iii) the protein
411 expression elevates. These observations clearly point to their role in conferring resistance to the
412 H_2O_2-mediated stress. These genes were described in the *Salmonella* reference genome as follows:
413 *ybbM*, putative YbbM family transport protein, metal resistance protein; *ybbK*, putative inner

414 membrane proteins; *ybbL*, putative ABC transporter, ATP-binding protein YbbL. To the best of
415 our knowledge, there is only one published study on the *ybbM* and *ybbL* (42). Based on their
416 findings, YbbL and YbbM have a role in iron homeostasis in *E. coli* and are important for survival
417 when the bacterium was challenged with 10 mM H_2O_2 for 30 min; this putative ABC transporter
418 transports iron and lessens ROS species formation that generates via H_2O_2. In this study, we
419 identified an additional gene, *ybbK*, in the same pathway as the gene required for resistance to
420 H_2O_2, strongly establishing the role of these 3 genes in resistance to H_2O_2.

421 **The H_2O_2 scavenging and degrading genes**

422 *Salmonella* employs redundant enzymes to degrade or scavenge ROS. The *katE*, *katG*, and *katN*
423 genes encode catalases, which are involved in H_2O_2 degradation. The *ahpCF*, *tsaA*, and *tpx* genes
424 encode peroxidases, which scavenge H_2O_2. The *sodA*, *sodB*, *sodCI*, and *sodCII* genes encode
425 superoxide dismutases and these enzymes specifically scavenge O_2 (11, 12, 43 - 45). However,
426 none of these were present in the list of genes identified by Tn-Seq. Even though *katE*, *katG*, *ahpC*,
427 *sodA*, *sodCI*, and *sodCII* showed reduced fitness, they did not meet the statistical threshold.
428 However, the proteomics data indicated that AhpC, SodB, and TpX were upregulated in the
429 presence of H_2O_2 (Data Set S3) and KatG was also upregulated, but its *p* value was 0.054. This
430 reveals that these 4 proteins were the most important enzymes to scavenge H_2O_2 under our
431 experimental conditions. *Salmonella* containing an *ahpC* promoter-gfp fusion shows that
432 expression of the *ahpC* is regulated by ROS that is generated from macrophages or exogenous
433 H_2O_2 and the response to H_2O_2 is in a dose-dependent manner (46). *Salmonella* mutant that lacks
434 *katE*, *katG*, or *ahpCF* can degrade micromolar concentrations of H_2O_2. However, *Salmonella*
435 mutant that has deletions in the all 5 genes, *katE*, *katG*, *katN*, *ahpCF* and *tsaA* (HpxF), cannot
436 degrade H_2O_2, is unable to proliferate in macrophages, and showed reduced virulence in mice (11).

437 This data emphasizes that *ahpC*, *sodB*, and *tpx* may be the primary players in scavenging and
438 degrading H_2O_2 in our experiment. Why Tn-seq did not detect any of these genes, while proteomics
439 detected only these 3 proteins among others? It may reflect the functional redundancy in the
440 genetic network that prevented single deletions in one of the these genes from exhibiting fitness
441 defect. Alternatively, when these mutants were grown together with all other mutants in the library,
442 the functional protein lacking in one mutant due to Tn5 insertion could have been compensated by
443 the other mutants in the library.
444 In addition to these genes, *oxyR* was detected by Tn-seq (Fig. 1C) and DDA proteomics. The *oxyR*
445 was identified by all 3 analysis methods of Tn-seq data and it was on the top of the list, indicating
446 a severe fitness defect of the mutant. The *oxyR* gene encodes H_2O_2 sensor and transcription factor,
447 which mediates protection against ROS. The *katG* and *ahpCF* are regulated by OxyR, peroxide
448 response regulator (13, 14). Although *Salmonella* OxyR regulon is induced in the *Salmonella*-
449 containing vacuole in macrophage, the *oxyR* mutant was virulent in a BALB/c mouse and can grow
450 well in human neutrophils *in vitro* (47, 48). The fitness of *oxyR* mutant was reduced for both H_2O_2L
451 and H_2O_2H with the respective fitness score of -4.96 and -5.94. *Salmonella oxyR* mutant exhibited
452 the growth rate reduced by 24% and 40% for H_2O_2L and H_2O_2H, respectively. Comparing this
453 reduction in growth rate to the other mutants such as *rpoS* or *aroK*, we observed that the *oxyR*
454 mutant did not show severe phenotype and the mutant escaped from the lag phase easily.
455 Moreover, our targeted proteomics indicated that the OxyR was not upregulated significantly.
456 Further studies are needed to uncover the exact role of OxyR in response to ROS. However,
457 previous studies implied that OxyR plays an essential role in resistance to H_2O_2 by regulating other
458 proteins. OxyR induces Dps in *E. coli*, a ferritin-like protein that sequesters iron (49). Sequestering
459 of iron impairs the Fenton reaction, which consequently provides protection against ROS and

460 reduces the damage of biomolecules. The *dps* gene was identified by the Tn-seq and its fitness
461 score was -2.48. However, the Dps protein was downregulated based on the DDA proteomic
462 analysis. To confirm this unexpected finding, we conducted the proteomic assay twice and each
463 time with at least 4 technical replicates, but the Dps protein was significantly downregulated with
464 $p = 0.001$. Further, the targeted-proteomics gave the same result, pointing to the downregulation
465 of Dps in response to H_2O_2. This is contrary to the previously reported works on Dps in *Salmonella*
466 and the reason for the discrepancy is unclear.

467 **DNA repair system and H_2O_2**

468 The imposed exogenous H_2O_2 activates DNA repair system in *Salmonella* in order to repair or
469 eliminate the damage that occurred on the nucleotides. The *E. coli* RecA protein repairs double-
470 strand DNA lesions through recombination (50). In our Tn-seq analysis, the fitness score of this
471 mutant was -5.36 for both concentrations, and in proteomics, the RecA was upregulated.
472 *Salmonella recA* mutant decreased the maximum OD_{600} by 16% for H_2O_2L and 22% for H_2O_2H
473 as compared to the same mutant grown in LB. The *Salmonella recA* mutant was also sensitive to
474 exogenous H_2O_2 in aerated rich medium (26). Moreover, *recG*, recombination and DNA repair
475 gene (51), showed a stronger phenotype than *recA* mutant. The *regG* deletion in *Salmonella* caused
476 the growth rate reduction by 52% for H_2O_2L and 60% for H_2O_2H. This disruption in *recG* also
477 caused the cells to stay in lag phase for a longer time in the presence of H_2O_2 as compared to LB;
478 the lag time increased by a 62-fold and a 159-fold for H_2O_2L and H_2O_2H, respectively. In the blood
479 of patients with *Salmonella* Typhi bacteremia, the proteins encoded by *recA*, *recG*, and *xthA* genes
480 were detected, suggesting these proteins are actively expressed in the blood environment (52). The
481 XthA protein is another enzyme that participates in DNA repair that is induced by H_2O_2 and iron-
482 mediated Fenton reaction. The *xthA* encodes exonuclease III, which repairs the damaged DNA.

483 We found that the *xthA* gene was required based on the Tn-seq assay and its mutant had a reduced
484 fitness score of -3.06 for H_2O_2L and a -4.38 for the H_2O_2H. Further, *Salmonella* lacking the *xthA*
485 increased the lag time by 8-fold and a 12-fold for H_2O_2L and H_2O_2H, respectively. Targeted-
486 proteomics showed upregulation of XthA in response to H_2O_2. *Salmonella* enterica serovar
487 Enteritidis defective in *xthA* is susceptible to egg albumin (53). *E. coli xthA* mutant is
488 hypersensitive to H_2O_2 (54). The *xthA* is also required for *Mycobacterium tuberculosis* to infect
489 C57BL/6J mice (55). In addition to these genes that repair damaged DNA, *uvrA* encoding Holliday
490 junction DNA helicase motor protein, *uvrA* encoding exonuclease ABC subunit A, *uvrD* encoding
491 DNA-dependent helicase II, and *polA* encoding DNA polymerase I were among top of the list of
492 the genes identified by Tn-seq as required for resistance to H_2O_2. Collectively, DNA repair system
493 is crucial for the survival of the *Salmonella* in a niche that contains H_2O_2.

494 **Flageller genes and H_2O_2**

495 Some flageller genes, *fliC* and *fliB*, were shown to be important for resistance to H_2O_2. These two
496 genes were identified by Tn-seq and their proteins were shown to be upregulated based on
497 proteomics. *Salmonella* lacking either of these genes exhibited a strong phenotype in the presence
498 of H_2O_2. During 24 h of incubation, *fliC* and *fliB* mutants could not grow in both H_2O_2 conditions.
499 However, growth of *fliD* mutant was not affected by H_2O_2. The *fliC* was shown to have a role in
500 *Salmonella* Typhi interaction with human macrophages (56) and *Salmonella* Typhimurium *fliB*
501 mutant was defective in swarming motility (57). Currently it remains unclear how flagella genes
502 can be involved in the resistance of *Salmonella* to oxidative stress, which warrants future study
503 into this direction.

504 **Fe-S cluster biogenesis system and H_2O_2**

505 *Salmonella* requires the genes from Fe-S cluster biogenesis system in order to resist H_2O_2. Our

506 Tn-seq analysis identified 5 genes in this system as required for the resistance. In *isc* operon (Fe-
507 S cluster), *iscA*, *hscB*, and *hscA* were among the genes required to resist H_2O_2. Particularly, the
508 *hscA* is on the top of the gene list identified by Tn-seq. In *E. coli*, this operon is regulated by *iscR*,
509 iron sulfur cluster regulator (58); in *Salmonella* the gene *iscR* encoding this transcription regulator
510 is named *yfhP*. The HscB and HscA chaperones are believed to be involved in the maturation of
511 [2Fe–2S] proteins (59, 60). The second operon that is involved in Fe-S protein biogenesis is the
512 *suf*, sulfur mobilization operon. Tn-seq found that two genes in this operon were required for
513 *Salmonella* to resist H_2O_2; *sufS* and *sufC*. *Salmonella* mutant lacking *sufS* exhibited a strong
514 phenotype in the presence of H_2O_2 and could not grow during the 24 h of incubation as compared
515 to LB. The SufS with SufE in *E. coli* form a heterodimeric cysteine desulphurase and SufB, SufC,
516 and SufD form a pseudo-ABC-transporter that could act as a scaffold (60); this operon is regulated
517 by OxyR (14). The other known genes in these two operons that are present in *Salmonella* are *iscA*,
518 *sufA*, *sufB*, and *sufD*; they showed a reduced fitness, while their *p* values were greater than 0.05.
519 The damage of Fe-S clusters is not only problem for the defective proteins, but also it fuels the
520 Fenton reaction via the released iron and H_2O_2 (10). Thus, *Salmonella* uses Fe-S cluster repair
521 system as an arsenal to overcome the damage imposed by H_2O_2.

522 **DNA adenine methylase**

523 DNA adenine methylase genes, *dam* and *damX*, are involved in *Salmonella* resistance against
524 H_2O_2. Our Tn-seq data showed that fitness of *dam* and *damX* mutants were reduced in the presence
525 of H_2O_2. To confirm this, *Salmonella dam* mutant was grown in both conditions. Under H_2O_2L,
526 the growth rate was reduced by 42% as compared to the mutant in LB and the mutant could not
527 grow under H_2O_2H during the 24 h of incubation. In addition, the lag time of the *Salmonella dam*
528 mutant was extended by a 19-fold for H_2O_2L. While the *Salmonella damX* mutant displayed a

529 moderate phenotype as compared to the *dam* mutant, the *damX* mutant also showed the growth
530 rate decreased by 23% and 33% for H_2O_2L and H_2O_2H, respectively. The lag time was extended
531 for this mutant by a 25-fold and a 49-fold for H_2O_2L and H_2O_2H, respectively. The *dam* regulates
532 virulence gene expression in *S*. Typhimurium (61). The different levels of Dam affects virulence
533 gene expression, motility, flagellar synthesis, and bile resistance in the pathogenic *S*. Typhimurium
534 14028S (62). Dam methylation activates the gene that are involved in lipopolysaccharide synthesis
535 (63). Moreover, *Salmonella* defective in *damX* is very sensitive to bile (64). Collectively, our study
536 demonstrates the critical role of DNA adenine methylase in *Salmonella* resistance against H_2O_2.

537 **Other genes for H_2O_2 resistance**

538 Beside the important pathways described above, there were many additional genes also important
539 for resistance to H_2O_2. Among those, the 3 unrelated genes, *rpoS*, *pgm*, and *tonB*, are important
540 ones that deserve more attention. The *rpoS* mutant showed reduced fitness and its protein was
541 upregulated in the presence of H_2O_2. *Salmonella* mutant defective in *rpoS* showed a strong
542 phenotype when grown in the presence of H_2O_2. The *rpoS* encodes the alternative sigma factor σ^S,
543 subunit of RNA polymerase; it is the master regulator of stress response (65). This implies that
544 *rpoS* is an important component of the genetic regulatory network that *Salmonella* employs in
545 order to resist H_2O_2. Furthermore, the fitness of *pgm* mutant was reduced and its protein was
546 upregulated in the presence of H_2O_2. Knock out of *pgm* in *Salmonella* caused a decrease in growth
547 rate, increase the lag phase time, and reduce the maximum OD_{600} in the presence of H_2O_2. The
548 *pgm* encodes phosphoglucomutase required for catalysis of the interconversion of glucose 1-
549 phosphate and glucose 6-phosphate (66). This gene contributes to resistance against antimicrobial
550 peptides, is required for *in vivo* fitness in the mouse model, and participates in LPS biosynthesis
551 (67). Lastly, the fitness of *tonB* mutant was also reduced. *Salmonella* lacking *tonB* exhibited a

552 strong phenotype in the presence of H_2O_2 as compared to the mutant grown in LB. The gene
553 mediates iron uptake in the *Salmonella* (45). In addition, seven of the genes identified in our study
554 (*proC*, *arcA*, *barA*, *exbD*, *flhD*, *fliC*, and *fliD*) were previously shown to be important for
555 interaction of *Salmonella* Typhi with human macrophages (56). To see the full list of *S.*
556 Typhimurium genes required for H_2O_2 resistance, the supplementary information contains data of
557 Tn-seq, individual mutant assays, and proteomic analysis in Data Set S1, S2, and S3, respectively.
558

559 **Conclusions**

560 We applied Tn-seq and proteomic analysis to find the genes and proteins that are required in *S.*
561 Typhimurium to resist H_2O_2 *in vitro*. As the concentration of H_2O_2 increased, the growth rate
562 reduced, the lag time extended, the fitness of mutants decreased, and some proteins were
563 differentially expressed. Validation of Tn-seq results with individual mutant assays indicated the
564 accuracy of the identified genes in response to the two H_2O_2 concentrations. The targeted-
565 proteomics had a good agreement with Tn-seq. We found about 80 genes that have not been
566 associated to resistance to H_2O_2 previously. *Salmonella* employs multiple pathways to resist H_2O_2
567 and the most important ones are ROS detoxifying enzymes, amino acid biosynthesis (*aroK* and
568 *aroB*), putative iron transporters (*ybbK*, *ybbL*, *ybbM*), iron homeostasis, Fe-S cluster repair, DNA
569 repair, flagellar and DNA adenine methylase genes. The genes identified in this study will broaden
570 our understanding on the mechanisms used by *Salmonella* to survive and persist against ROS in
571 macrophages.

572 Our unbiased system-wide approach, Tn-seq, was successful in identifying novel genetic
573 determinants that have not been implicated previously in *Salmonella* resistance to oxidative stress.
574 Furthermore, the combined use of quantitative proteomic approach has provided additional

575 insights on the function or mode of action of the identified genetic determinants in resisting
576 oxidative stress. As expected, the majority of the proteins important for resistance to H_2O_2 were
577 upregulated in response to the same stressor. However, the expression level did not increase for
578 some proteins, in spite of their known roles in resistance to H_2O_2. Interestingly, the downregulation
579 of Dps and other proteins was counterintuitive to the common mode of protein regulation and
580 function, yet it may point to some unknown aspects of how *Salmonella* regulates the expression
581 of those proteins to better cope with the oxidative stress during infection in macrophage.
582

583 **Materials and Methods**

584 **Construction of Tn5 mutant library**

585 We mutagenized *Salmonella* enterica subsp. enterica serovar Typhimurium str. ATCC 14028S
586 (with spontaneous mutation conferring resistance to nalidixic acid (NA)), by biparental mating
587 using *Escherichia coli* SM10 λ*pir* carrying a transposon-delivery plasmid vector pBAM1
588 (Ampicillin (Amp) resistance) as a donor strain (68). The plasmid pBAM1 was generously
589 provided by Victor de Lorenzo (Systems and Synthetic Biology Program, Centro Nacional de
590 Biotecnología, Madrid, Spain). The donor strain, *E. coli* Sm10 λpir (pBAM1), was grown
591 overnight in LB with 50 µg/ml Amp and recipient strain was grown in LB with 50 µg/ml NA at
592 37 °C . Equal volumes (1 ml) of donor and recipient were mixed, centrifuged, washed in 10 mM
593 $MgSO_4$, and re-suspended in 2 ml PBS (pH 7.4). Then, the mating mixture was concentrated and
594 laid on a 0.45-µm nitrocellulose filters (Millipore). The filter was incubated at 37°C on the surface
595 of LB agar plate. After 5 h of conjugation, the cells on the filter was washed in 10 mM $MgSO_4$,
596 and resuspended in 1 ml MgSO4. The conjugation mixture was plated on LB agar containing 50
597 µg/ml NA and 50 µg/ml kanamycin (Km). After approximately 24 h at 37°C, we scraped the

598 colonies into LB broth containing 50 μg/ml Km and 7% DMSO. The yield was approximately
599 68,000 individual colonies from each conjugation. Five independent conjugations were conducted
600 to yield approximately 325,000 mutants. The library was stored at −80°C in aliquots. To determine
601 the frequency of the mutants that have been produced by integration of the entire delivery plasmid,
602 the colonies were picked randomly and streaked on LB plates (Km) and LB plates (Km and Amp).
603 It was shown that ~20% of the cells in the library were resistant to Amp, indicating a significant
604 portion of the Km-resistant colonies was not from authentic transposition events.

605 **Measuring growth responses of *S.* Typhimurium to H_2O_2**

606 To determine the effect of H_2O_2 concentrations on growth parameters, overnight culture of the
607 wild type *S.* Typhimurium 14028s was inoculated into fresh LB broth media (1:200 dilution) to
608 give a seeding concentration corresponding to OD_{600} of ~0.1. The LB broth contained freshly
609 prepared H_2O_2 to give the final concentrations ranging from 0.05 to 10 mM. The cultures were
610 directly added into 96-well microplate (200 μl/well). The microplate was incubated in a Tecan
611 Infinite 200 microplate reader at 37°C, with shaking duration 5 s, shaking amplitude 1.5 mm, and
612 reading OD_{600} every 10 min. The number of replicates were at least three. The lag time, growth
613 rate, and maximum OD_{600} were calculated using GrowthRates script (69). Growth Rate % decrease
614 was calculated as follows (70): Growth Rate % decrease = $((\mu_{PC} - \mu_S)/\mu_{PC}) \times 100$; where μ = the
615 maximum slope (growth rate), μ_{PC} = growth rate of positive control (without H_2O_2), μ_S = growth
616 rate in the presence of H_2O_2.

617 **Selection of the mutant library for Tn-seq analysis**

618 The transposon library was thawed at room temperature and diluted 10^{-1} in fresh LB broth. To
619 activate the library, the diluted library was incubated at 37°C with shaking at 225 rpm for an hour.
620 Then, the culture was washed twice with PBS and resuspended in LB broth medium. The library

621 was inoculated to 20 ml LB broth and LB broth supplemented with either 2.5 or 3.5 mM H_2O_2
622 (H_2O_2L and H_2O_2H, respectively), seeding CFU was 3.5 x 10^6 per ml. Then, when the cultures
623 reached mid-exponential phase, OD_{600} of 2.7 (~1.17 x 10^8 CFU/ml), the incubation was stopped,
624 and the culture was immediately harvested by centrifugation, and stored at -20°C.

625 **Preparation of Tn-seq amplicon libraries**

626 Genomic DNA was extracted from the harvested cells using DNeasy Blood & Tissue kit (Qiagen),
627 and quantified using Qubit dsDNA RB Assay kit (Invitrogen). As described above, 20% of the
628 mutants in the library were the result of the integration of pBAM1 into chromosome. To remove
629 the Tn5-junction sequences originated from the plasmid in the Tn-seq amplicon libraries, genomic
630 DNA was digested with PvuII-HF (New England Biolabs), which digests immediately outside the
631 inverted repeats on both sides of Tn5 in pBAM1, and purified with DNA Clean & Concentrator-5
632 kit (Zymo Reaerch). Then, a linear PCR extension was performed using a Tn5-specific primer in
633 order to produce single stranded DNA corresponding to Tn5-junction sequences. To increase the
634 specificity in extending into Tn5-junction sequences, the linear PCR was conducted with a dual
635 priming oligonucleotide Tn5-DPO (5'-AAGCTTGCATGCCTGCAGGTIIIIICTAGAGGATC-3')
636 that is specific to Tn5 end (71). The PCR reaction contained 25 µl Go Taq Colorless Master Mix
637 (Promega), 20 µM Tn5-DPO primer, 100 ng gDNA, and 50 µl MQ-H_2O. The PCR cycle consisted
638 of the initial denaturation at 95°C for 2 min, followed by 50 cycles at 95°C for 30 sec, 62°C for 45
639 sec, and 72°C for 10 sec. The PCR product was purified with DNA Clean & Concentrator-5 kit
640 and eluted in 13 µl TE buffer. After that, C-tail was added to the 3' end of the single-stranded
641 DNA. The C-tailing reaction was consisted of 2 µl terminal transferase (TdT) buffer (New England
642 Biolabs), 2 µl $CoCl_2$, 2.4 µl 10 mM dCTP, 1 µl 1 mM ddCTP, 0.5 µl TdT and 13 µl purified linear
643 PCR product. The reaction was performed at 37°C for 1 h and the enzyme was inactivated by

644 incubation at 70°C for 10 min. The C-tailed product was purified with DNA Clean & Concentrator-
645 5 kit and eluted in 12 µl TE. Next, the exponential PCR was performed with forward primer, P5-
646 BRX-TN5-MEO,
647 AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC
648 TNNNNAG-6 nt barcode-CCTAGGCGGCCTTAATTAAAGATGTGTATAAGAG and reverse
649 primer, P7-16G,
650 CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGGGGGGGGGGGGGGGG to attach
651 Illumina adapter sequences along with the sample barcodes. The PCR reaction contained 25 µl Go
652 Taq Green Master Mix, 10 µM P5-BRX-TN5-MEO primer, 10 µM P7-16G primer, 1 µl purified
653 C-tailed genomic junctions, and MQ-H_2O to 50 µl; the PCR condition started with initial
654 denaturation at 95°C for 2 min, followed by 36 cycles of 95°C for 30 sec, 60°C for 30 sec, and
655 72°C for 20 sec, with the final extension at 72°C for 5 min. Then, the size selection of the DNA
656 was performed using agarose gel electrophoresis. The 50 µl PCR products were incubated at 60°C
657 for 15 min and incubated on ice for 5 min, and immediately loaded on the 1% agarose gel in 0.5%
658 TAE buffer. After running the gel, the DNA fragment of size 325 – 625 bp was cut and put in a
659 microtube for each sample. The DNA was extracted from the gel using Zymoclean Gel DNA
660 Recovery kit (Zymo Reaerch). The prepared DNA libraries were quantified using Qubit dsDNA
661 RB Assay kit. Since each library has its own barcode, the libraries were combined and sequenced
662 on a flow cell of HiSeq 3000 using single end read and 151 cycles (Illumina) at the Center for
663 Genome Research & Biocomputing in Oregon State University.

664 **Tn-seq data analysis**

665 The preliminary data analysis was conducted by using a super computer in the High Performance
666 Computing Center (AHPCC) at the University of Arkansas. The libraries that were multiplexed

667 for sequencing were de-multiplexed using a custom Python script. The script searched for the six-
668 nucleotide barcode for each library for perfect matches. In order to extract the transposon genomic
669 junctions, we used Tn-Seq Pre-Processor (TPP) tool (30) with some modifications in the script.
670 The TPP searched for the 19 nucleotide inverted repeat (IR) sequence and identified five
671 nucleotides (GACAG) at the end of the IR sequence, allowing one nucleotide mismatch. The Tn5-
672 junctions that start immediately after GACAG were extracted and the C-tails at the end of junctions
673 were removed. Tn5-junction sequences less than 20 nucleotides were discarded and remaining
674 Tn5-junctions were mapped to the *Salmonella* enterica serovar Typhimurium 14028S genome and
675 plasmid using BWA-0.7.12 (72). To identify genes that are required for H_2O_2 resistance, the
676 following three Tn-seq analysis tools were used for comparative analysis: (i) ARTIST (28): the
677 genomic junctions were mapped to the reference genome using Bowtie 2.2.7 (73). The number of
678 insertions and reads were determined for genes and intergenic regions. The data were normalized
679 with default script in the ARTIST. Then, the relative abundance of Tn5 insertions in the output
680 library versus the input were calculated. Later, the *p* values were calculated from a 100 independent
681 Mann–Whitney U test (MWU) analysis that were carried out on input and output data for each
682 gene. Finally, the genes were considered conditionally essential if the *p* values were ≤ 0.05 in the
683 90 of the 100 MWU tests. (ii) Tn-seq Explorer (29): The output SAM files from the TTP were
684 used as input to the Tn-seq Explorer. The unique insertions with less than 20 reads were removed
685 from the input and outputs. Using the window size of 550 and excluding 5% of beginning of genes
686 and 20% of the end of genes, Essentiality Index (EI), number of unique insertions, and total number
687 of reads per gene were counted. The EI of more than 10 were removed from the input. Genes with
688 less than 300 nucleotides were removed. Deferential EI were calculated from input and outputs
689 (ΔIE = output EI–input EI) and genes with ΔIE more than -1 were removed. Log2 fold change of

690 reads were calculated from input and output (Log2FC = log2 (output reads/input reads)) and the
691 genes were ranked based on the Log2FC value from least value to highest. The genes with Log2FC
692 value of less than -2 from the H_2O_2H and present in H_2O_2L were considered conditionally essential.
693 (iii) TRANSIT (30): The output wig files from the TTP was used as input data file for TRANSIT.
694 The comparative analysis was conducted with Tn5 resampling option. The reads were normalized
695 with trimmed total reads (TTR). Insertions outside the 5% and 10% sequences from 5'- and 3'-
696 ends were removed, respectively. The genes were considered conditionally essential if p values ≤
697 0.05.

698 **Phenotypic evaluation of individual deletion mutants**

699 The mutants were obtained from *Salmonella enterica* subsp. *enterica*, 14028s (Serovar
700 Typhimurium) Single-Gene Deletion Mutant Library through BEI Resources
701 (www.beiresources.org). The overnight cultures of *S.* Typhimurium mutants were added into fresh
702 LB broth media containing freshly prepared H_2O_2 (2.5, or 3.5 mM/ml) (1:200 dilution) to give
703 seeding OD_{600} of 0.1. The cultures were directly added into 96-well microplates and incubated in
704 Tecan Infinite 200 at 37ºC for 24 h. The lag time, growth rate, and maximum OD_{600} were calculated
705 using GrowthRates (69).

706 **Sample preparation for proteomics and mass spectrometry analysis**

707 The overnight culture of the wild type *S.* Typhimurium 14028 was diluted 1:200 in 50 ml LB
708 medium, and LB containing either 2.5 or 3.5 mM H_2O_2 in a 300-ml flask. The cultures were grown
709 until mid-exponential phase (OD_{600} of 2.7), and the 50 ml volume of cultures were used for a total
710 protein extraction by using Qproteome Bacterial Protein Prep kit (Qiagen). The proteins were
711 denatured with mercaptoethanol and separated by SDS-PAGE gel electrophoresis. For each
712 condition, there were three lanes with approximately 300 μg of proteins. The gel portions of 3

713 lanes for each condition were cut out and chopped into small pieces, pooled together, washed twice
714 with 50 mM NH_4HCO_3, destained with NH_4HCO_3/ 50% Acetonitrile (ACN), and dried with pure
715 ACN. Then, the proteins were reduced using 10 mM Dithiothreitol in 50 mM NH_4CO_3 and the
716 alkylation was conducted with 10 mg/ml Iodoacetamide Acid in 50 mM NH_4CO_3. After that, the
717 proteins were washed with NH_4HCO_3, and dried with pure ACN. Mass spectrometry grade Trypsin
718 gold from Promega (~20 ng/µl in 50 mM NH_4HCO_3) was added to dried gels, and left it overnight
719 for efficient in-gel digestion of the proteins at 37ºC. During the digestion, tryptic peptides diffused
720 out into the solution. Later, these digests were analyzed by ESI-LC-MS/MS at State Wide Mass
721 Spectrometry Facility, University of Arkansas at Fayetteville. Data dependent analysis (DDA) for
722 the in-gel trypsin digested samples from each condition were performed by using an Agilent 1200
723 series micro flow HPLC in line with Bruker Amazon-SL quadrupole ion trap ESI mass
724 spectrometer (QIT-ESI-MS). All the ESI-MS analyses were performed in a positive ion mode
725 using Bruker captive electrospray source with a dry nitrogen gas temperature of 200°C, with
726 nitrogen flow rate of 3 L/minute. LC-MS/MS data were acquired in the Auto MS(n) mode with
727 optimized trapping condition for the ions at m/z 1000. MS scans were performed in the enhanced
728 scanning mode (8100 m/z/second), while the collision-induced dissociation or the MS/MS
729 fragmentation scans were performed automatically for top ten precursor ions with a set threshold
730 for one minute in the UltraScan mode (32,500 m/z/second). Tryptic peptides were separated by
731 reverse-phase high-performance liquid chromatography (RP-HPLC) using a Zorbax SB C18
732 column, (150 × 0.3 mm, 3.5 µm particle size, 300 Å pore size, Agilent Technologies), with a
733 solvent flow rate of 4 µL/minute, and a gradient of 5%–38% consisting of 0.1% FA (solvent A)
734 and ACN (solvent B) over a time period of 320 minutes. Tryptic peptides were then identified by
735 searching MS/MS data in *S.* Typhimurium 14028S reference proteome database (24, 35) by using

736 MASCOT database search software (34). MS1 intensities of the integrated areas of these identified
737 tryptic peptides were compiled and grouped in skyline software according to the
738 replicates/conditions to perform statistical analysis. Targeted protein work were performed using
739 Shimadzu UPLC-20A coupled to 8050 triple quadrupole ESI-MS with heated probe. Sequence
740 specific fragment ion intensities from at least three unique tryptic peptides from the protein of
741 interest were used in the protein quantitation. Multiple reaction monitoring (MRM) events
742 corresponding to sequence specific fragment ions derived from the precursor tryptic peptides were
743 targeted to operate at a certain specific retention time intervals predicted by in house retention time
744 library. This library was generated using the correlation of relative hydrophobicity of the tryptic
745 peptides with their retention times (RT) from highly common housekeeping proteins, for the UPLC
746 method used in this analysis as described below. While the RT were correlated well within 99%
747 confidence, sufficient number of sequence specific fragment ions were used as basis for
748 identification of the tryptic peptide by MS/MS alone. Specificity and the confidence was achieved
749 by incorporating RT prediction. In addition to the application of skyline in quantitation, skyline
750 software was also used in predicting RT and optimizing parameters such as collision energies and
751 voltages with the help of Shimadzu Labsolution software. Tryptic peptides were separated by
752 reverse-phase ultra-high-performance liquid chromatography (RP-UPLC) compatible Shimadzu
753 C18, 1.9-micron particle size, 50x2.1 mm column (SN # 16041880T), with a solvent flow rate of
754 0.3 mL/minute, and a gradient of 5%–90% consisting of 0.1% FA (solvent A) and 0.1% FA in
755 ACN (solvent B) over a time period of 10 minutes.

756 **Accession numbers**

757 Tn-seq sequencing data are available on NCBI Sequence Read Archive. LB: PRJNA352537;
758 H_2O_2L: PRJNA352862; H_2O_2H: PRJNA352865.

759 **Funding**

760 The bench fee of this work was came from a fund from Arkansas Biosciences Institute. The first
761 author was supported by his parents, Cell and Molecular Biology (CEMB) program at the
762 University of Arkansas, and Human Capacity Development Program-Kurdistan Regional
763 Government (HCDP-KRG).

764 **Acknowledgment**

765 We thank Jack Jiang, a high school student, for writing the demultiplexing Python script. We are
766 thankful for Dr. Thomas R. Ioerger and Dr. Michael A. DeJesus at Texas A&M University for the
767 help on the use of TRNASIT. We extend our special thanks to Dr. Jeff F. Pummill and Dr. Pawel
768 Wolinski at Arkansas High Performance Computing Center for their help to use the facility.

769 **Author Contributions**

770 Conceived and designed the experiments: YK SK. Performed the experiments, analyzed the data,
771 wrote the manuscript: SK. Proteomics work: RL SK AQ JL. Revised the manuscript: YK SK. All
772 authors read the final version of the manuscript.

773 **Supplementary Information**

774 **Data Set S1**

775 Full list of *Salmonella* Tn-seq genome for H_2O_2L and H_2O_2H analyzed by ARTIST, Tn-Seq
776 Explorer, and ARTIST. The list of 137 *Salmonella* genes for H_2O_2 resistance.

777 **Data Set S2**

778 Lag time, growth rate, and maximum $OD6_{00}$ of *Salmonella* grown in LB, H_2O_2L, H_2O_2H.

779 **Data Set S3**

780 *Salmonella* proteomic analysis in response to the H_2O_2, non-targeted and targeted proteomics.

781

782 **Table S1**

783 *Salmonella* Tn-seq sequencing in numbers for LB, H_2O_2L, and H_2O_2H

784 **Table S2**

785 List of functional categories required for *Salmonella* H_2O_2 resistance.

786 **Table S3**

787 List of deferentially expressed proteins of *Salmonella* in response to H_2O_2 and their pathways.

788

789 **Fig S1**

790 Stability of H_2O_2 in the medium and during the experiment time.

791 **Fig S2**

792 Tn-seq library preparation diagram for Illumina sequencing.

793 **Fig S3**

794 Correlation between reads of *Salmonella* Tn-seq conditions, H_2O_2L and H_2O_2H.

795

796 **REFERENCES**

797 1. **LaRock DL, Chaudhary A, Miller SI.** 2015. Salmonellae interactions with host

798 processes. Nat Rev Microbiol **13:**191-205.

799 2. **Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL,**

800 **Griffin PM.** 2011. Foodborne illness acquired in the United States – major pathogens.

801 Emerg Infect Dis **17:**1-15.

802 3. **Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A,**

803 **Hoekstra RM.** 2010. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin

804 Infect Dis **50:**882-889.

805 4. **Haraga A, Ohlson MB, Miller SI.** 2008. Salmonellae interplay with host cells. Nat Rev
806 Microbiol **6:**53-66.
807 5. **Nauseef W.** 2004. Assembly of the phagocyte NADPH oxidase. Histochem Cell Biol
808 **122:**277-291.
809 6. **Seaver LC, Imlay JA.** 2001. Hydrogen peroxide fluxes and compartmentalization inside
810 growing *Escherichia coli*. J Bacteriol **183:**7182-7189.
811 7. **Park S, You X, Imlay JA.** 2005. Substantial DNA damage from submicromolar
812 intracellular hydrogen peroxide detected in Hpx- mutants of *Escherichia coli*. Proc Natl
813 Acad Sci U S A **102:**9317-9322.
814 8. **Imlay JA, Linn S.** 1988. DNA damage and oxygen radical toxicity. Science **240:**1302-
815 1309.
816 9. **Imlay JA.** 2006. Iron-sulphur clusters and the problem with oxygen. Mol Microbiol
817 **59:**1073-1082.
818 10. **Slauch JM.** 2011. How does the oxidative burst of macrophages kill bacteria? Still an open
819 question. Mol Microbiol **80:**580-583.
820 11. **Hébrard M, Julie P. M. Viala, Méresse S, Barras F, Aussel L.** 2009. Redundant
821 Hydrogen Peroxide Scavengers Contribute to *Salmonella* Virulence and Oxidative Stress
822 Resistance. J Bacteriol **191:**4605-4614.
823 12. **Horst SA, Jaeger T, Denkel LA, Rouf SF, Rhen M, Bange F.** 2010. Thiol peroxidase
824 protects *Salmonella enterica* from hydrogen peroxide stress *in vitro* and facilitates
825 intracellular growth. J Bacteriol **192:**2929-2932.
826 13. **Morgan RW, Christman MF, Jacobson FS, Storz G, Ames BN.** 1986. Hydrogen
827 Peroxide-Inducible Proteins in *Salmonella* typhimurium Overlap with Heat Shock and

828 Other Stress Proteins. Proc Natl Acad Sci U S A **83:**8059-8063.

829 14. **Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA, Storz G.** 2001. DNA
830 microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen
831 peroxide. J Bacteriol **183:**4562-4570.

832 15. **Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran S, Hammarlof D, Canals**
833 **R, Grissom J, Conway T, Hokamp K, Hinton J.** 2013. An infection-relevant
834 transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. Cell Host
835 Microbe **14:**683-695.

836 16. **Turner KH, Wessel AK, Palmer GC, Murray JL, Whiteley M.** 2015. Essential genome
837 of *Pseudomonas aeruginosa* in cystic fibrosis sputum. Proc Natl Acad Sci U S A **112:**4110-
838 4115.

839 17. **Chan K, Kim CC, Falkow S.** 2005. Microarray-based detection of *Salmonella enterica*
840 Serovar Typhimurium transposon mutants that cannot survive in macrophages and mice.
841 Infect Immun **73:**5438-5449.

842 18. **Silva-Valenzuela C, Molina-Quiroz R, Desai P, Valenzuela C, Porwollik S, Zhao M,**
843 **Hoffman R, Andrews-Polymenis H, Contreras I, Santiviago C, McClelland M.** 2016.
844 Analysis of two complementary single-gene deletion mutant libraries of *Salmonella*
845 Typhimurium in intraperitoneal infection of BALB/c Mice. Front Microbiol **6:**1455.

846 19. **van Opijnen T, Bodi KL, Camilli A.** 2009. Tn-seq: high-throughput parallel sequencing
847 for fitness and genetic interaction studies in microorganisms. Nat Methods **6:**767-772.

848 20. **Khatiwara A, Jiang T, Sung S, Dawoud T, Kim JN, Bhattacharya D, Kim H, Ricke**
849 **SC, Kwon YM.** 2012. Genome scanning for conditionally essential genes in *Salmonella*
850 *enterica* Serotype Typhimurium. Appl Environ Microbiol **78:**3098-3107.

851 21. **Chaudhuri R, Morgan E, Peters S, Pleasance S, Hudson D, Davies H, Wang J, van**
852 **Diemen P, Buckley A, Bowen A, Pullinger G, Turner D, Langridge G, Turner A,**
853 **Parkhill J, Charles I, Maskell D, Stevens M.** 2013. Comprehensive assignment of roles
854 for *Salmonella* Typhimurium genes in intestinal colonization of food-producing animals.
855 PLoS Genet **9:**e1003456.
856 22. **Langridge GC, Phan M, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell**
857 **DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK.** 2009. Simultaneous assay of
858 every *Salmonella* Typhi gene using one million transposon mutants. Genome Res **19:**2308-
859 2316
860 23. **Barquist L, Langridge GC, Turner DJ, Phan M, Turner AK, Bateman A, Parkhill J,**
861 **Wain J, Gardner PP.** 2013. A comparison of dense transposon insertion libraries in the
862 *Salmonella* serovars Typhi and Typhimurium. Nucleic Acids Res **41:**4549-4564.
863 24. **Domon B, Aebersold R.** 2010. Options and considerations when selecting a quantitative
864 proteomics strategy. Nat Biotechnol **28:**710-721.
865 25. **Jensen ON, Mann M.** 2003. Proteomic analysis of post-translational modifications. Nat
866 Biotechnol **21:**255-261.
867 26. **Bogomolnaya L, Andrews K, Talamantes M, Maple A, Ragoza Y, Vazquez-Torres A,**
868 **Andrews-Polymenis H.** 2013. The ABC-Type efflux pump MacAB protects *Salmonella*
869 *enterica* serovar Typhimurium from oxidative stress. Mbio **4:**e00630.
870 27. **Kwon YM, Ricke SC, Mandal RK.** 2016. Transposon sequencing: methods and
871 expanding applications. Appl Microbiol Biotechnol **100:**31-43.
872 28. **Pritchard JR, Chao MC, Abel S, Davis BM, Baranowski C, Zhang YJ, Rubin EJ,**
873 **Waldor MK.** 2014. ARTIST: high-resolution genome-wide assessment of fitness using

874 transposon-insertion sequencing. PLoS Genet **10:**e1004782-e1004782.

875 29. **Solaimanpour S, Sarmiento F, Mrazek J.** 2015. Tn-Seq Explorer: A tool for analysis of
876 high-throughput sequencing data of transposon mutant libraries. Plos One **10:**e0126070.

877 30. **DeJesus MA, Ambadipudi C, Baker R, Sassetti C, Ioerger TR.** 2015. TRANSIT--A
878 software tool for Himar1 TnSeq analysis. PLoS Comput Biol **11:**e1004401-e1004401.

879 31. **Huang DW, Sherman BT, Lempicki RA.** 2009. Bioinformatics enrichment tools: paths
880 toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res **37:**1-
881 13.

882 32. **Kanehisa M, Goto S.** 2000. KEGG: Kyoto Encyclopedia of genes and genomes. Nucleic
883 Acids Res **28:**27-30.

884 33. **MacLean B, Tomazela D, Shulman N, Chambers M, Finney G, Frewen B, Kern R,**
885 **Tabb D, Liebler D, MacCoss M.** 2010. Skyline: an open source document editor for
886 creating and analyzing targeted proteomics experiments. Bioinformatics **26:**966-968.

887 34. **Perkins D, Pappin D, Creasy D, Cottrell J.** 1999. Probability-based protein identification
888 by searching sequence databases using mass spectrometry data. Electrophoresis **20:**3551-
889 3567.

890 35. **Cravatt BF, Yates III JR, Simon GM.** 2007. The biological impact of mass-
891 spectrometry-based proteomics. Nature **450:**991-1000.

892 36. **Carlioz A, Touati D.** 1986. Isolation of superoxide dismutase mutants in *Escherichia coli*:
893 is superoxide dismutase necessary for aerobic life? Embo j **5:**623-630.

894 37. **Weatherspoon-Griffin N, Yang D, Kong W, Hua Z, Shi Y.** 2014. The CpxR/CpxA Two-
895 component Regulatory System Up-regulates the Multidrug Resistance Cascade to
896 Facilitate *Escherichia coli* Resistance to a Model Antimicrobial Peptide. J Biol Chem

897 **289:**32571-32582.

898 38. **Vinella D, Gagny B, Joseleau-Petit D, D'Ari R, Cashel M.** 1996. Mecillinam resistance
899 in *Escherichia coli* is conferred by loss of a second activity of the AroK protein. J Bacteriol
900 **178:**3818-3828.

901 39. **Kaur P, Datta S, Shandil RK, Kumar N, Robert N, Sokhi UK, Guptha S, Narayanan**
902 **S, Anbarasu A, Ramaiah S.** 2016. Unravelling the secrets of mycobacterial cidality
903 through the lens of antisense. PLoS One **11:**e0154513.

904 40. **Van Parys A, Boyen F, Leyman B, Verbrugghe E, Haesebrouck F, Pasmans F.** 2011.
905 Tissue-specific *Salmonella* Typhimurium gene expression during persistence in Pigs. Plos
906 One **6:**e24120.

907 41. **Günel-Özcan A, Brown KA, Allen AG, Maskell DJ.** 1997. *Salmonella* typhimurium
908 aroB mutants are attentuated in BALB/ c mice. Microb Pathog **23:**311-316.

909 42. **Nicolaou SA, Fast AG, Nakamaru-Ogiso E, Papoutsakis ET.** 2013. Overexpression of
910 fetA (ybbL) and fetB (ybbM), encoding an Iron exporter, enhances resistance to oxidative
911 stress in *Escherichia coli*. Appl Environ Microbiol **79:**7210-7219.

912 43. **Robbe-Saule V, Coynault C, Ibanez-Ruiz M, Hermant D, Norel F.** 2001. Identification
913 of a non-haem catalase in *Salmonella* and its regulation by RpoS (σS). Mol Microbiol
914 **39:**1533-1545.

915 44. **Craig M, Slauch J.** 2009. Phagocytic superoxide specifically damages an
916 extracytoplasmic target to inhibit or kill *Salmonella*. Plos One **4:**e4975.

917 45. **Tsolis RM, Bäumler AJ, Heffron F, Stojiljkovic I.** 1996. Contribution of TonB- and Feo-
918 mediated iron uptake to growth of *Salmonella* typhimurium in the mouse. Infect Immun
919 **64:**4549-4556.

920 46. **Aussel L, Zhao W, Hebrard M, Guilhon AA, Viala JPM, Henri S, Chasson L, Gorvel**
921 **JP, Barras F, Meresse S.** 2011. *Salmonella* detoxifying enzymes are sufficient to cope
922 with the host oxidative burst. Mol Microbiol **80:**628-640.
923 47. **Taylor PD, Inchley CJ, Gallagher MP.** 1998. The *Salmonella* typhimurium AhpC
924 polypeptide is not essential for virulence in BALB/c Mice but is recognized as an antigen
925 during infection. Infect Immun **66:**3208-3217.
926 48. **Papp-Szabò E, Firtel M, Josephy PD.** 1994. Comparison of the sensitivities of
927 *Salmonella* typhimurium oxyR and katG mutants to killing by human neutrophils. Infect
928 Immun **62:**2662-2668.
929 49. **Imlay J.** 2015. Diagnosing oxidative stress in bacteria: not as easy as you might think. Curr
930 Opin Microbiol **24:**124-131.
931 50. **Carlsson J, Carpenter VS.** 1980. The recA+ gene product is more important than catalase
932 and superoxide dismutase in protecting *Escherichia coli* against hydrogen peroxide
933 toxicity. J Bacteriol **142:**319-321.
934 51. **Ishioka K, Iwasaki H, Shinagawa H.** 1997. Roles of the recG gene product of *Escherichia*
935 *coli* in recombination repair: Effects of the Delta recG mutation on cell division and
936 chromosome partition. Genes Genet Syst **72:**91-99.
937 52. **Sheikh A, Charles R, Sharmeen N, Rollins S, Harris J, Bhuiyan M, Arifuzzaman M,**
938 **Khanam F, Bukka A, Kalsy A, Porwollik S, Leung D, Brooks W, LaRocque R,**
939 **Hohmann E, Cravioto A, Logvinenko T, Calderwood S, McClelland M, Graham J,**
940 **Qadri F, Ryan E.** 2011. In Vivo Expression of *Salmonella enterica* Serotype Typhi Genes
941 in the Blood of Patients with Typhoid Fever in Bangladesh. PLoS Negl Trop Dis **5:**e1419.
942 53. **Lu S, Killoran PB, Riley LW.** 2003. Association of *Salmonella enterica* Serovar

943 Enteritidis YafD with resistance to chicken egg albumen. Infect Immun **71:**6734-6741.

944 54. **Demple B, Halbrook J, Linn S.** 1983. *Escherichia coli* xth mutants are hypersensitive to
945 hydrogen peroxide. J Bacteriol **153:**1079-1082.

946 55. **Sassetti CM, Rubin EJ.** 2003. Genetic requirements for mycobacterial survival during
947 infection. Proc Natl Acad Sci U S A **100:**12989-12994.

948 56. **Sabbagh S, Lepage C, McClelland M, Daigle F.** 2012. Selection of *Salmonella enterica*
949 Serovar Typhi genes involved during interaction with human macrophages by screening of
950 a transposon mutant library. Plos One **7:**e36643.

951 57. **Bogomolnaya L, Aldrich L, Ragoza Y, Talamantes M, Andrews K, McClelland M,**
952 **Andrews-Polymenis H.** 2014. Identification of novel factors involved in modulating
953 motility of *Salmonella enterica* Serotype Typhimurium. Plos One **9:**e111513.

954 58. **Schwartz CJ, Giel JL, Patschkowski T, Luther C, Ruzicka FJ, Beinert H, Kiley PJ.**
955 2001. IscR, An Fe-S cluster-containing transcription factor, represses expression of
956 *Escherichia coli* genes encoding Fe-S cluster assembly proteins. Proc Natl Acad Sci U S
957 A **98:**14895-14900.

958 59. **Unciuleac M, Chandramouli K, Naik S, Mayer S, Huynh B, Johnson M, Dean D.** 2007.
959 Vitro activation of apo-aconitase using a [4Fe-4S] cluster-loaded form of the IscU [Fe-S]
960 cluster scaffolding protein. Biochemistry **46:**6812-6821.

961 60. **Py B, Barras F.** 2010. Building Fe–S proteins: bacterial strategies. Nat Rev Microbiol
962 **8:**436-446.

963 61. **Balbontín R, Rowley G, Pucciarelli MG, López-Garrido J, Wormstone Y, Lucchini**
964 **S, Portillo FG, Jay C. D. Hinton, Casadesús J.** 2006. DNA adenine methylation regulates
965 virulence gene expression in *Salmonella enterica* Serovar Typhimurium. J Bacteriol

966 **188:**8160-8168.

967 62. **Badie G, Heithoff DM, Sinsheimer RL, Mahan MJ.** 2007. Altered levels of *Salmonella*
968 DNA adenine methylase are associated with defects in gene expression, motility, flagellar
969 synthesis, and bile resistance in the pathogenic strain 14028 but not in the laboratory strain
970 LT2. J Bacteriol **189:**1556-1564.

971 63. **Sarnacki SH, Marolda CL, Llana MN, Giacomodonato MN, Valvano MA, Cerquetti**
972 **MC.** 2009. Dam methylation controls O-antigen chain length in *Salmonella enterica*
973 Serovar Enteritidis by regulating the expression of Wzz protein. J Bacteriol **191:**6694-
974 6700.

975 64. **Lopez-Garrido J, Cheng N, Garcia-Quintanilla F, Garcia-del Portillo F, Casadesus J.**
976 2010. Identification of the *Salmonella enterica* damX gene product, an inner membrane
977 protein involved in bile resistance. J Bacteriol **192:**893-895.

978 65. **Hengge-Aronis R.** 2002. Signal transduction and regulatory mechanisms involved in
979 control of the σS (RpoS) subunit of RNA polymerase. Microbiol Mol Biol Rev **66:**373-
980 395.

981 66. **Lu M, Kleckner N.** 1994. Molecular cloning and characterization of the pgm gene
982 encoding phosphoglucomutase of *Escherichia coli*. J Bacteriol **176:**5847-5851.

983 67. **Paterson GK, Cone DB, Peters SE, Maskell DJ.** 2009. The enzyme phosphoglucomutase
984 (Pgm) is required by *Salmonella enterica* serovar Typhimurium for O-antigen production,
985 resistance to antimicrobial peptides and in vivo fitness. Microbiology **155:**3403-3410.

986 68. **Martínez-García E, Calles B, Arévalo-Rodríguez M, de Lorenzo V.** 2011. pBAM1: an
987 all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes.
988 BMC Microbiol **11:**38-38.

989 69. **Hall B, Acar H, Nandipati A, Barlow M.** 2014. Growth rates made easy. Mol Biol Evol
990 **31:**232-238.
991 70. **Yao Q, Luo Z, Yuan X, Yu Y, Zhang C, Xie J, Lee JY.** 2014. Assembly of nanoions via
992 electrostatic interactions: ion-like behavior of charged noble metal nanoclusters. Sci Rep
993 **4:**3848-3848.
994 71. **Chun J, Kim K, Hwang I, Kim Y, Lee D, Lee I, Kim J.** 2007. Dual priming
995 oligonucleotide system for the multiplex detection of respiratory viruses and SNP
996 genotyping of CYP2C19 gene. Nucleic Acids Res **35:**e40-e40.
997 72. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler
998 transform. Bioinformatics **25:**1754-1760.
999 73. **Langmead B, Trapnell C, Pop M, Salzberg S.** 2009. Ultrafast and memory-efficient
1000 alignment of short DNA sequences to the human genome. Genome Biol **10:**R25-R25.
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011

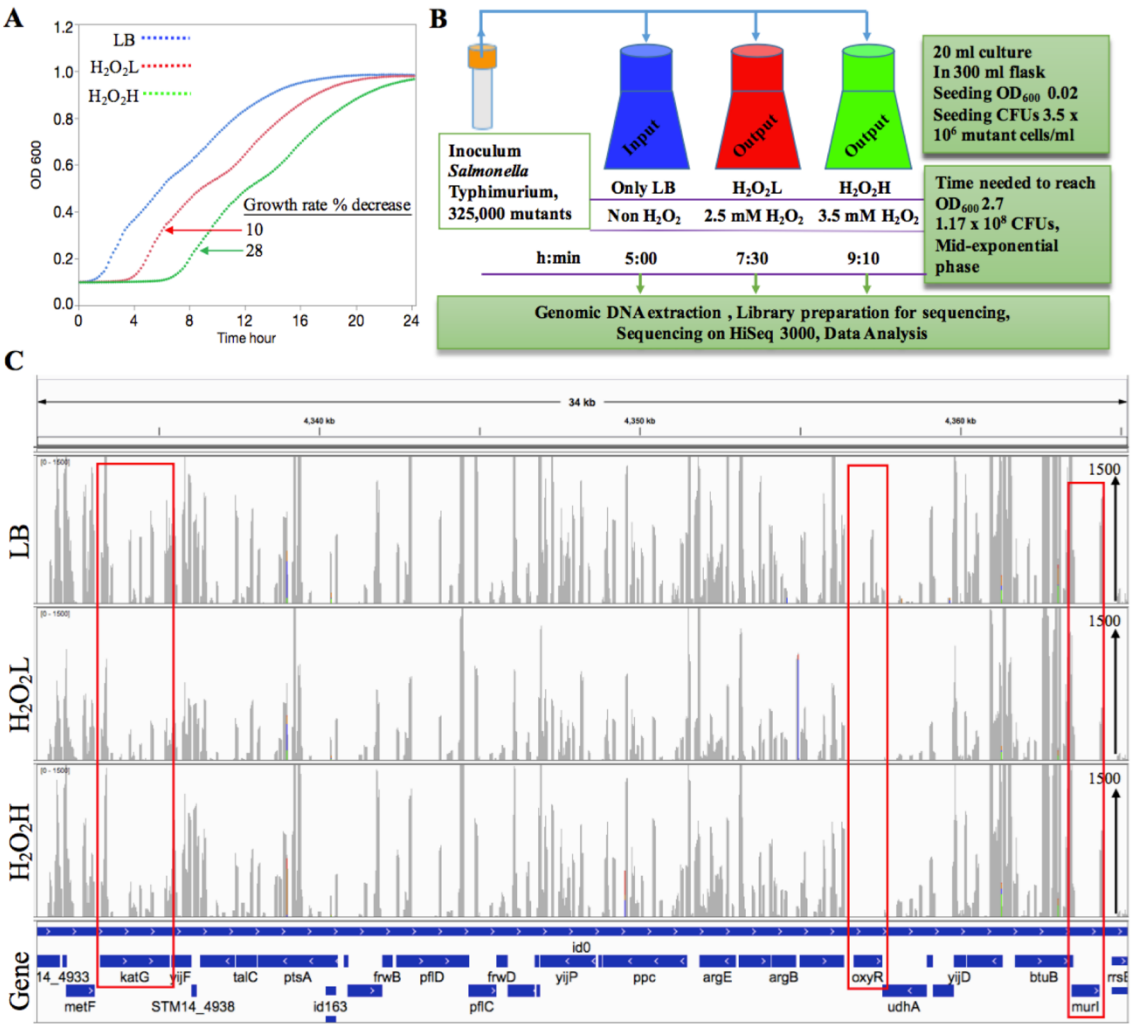
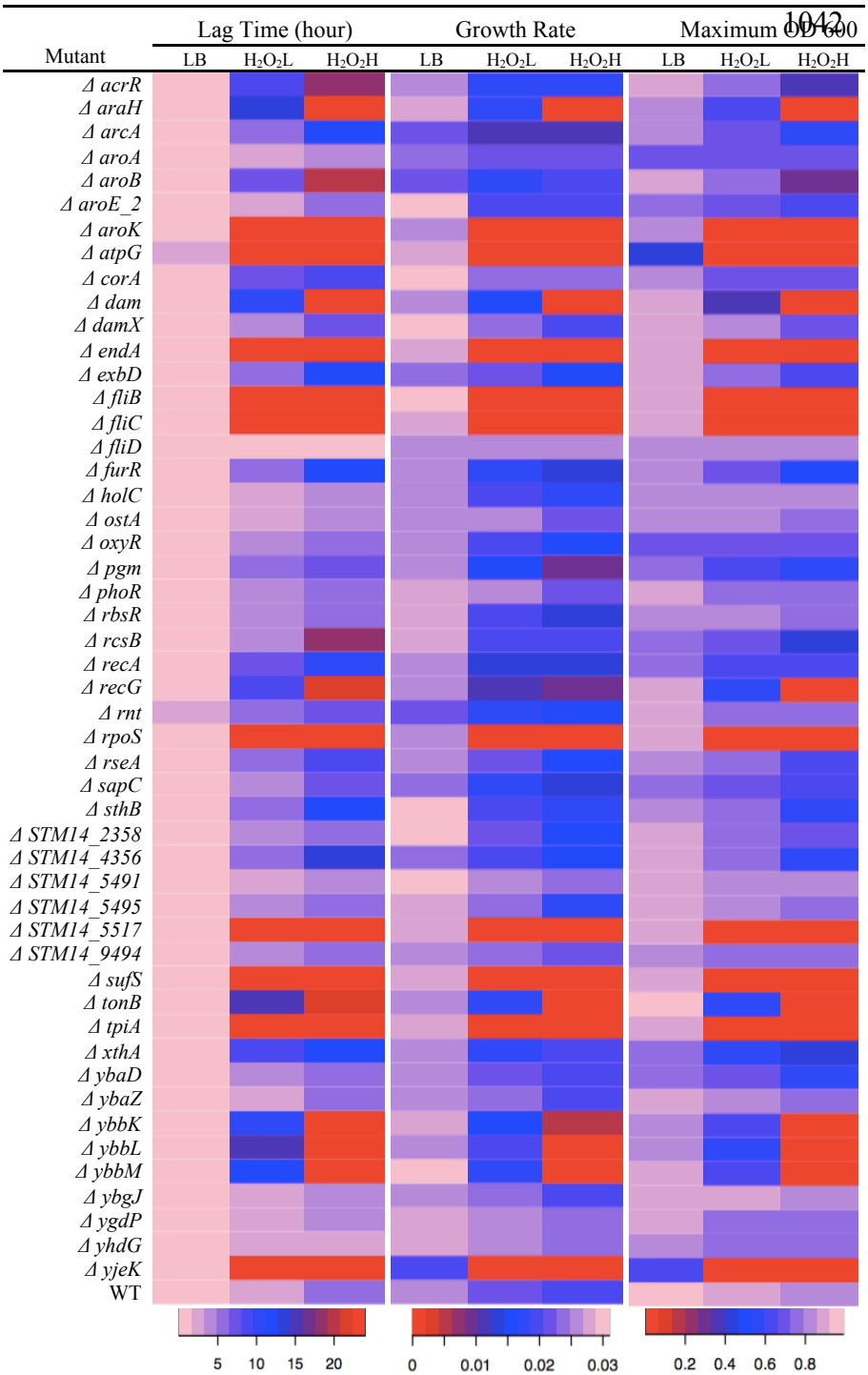


FIG 1

The Tn-seq can identify the genes that are required H_2O_2 resistance. (A) The effect of H_2O_2 on the

1021 growth rate of wild type *Salmonella* Typhimurium. An overnight culture of bacteria was diluted
1022 1:200 in the LB medium contains either 2.5 mM H_2O_2 (H_2O_2L) or 3.5 mM H_2O_2 (H_2O_2H), LB
1023 without H_2O_2 was used as control. The cultures were incubated at 37°C for 24 h in a 96-well plate.
1024 The reduced growth rates for the H_2O_2 were in comparison to the control. In the all growth curve
1025 figures in this work, the blue color represents just LB (no H_2O_2), the red is H_2O_2L, and green is
1026 H_2O_2H. (B) Schematic representation of the Tn-seq study. The *Salmonella* transposon mutant
1027 library was inoculated into LB and LB contains H_2O_2L or H_2O_2H. The three cultures were grown
1028 until they reached mid-exponential phase. The DNA was extracted form each culture and subjected
1029 to library preparation, sequencing, and data analysis. (C) The Tn-seq profile of the three
1030 conditions. It shows 34 kb of the *Salmonella* genome which starts with *metF* gene and ends with
1031 *rrsB*, horizontal axis. The height of vertical axis represents number of reads which is 1500 reads.
1032 The highlighted genes in red are *katG*, catalase peroxidase, was tolerated insertions in the both
1033 H_2O_2 conditions; *oxyR* was not tolerated insertions in presence of H_2O_2 and indicated that the gene
1034 is required to H_2O_2 resistance; *murI*, glutamate racemase, was not tolerated any insertions at all
1035 and considered as essential gene, its required for the biosynthesis of a component of cell wall
1036 peptidoglycan.
1037
1038
1039
1040
1041



1062 FIG 2

1063 Growth curve of 50 mutants and a wild type *Salmonella* in the form of heat map. The lag phase

1064 time, growth rate, and maximum OD_{600} of the individual *Salmonella* Typhimurium mutants and

1065 the wild type in the growth conditions of LB (no H_2O_2), H_2O_2L and H_2O_2H. The overnight culture
1066 of the mutants and the wild type bacteria were diluted 1:200 in the LB medium, and the LB
1067 contains 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
1068 37°C for 24 h in 96-well plate and the OD_{600} was recorded every 10 min. The lag phase time,
1069 growth rate and maximum OD_{600} were calculated and shown here as a graphical representation.
1070 The pale pink color indicates short lag phase time, high growth rate, and high OD_{600}. The red color
1071 indicates that the bacteria was stayed in lag phase, growth rate was close to zero, and the OD_{600} of
1072 the culture was not raised in the 24 h time of the assays.
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087

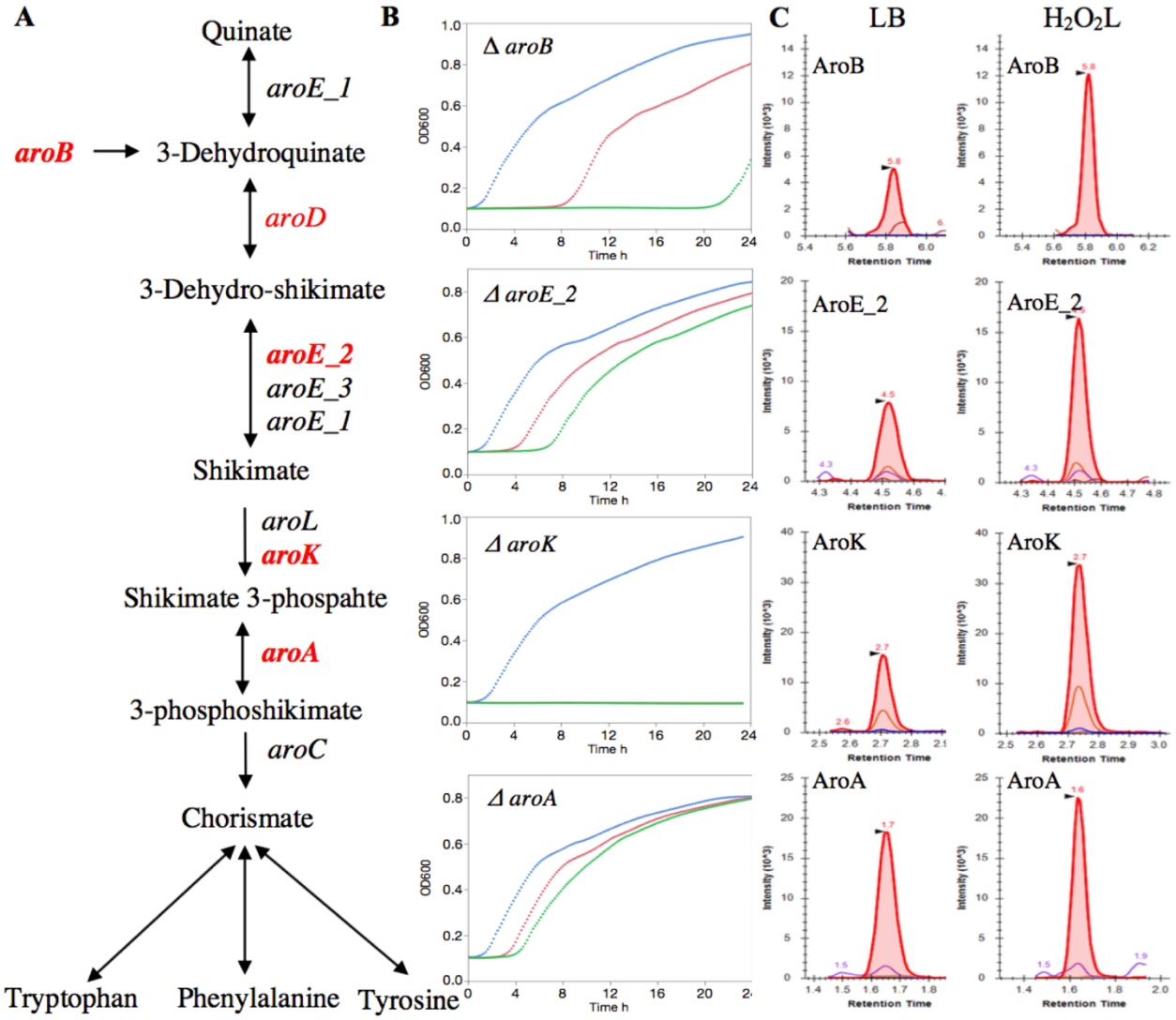


FIG 3

The role of aromatic amino acid biosynthesis genes in resistance to the H_2O_2. (A) Schematic representation of aromatic amino acid biosynthesis, adapted from the KEGG pathway database. The genes in red color were identified by the Tn-seq for H_2O_2 resistance in *Salmonella*. The red bold color genes were identified by the Tn-seq and the phenotypes were validated by the individual mutant assays. (B) The overnight cultures of the individual mutants were diluted 1:200 in the LB (no H_2O_2) and the LB contains either 2.5 mM H_2O_2 (H_2O_2L) or 3.5 mM H_2O_2 (H_2O_2H). The

1096 cultures were incubated at 37°C for 24 h in 96-well plate. The color of growth curve figures is blue
1097 for the LB, red is H_2O_2L, and green is H_2O_2H. In the *Δ aroK* growth curve, the red color is under
1098 the green color. (C) Differential expression of *Salmonella* proteins in response to the H_2O_2L
1099 compared to the LB. Wild type *Salmonella* was grown in LB, H_2O_2L, and H_2O_2H until mid-
1100 exponential phase. Targeted-proteomics was quantified AroB, AroE_2, AroK, and AroA protein
1101 expressions in response to H_2O_2L. The shown peaks represent a unique peptide of the protein.
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117

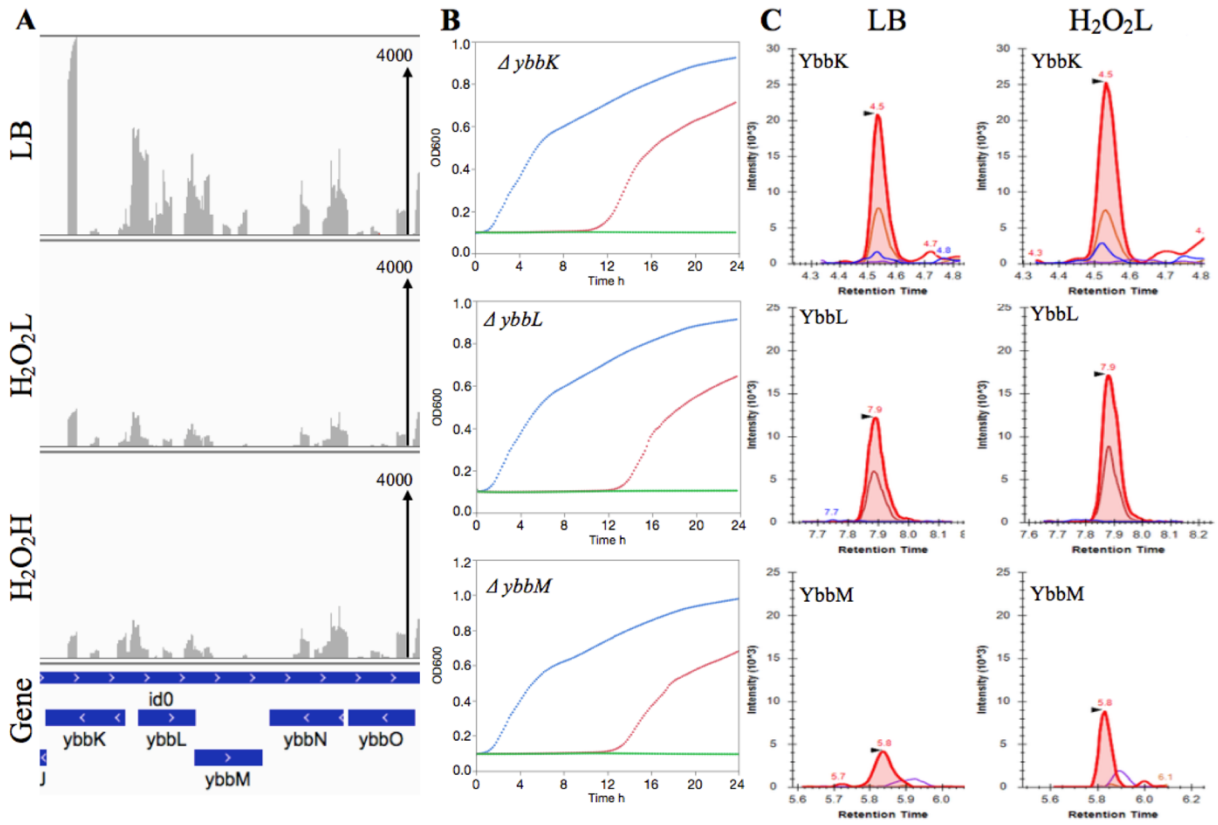


FIG 4

The *ybbK*, *ybbL*, and *ybbM* have a role in resistance to H_2O_2. (A) The Tn-seq profile of the LB (no H_2O_2), H_2O_2L (2.5 mM), and H_2O_2H (3.5 mM). It shows ~6 kb of *Salmonella* Typhimurium genome starts with *ybbK* and ends with *ybbO*, horizontal axis. The read scale for the conditions are 4000, vertical axis. (B) The growth curve of *ΔyybK*, *ΔyybL*, and *ΔyybM*. The overnight cultures of these three mutants were diluted 1:200 in LB, H_2O_2L, and H_2O_2H. The cultures were incubated at 37°C for 24 h in 96-well plate. The growth curve colors are blue which represents just LB or no H_2O_2, the red is H_2O_2L, and green is H_2O_2H. (C) Wild type *Salmonella* was grown in LB, H_2O_2L, and H_2O_2H until mid-exponential phase. Targeted-proteomics was quantified YbbK, YbbL, and YbbM protein expressions in response to H_2O_2L. The shown peaks represent a unique peptide of the protein.