Screening and Identification of Key Genes for the Survival and Multiplication of Salmonella typhimurium in the Host

Na Sun
College of veterinary medicine, Shandong Agricultural University
Yanying Song
College of Veterinary Medicine, SDAU
Cong Liu
College of Veterinary Medicine, SDAU
Yu Dai
College of Veterinary Medicine, SDAU
Peng Wang
College of Veterinary Science, SDAU
Mengda Liu
Laboratory of Zoonoses, China animal health and epideminology center
Lanping Yu
College of Veterinary Science, SDAU
Chenglian Feng
Chinese research academy of environmental sciences
Shuhong Sun
College of Veterinary Science, SDAU
Yingli Shang
College of Veterinary Science, SDAU
Fangkun Wang (wangfangkun@sdau.edu.cn)
Shandong Provincial key laboratory of animal biotechnology and disease control and prevention
https://orcid.org/0000-0001-8844-8974

Research article

Keywords: Salmonella typhimurium, RNA-seq, Comparative transcriptomics, Pathogenic mechanism

Posted Date: December 16th, 2019

DOI: https://doi.org/10.21203/rs.2.18938/v1

Abstract

Background

Salmonella typhimurium is an important intracellular pathogen that poses a health threat to humans. The key to studying the pathogenesis of *Salmonella* is to clarify the mechanisms responsible for its survival and reproduction in macrophages. In this study, RNA was extracted from *S. typhimurium* reference strain (ATCC 14028) and *S. typhimurium* isolated from the spleen of infected mice for RNA high-throughput sequencing analysis, based on the BGISEQ-500 platform.

Results

A total of 1,340 significant differentially expressed genes (**DEGs**) were screened through quantitative gene analysis and various analyses based on gene expression. Of these, 16 genes were randomly selected for fluorescence quantitative PCR verification. Two pairs of primers, *16S* and pyridoxol 5′-phosphate synthase (***pdxJ***), were used as internal references. The coincidence rate was determined to be 93.75%, which was consistent with the RNA-seq data, proving the reliability of the RNA-seq data. Functional annotation revealed DEGs associated with regulation, metabolism, transport and binding, pathogenesis, and motility. Through data mining and literature retrieval, 26 of the 58 upregulated DEGs (FPKM >10) were not reported to be related to the adaptation to intracellular survival, and were classified as candidate key genes (**CKGs**) for survival and proliferation *in vivo*. Among the CKGs, five were biotin synthetic bio family proteins. BioF is one of the first enzymes in the protein synthesis pathway. To evaluate the potential role of *bioF* in survival and proliferation, *bioF* mutants of *Salmonella* were constructed, and the virulence/attenuation was evaluated *in vivo*. Through the infection of BALB/c mice, *bioF* deficiency was found to significantly reduce the bacterial load and the fatality rate of mice.

Conclusions

Our results indicated that the *bioF* gene plays an important role in the survival and proliferation of *S. typhimurium in vivo*. These data contribute to our understanding of the mechanisms used by *Salmonella* to regulate virulence gene expression whilst replicating inside mammalian cells.

1. Background

Salmonella is a Gram-negative pathogen that poses a health threat to humans and livestock. According to the World Health Organization (WHO), salmonellosis is the most common food-borne disease, with nearly one-tenth of the world's population becoming infected each year and around 33 million deaths (http://www.who.int/mediacentre/factsheets/fs139/en/). Among the more than 2600 documented serotypes, *Salmonella typhimurium* is the predominant serotype associated with salmonellosis worldwide [1, 2].

The mechanism of survival and reproduction in macrophages may be key to studying the pathogenesis of *S. typhimurium*. During the infection process, *S. typhimurium* must adjust and adapt to rapidly changing intracellular and extracellular environments: downregulating the expression of non-essential genes, upregulating the expression of genes necessary for survival, and expressing essential virulence genes [3-6]. Many of these adaptive processes are regulated at the transcriptional level [7]. In recent years, a large number of studies have confirmed that many bacteria that induce the high expression of genes *in vivo* are pathogenic [8-12]. In studies investigating the pathogenesis of *Salmonella*, the induction of highly expressed genes *in vivo* is a research focus [13-16].

In this study, we systematically screened and identified the significant differentially expressed genes of intracellular bacteria in host cells. We studied the bacterial transcriptome rather than the proteome to avoid difficulties in separating out host cell proteins, and applied a new generation of high Flux BGISEQ-500 expression profiling [17] to obtain *S. typhimurium* transcriptome data for intracellular and extracellular cultures. RNA sequencing (RNA-seq) analysis showed that the transcription levels of genes were significantly altered in the pathogenic bacteria *in vivo*. We screened and identified key factors in the intracellular proliferation of *S. typhimurium* by comparative transcriptome analysis combined with literature mining. Our results will help to reveal novel mechanisms employed by *S. typhimurium* to adapt and survive *in vivo* and yielded novel candidate key genes (CKGs) for the survival and proliferation of this pathogen *in vivo*.

2. Results

2.1 Gene expression profiles

A total of six samples were analyzed by high-throughput sequencing using the BGISEQ-500 platform, with an average yield of 21.95 M of data per sample. TN6_2A, TN6_3A, and TN6_4A were three *in vivo* experimental replicates, and TW6_2A, TW6_3A, and TW6_4A were three *in vitro* controls. The quality of the filtered reads is shown in the filtered reads quality statistics table (Table 1). After obtaining the clean reads, we used HISAT to compare the clean reads to the reference genome sequence. A total of 4,479 genes were detected. Venn diagrams were generated to quantify the overlap between the three biological replicates. Among them, 4,477 transcripts were detected in the *in vitro* group, and 4,460 genes were identified among the three biological replicates (Fig. 1a). In total, 3,737 transcripts were detected in the *in vivo* group, of which 2,853 genes were identified among the three biological replicates (Fig. 1b). To determine the correlation in gene expression between the samples, the Pearson correlation coefficient of all gene expression levels between each sample was calculated. The correlation coefficient between *in vivo* samples was more than 84.2%, compared with more than 93.3% for the *in vitro* samples (Fig. 2a). To show the number of genes in each FPKM group (FPKM $\leq$ 1, FPKM 1–10, FPKM $\geq$ 10) for each sample, we performed statistical analysis and the results are shown in Fig. 2b.

2.2 Identification of DEGs

To determine the DEGs among the 4,479 genes detected by RNA-seq, the DEGseq software was used, with the difference multiple set to more than twice and a Q-value ≤ 0.001 set as the screening parameter. Comparative transcriptomic analysis (*in vivo* vs*. in vitro*) revealed a total of 1,340 DEGs (Additional file 1: Table S3), of which 633 were upregulated in expression, and 707 were downregulated in expression (Fig. 3).

2.3 Gene ontology (GO) enrichment analysis for DEGs

Gene ontology divided the genes into three major functional categories: molecular function, cellular component, and biological process (Fig. 4a). The molecular function classification results for the DEGs showed that they were mainly distributed into the categories: binding, catalytic activity, transporter activity, and transcription regulator activity. The cellular component classification results for the DEGs showed that they were mainly distributed into the categories: membrane, membrane part, protein-containing complex, and extracellular region. The biological process classification results for the DEGs showed that they were mainly distributed into the categories: cellular process, metabolic process, localization, and response to stimulus.

The enriched bubble map shows the enrichment of GO terms from three dimensions. The figure below shows the GO enrichment results for the DEGs (Fig. 4b). GO functional enrichment analysis of DEGs revealed that the DEGs were significantly enriched in the categories: pathogenesis, interspecies interaction between organisms, multi-organism process, locomotion, and extracellular region(Additional file 2: Table S4).

2.4 Pathway enrichment analysis for DEGs

The gene-involved KEGG metabolic pathway is divided into seven branches: Cellular Processes, Environmental Information Processing, Genetic Information Processing, Human Disease, Metabolism, Organismal Systems, and Drug Development. Further classification statistics are performed under each branch (Additional file 3: Table S5). The KEGG pathway annotation classification results for the DEGs are shown in Fig. 5a.

KEGG pathway enrichment analysis showed that the DEGs were significantly enriched in multiple pathways (Additional file 4: Table S6), such as: microbial metabolism in diverse environments, bacterial chemotaxis, *Salmonella* infection, bacterial secretion system, two-component system, and flagellar assembly (Fig. 5b).

2.5 Validation of the RNA-seq data by qPCR

Sixteen genes were randomly selected for fluorescence quantitative PCR verification (Fig. 6a). Using the *16S* gene (Fig. 6b) and the *pdxJ* gene (Fig. 6c) as internal controls, qRT-PCR analysis was performed to validate the expression profiles of DEGs in *S. typhimurium*-infected mice. Fifteen genes, with the exception of *menG,* were significantly upregulated or downregulated, consistence with the RNA-seq data. The overall coincidence rate was 93.75%.

2.6 Screening of CKGs for survival and proliferation *in vivo*

A further in-depth analysis of the DEGs was conducted, and 58 genes that were more than 10 times upregulated *in vivo* (FPKM >10) were excluded. Through a literature review, we eliminated the genes reported to be associated with intracellular survival and proliferation in *Salmonella*. Finally, 26 of the 58 genes were classified as CKGs for survival and proliferation *in vivo* (Table 2).

2.7 Virulence of *bioF* mutant strains in mice

When mice were challenged with the *invA* and *bioF* deletion strains of *S. typhimurium*, mortality occurred two days later than those challenged with the reference strain (Fig. 7a). No mortalities occurred among the control mice injected with sterile saline. Similar survival rates in mice were observed following challenge with the deletion strains as observed following challenge with a knockout strain of the virulence gene *invA*. Knockout of the *bioF* gene, significantly reduced the mortality rates among mice following bacterial challenge ($P < 0.05$). On day 1, the *S. typhimurium* load in the spleen was significantly lower for the Δ*invA* group than for the D*bioF* group and the wild-type group ($P < 0.05$), and the D*bioF* group showed a lower bacterial load in the spleen than the reference strain group, although this difference was not significant (Fig. 7b). The bacterial loads showed an upward trend with time, but the knockout groups D*bioF* and D*invA* remained significantly lower than the wild-type group ($P < 0.05$).

3. Discussion

In this study, a total of 1,340 DEGs were screened through quantitative gene analysis and were subjected to various analyses based on gene expression levels (principal component, correlation, differential gene screening). These DEGs were further analyzed by GO function significant enrichment analysis, pathway significant enrichment analysis, and clustering analysis, and most of the genes found to be enriched were classified into: cellular processes and catalytic activity, structural constituent of ribosome, structural molecule activity, pathogenesis, locomotion, bacterial chemotaxis, flagellar assembly, and *Salmonella* infection. These DEGs therefore appeared to be functionally closely related to the adaptation of *Salmonella* to *in vivo* microenvironments, ensuring the survival of bacteria, and promoting bacterial proliferation. On this basis, we optimized the screening conditions and screened 58 significantly upregulated genes (FPKM > 10). These DEGs mostly encoded type III secretion system (T3SS) proteins, some enzymes, and the SPI2 effector proteins. The T3SS effectors provide Gram-negative bacteria with a unique virulence mechanism enabling them to inject bacterial effector proteins directly into the host cell cytoplasm, bypassing the extracellular milieu [18]. *Salmonella* replicates within both nonphagocytic epithelial cells and macrophages in *Salmonella*-containing vacuoles (SCVs) [19]. T3SS secretion effectors alter vacuole positioning by acting on host cell actin filaments, microtubule motors, and components of the Golgi complex. Once positioned, maturation is stalled and bacterial replication is initiated. The T3SS effectors could block phagocytosis or promote bacterial invasion of non-phagocytic cells, altering membrane trafficking, and modulating innate and adaptive immune responses [20, 21]. The natural immune system is the first line of defense against bacterial disease [22]. These effectors, which

are injected into host cells through a secretory system (such as T3SS), employ complex and sophisticated strategies to block and control the host's signal transduction pathways, particularly those that have important functions in host innate immunity. For example, *Salmonella* AvrA can be secreted into the host cytoplasm, where it inhibits inflammation and cellular pro-apoptotic responses [23, 24]. SseL protein can inhibit the intracellular NF-kB pathway and enhance the pathogenicity of *Salmonella pullorum* [16]. These activities enable bacteria to survive and replicate, causing disease.

In this study, most of the proteins in the T3SS are upregulated *in vivo*, and their secreted effector proteins such as *SpiC*, *SseB*, *SseC*, *SseD*, *SseE*, *SseF*, *SseL*, and STM1410 also showed significant high levels of expression. It has been confirmed that SifA is an SPI2 effector protein. It plays a vital role in maintaining the integrity of the SCV membrane, and is transported and localized in cells [25]. *SseJ* is necessary for bacterial survival, but *sifA* is not [26]. However, deletion of *sifA* in *S. typhimurium* reduces the proliferation of bacteria in cells and systemic toxicity in mice [27]. Although the role of *sifs* in the pathogenesis of *Salmonella* is unclear, *sifA* mutants are severely attenuated in virulence studies [28].

In the *Sse* family, in addition to *SseA*, which is essential for bacterial survival in cells and a key factor in determining bacterial virulence [29], *SseJ* also regulates SCV membrane dynamics and is associated with SCV and *Sifs* [28], and *SseB*, *SseC*, and *SseD* are located on the bacterial cell membrane and are also essential for the survival of bacteria in cells, being secreted by the SPI2-mediated T3SS and playing a toxic role outside of the bacterial membrane [30].

SsaM is essential for the secretion of SseB, SseC, and SseD. Following the deletion of *SsaM*, the bacteria cannot form *Sifs* and cannot transfer the effector protein *SseJ* to infected cells. The virulence of the deletion strain and its replication ability in infected cells are all decreased [31]; *Spic* also promotes *SseB* and *SseC* secretion. The decrease in virulence caused by *Spic* deletion is not due to a single effector deletion, but results from the deletion of all SPI2 effector proteins. *Spic* inhibits the interaction between SCVs and late endosomes and lysosomes, as well as the endocytosis and recycling of transferrin [32], so *SsaM* and *Spic* are also virulence genes of *S. typhimurium* [25].

SpiA of *S. typhimurium* is critical for the virulence of host cells. After deletion of this gene, phenotypic and biological analysis showed a reduction in growth rate, changes in morphology, and changes in adhesion and invasion of epithelial cells compared with wild-type strains. Similarly, the deletion strain showed a decrease in biofilm formation, as observed by scanning electron microscopy in a quantitative microtiter assay, and the proliferation of the deletion strain in cells was significantly decreased during the biofilm formation stage and the exponential growth period. This suggests that the *spiA* gene is involved in biofilm formation and the virulence of *Salmonella* [33].

To ensure intracellular survival, *Salmonella* must not only avoid removal by the host immune system but also strive to obtain nutrients, which requires the adjustment of gene expression levels to adapt to a rapidly changing environment [6], and these processes are mostly dependent on certain special biological macromolecules. At present, it is still unclear which regulatory networks are affected by effector proteins,

the SPI, and transport processes, and their role in adapting to the SCV microenvironment in the host cell [6].

To date, the biotin synthesis pathway has been studied in detail in *E. coli*, *Bacillus sphaericus*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, and plants [34]. Studies have shown that after *S. typhimurium* infection of macrophages, SPI2-mediated secretion of the T3SS proteins SsaQ and SseE and expression of the biotin biosynthesis proteins BioB and BioD increased. After loss of the *bioB* gene, the ability of the bacteria to proliferate in phagocytes was reduced, indicating that biotin plays a role in the survival of *S. typhimurium* in macrophages [35]. In our experiments, not only *bioB* and *bioD* were found to be upregulated, but also *bioA*, *bioC*, and *bioF*. BioF is an enzyme that catalyzes the first step in the biotin synthesis pathway. Its structure and function have been well studied in microorganisms such as *E. coli*, *B. subtilis*, and *B. sphaericus* [36, 37], but not in *S. typhimurium* [35].

In this study, we used homologous recombination to construct derivatives of *S. typhimurium* in which the *bioF* and *invA* genes were disrupted. *S. typhimurium invA* mutants (Δ*invA*) were included as a control. Gene disruption had no evident effect on growth *in vitro* or on cellular morphology. The parent strains (WT) and their *bioF* or *invA* mutant derivatives were inoculated into mice. The Δ*bioF* strain was significantly attenuated compared with the WT strain. After knocking out the candidate gene *bioF*, the mortality of the mice was significantly reduced ($P < 0.05$). Unlike Δ*invA*, there was no significant difference in the bacterial load in the spleen between the initial stage of Δ*bioF* infection and infection with the parental strain. Consistent with the Δ*invA* strain, the bacterial load was significantly lower with the Δ*bioF* strain than with the parental strain as the infection time became prolonged. The reason for this may be that *invA* is a virulence gene that has been confirmed to be involved in the invasion of the T3SS. The *invA* gene plays a role in the process of bacterial invasion, and deletion of this gene directly reduces the number of invading cells [38]; by contrast, *bioF* is a gene involved in bacterial metabolism and its function does not affect the invasion stage of bacteria pathogenesis. Later on in the infection, the bacterial load was significantly lower with the mutant strain than with the reference strain, suggesting that in a microenvironment of intracellular nutrient deficiency, the loss of *bioF* affects the proliferation of the bacteria. These results indicated that the *bioF* gene plays an important role in the survival and proliferation of *S. typhimurium in vivo*.

In host **SCVs** [39], a variety of metabolic changes, including bacterial oxidation and nitrification, are initiated to accommodate changes in the microenvironment [40]. Physiological, metabolic, and effector protein-mediated adaptation strategies allow the bacteria to replicate within the SCVs, and to form persister cells [39, 41]. Sulfur is an essential element for microorganisms that can be obtained from a variety of compounds, with sulfate being the preferred source [42]. Sulfate uptake is carried out by sulfate permease belonging to the SulT (CysPTWA), SulP, CysP / (PiT), and CysZ families. The main proteins of the sulfur metabolism transport pathway are significantly upregulated in cells, which may be an adaptation to the SCV microenvironment. Among them, CysA is a sulfate/thiosulfate transfer ATP-binding protein, which is involved in the formation of the ABC transporter complex CysA WTP located inside the cell membrane, and is responsible for energy coupling with the transport system. Within the

bacterial cytoplasm, CysH is a phosphate adenosine phosphate reductase (thioredoxin), which is a member of the PAPS reductase family. The three substrates of the CysH enzyme are adenosine 3', 5'-diphosphate, sulfite, and thioredoxin disulfide, and the two products are 3'-adenosyl sulfate and thioredoxin(https://www.prospecbio.com/cysh_ecoli). In a study of the intracellular pathogenic bacteria *Mycobacterium tuberculosis*, it was confirmed that *cysA* deficiency leads to sulfate auxotrophy [43]. The activity of *cysH* can protect bacteria against free radicals during chronic infection, and nitrosation and oxidation exert excitatory resistance, which may be the mechanism that guarantees chronic persistent infection [44]. Nitrosation and oxidative stress play a key regulatory role in inflammation and the immune response. Nitric oxide produced by the host's several immune cell NO synthases has a severe impact on the major carbon metabolic pathways of *Salmonella* [45]. Following *Salmonella* infection in mice, both *cysA* and *cysH* were upregulated, suggesting that *Salmonella* has a similar mechanism of action to *M. tuberculosis* and plays an important role in regulating host cell nitrosation and oxidative stress.

Clusters of antibiotic resistance-related genes were significantly upregulated, such as those encoding the UgtL, VirK, Mig-14, and PagP proteins. The *mig-14* gene regulates bacterial gene expression during the infection of cells and is an important virulence gene of *S. typhimurium* [46]. *VirK* is a gene adjacent to *mig-14* on the *Salmonella* chromosome. It contributes to the regeneration of the bacterial outer membrane in response to the host environment in the late stage of *Salmonella* infection. Like *mig-14*, deletion of *virK* changes the resistance of the bacteria to antimicrobial peptides, such as bacteriocin B, but does not significantly reduce the proliferation of *Salmonella* in infected cells [47]. The synergy between *pagP* and *ugtL* facilitates the continued colonization of *Salmonella* in the gut and the construction of a robust outer membrane barrier [48].

The body temperature of the host can also affect the virulence of the pathogen. Moderate fever can reduce the severity of bacterial and fungal infections; however, *Salmonella* can regulate the expression of its own genes to adapt to a variety of rapidly changing temperatures, thereby allowing it to multiply in various ecosystems. MgtC and MgtB are Mg^{2+} transporters that have been shown to have complex functions and are required for bacterial virulence, flagellar independence, and cellulose synthesis. However, when the host temperature rises, the expression of *mgtC* increases, but *mgtB* is not necessary gene for the cell to adapt to temperature changes [49]. Both PipB and PipB2 are effector proteins that bind to the host cell membrane. Although neither of these genes are required for the proliferation of *Salmonella* in cells, their function is not completely redundant, and PipB2 is a known virulence factor. Studies have shown that *pipB2* deletion strains show reduced virulence after infection in mice [28]. The *Salmonella* virulence factor SspH2 belongs to a growing class of bacterial effector proteins that use and disrupt the eukaryotic ubiquitination pathway, a central eukaryotic regulatory mechanism that controls a variety of biological processes, such as programmed cell death, cell cycle progression, and signal transduction, and when deleted, can lead to cancer and neurodegenerative diseases [50].

Through our literature review of these genes, we obtained a total of 26 DEGs that have not yet been reported to be related to intracellular survival and proliferation. These genes are highly likely to be associated with such functions and are candidate key proteins for intracellular survival and proliferation.

In future studies, we will further elucidate the expression of these proteins in host cells and their possible interaction with host cell proteins, enabling us to explore the immune signaling pathways that are involved *in vivo*. This will help us to clarify key issues such as the pathogenesis of intracellular bacteria and the mechanisms of evading host immune responses. Furthermore, such studies may provide target proteins for vaccine development and drug development, which may aid efforts to control infections by Salmonella as well as other important intracellular pathogens.

4. Conclusion

Comparative transcriptomic analysis of *S. typhimurium in vivo* and *in vitro* has highlighted a series of CKGs for bacterial survival and proliferation *in vivo*, including *bioF*. Inactivation of this gene in *S. typhimurium* ATCC 14028 led to significant attenuation *in vivo*. Future research will need to address the complex role of *bioF* in the life cycle of *S. typhimurium*. Further studies will also be required to identify other CKGs promoting intracellular survival and proliferation.

5. Methods

MPTP-induced PD motor deficits

We evaluated the exercise capacity of MPTP-induced mice by pole test and rotarod test. Fig. 1A, B showed the time required for the movement of the mice in the pole test and rotarod test. In the pole test, MPTP-induced mice were used longer time than the control group (Fig. 1a). Compared with the control group, the MPTP-induced mice had poor coordination and the rotarod test time was significantly shortened (Fig. 1b). This indicates that MPTP-induced PD mice were successfully constructed.

Summary of RNA-Seq data sets

To understand the transcriptome information of different brain regions in PD mice, 24 libraries were constructed in the CC, HP, ST, and CB regions. Subsequently, the libraries were sequenced on the Illumina HiSeq 2500 sequencing platform. The raw reads of each sample ranged from 27.4 to 63.2 million (Additional file 1: Table S1). After quality control and filtering, clean reads were obtained. We compared the clean reads to the reference genome with Hisat2. The reads for each gene were calculated using HTSeq-count. Principal component analysis (PCA) was performed based on the DESeq2 method to analyze whether the same brain region samples were together. The results showed that the four brain regions of the control group were divided into different clusters (Fig. 2a). Compared with the CC and HP regions, the transcriptional characteristics of the CB region were significantly different. In addition, the transcriptome characteristics of the ST were significantly different from those of the CC and HP regions. However, the transcriptome characteristics of the CC and HP regions were not significantly different. The four brain regions of the model group were divided into different clusters (Fig. 2b). The difference between the CB region and the other three brain regions was the most obvious.

Differentially expression gene in different brain regions

The transcriptome characteristics of the model group and control group mice were analyzed. DEG analysis explained the difference between the transcription in the model group and control group. According to the values of fold change > 2 and $P < 0.05$, DEGs were obtained using the DESeq2 R package.

Cluster analysis was performed using the correlation distance and hierarchical algorithm to show gene expression patterns in different brain regions of the control group and model group (Fig. 3a and b). In the same pathological state, the gene expression in the CB was significantly different from that in the CC, HP, and ST regions.

To analyze the effects of gene expression in different brain regions on PD, the DEGs in the model and control groups in the same brain region were compared. As shown in Fig. 3c, 110 DEGs were obtained in the CC region, of which 64 genes were upregulated and 46 genes were downregulated. A total of 179 DEGs were obtained in the HP region, of which 99 genes were upregulated and 80 genes were downregulated. A total of 521 DEGs were obtained in the ST region, of which 485 genes were upregulated and 36 genes were downregulated. There were 96 DEGs in the CB region, of which 67 were upregulated and 29 were downregulated. Fig. 3d shows the overlap of DEGs in the four brain regions. Of the four brain regions, only two genes (*Gxylt1* and *C920006O11 Rik*) were co-expressed.

Differentially expressed genes in HP and ST regions of Parkinson mice

To further understand the transcript information in the HP and ST regions of mice with PD, RNA sequencing analysis was undertaken. Based on the results, unbiased hierarchical clustering analysis was performed for all genes in the HP region. The data of all genes were visualized using a heat map (Fig. 4a). For data quality control, we performed PCA (Fig. 4b). The results showed that the gene expression in the HP region of the model group was different from that in the HP region of the control group. A total of 179 DEGs were found in the HP region in the model group. Additional file 1: Table S2 lists the top 10 upregulated genes and top 10 downregulated genes with the most significant differences in the HP region. In addition, hierarchical clustering analysis of the DEGs showed that there were differences in gene expression in the ST region of the PD model mice (Fig. 4c). Additional file 1: Table S3 lists the top 10 upregulated genes and top 10 downregulated genes with the most significant differences in the ST region.

Compared with the control group, there were 110 (64 upregulated, 46 downregulated) and 96 (67 upregulated, 29 downregulated) DEGs in the CC and CB regions of the model group mice, respectively. Additional file 1: Table S4 and Table S5 listed the top 10 up-regulated genes and the top 10 down-regulated genes in CC and CB regions, respectively. A heat map of the top 100 genes differentially expressed between the control group and model group in the CB region is provided in Additional file 2: Figure S1.

Gene ontology enrichment analysis

Gene ontology enrichment analysis of the DEGs in HP and ST regions

The biological functions of DEGs were identified by GO analysis. The GO analysis database can classify DEGs, including “biological processes,” “cellular components,” and “molecular functions.” The GO term corrected by $P < 0.05$ was defined as significant enrichment of DEGs. In Fig. 5, the significant enrichment of GO terms in the three main categories (cellular component, biological process, and molecular function) is shown. In the HP region, the main five subcategories of biological process were “DNA methylation or demethylation,” “posttranscriptional regulation of gene expression,” “rhythmic process,” “protein localization to organelle,” and “regulation of DNA metabolic process,” and the two subcategories for molecular function were “guanyl-nucleotide exchange factor activity” and “protein serine/threonine kinase activity” (Fig. 5a). In the ST region, the main five subcategories of cell component were “post synapse,” “dendritic tree,” “presynapse,” “perikaryon”, and “endoplasmic reticulum subcompartment.” The five subcategories for biological processes were “positive regulation of nervous system development,” “synapse organization,” “regulation of transmembrane transport,” “regulation of exocytosis,” and “regulation of intracellular transport”; and five subcategories for molecular function, which are “protein domain specific binding,” “protein kinase binding,” “protein heterodimerization activity,” “GABA receptor binding,” and “guanyl-nucleotide binding” (Fig. 5d). Fig. 5b, 5c and 5e, 5f show the top 20 clusters of significantly enriched terms in the HP and ST regions, respectively.

Gene ontology enrichment analysis of the DEGs in CC and CB regions

GO enrichment analysis was performed on DEGs in the CC and CB regions to determine the effects of PD on these two brain regions. The results showed that 10 GO terms were enriched in the CC region (Fig. 6a), in which the cell components were significantly enriched in two GO terms (“axon” and “filopodium”), the biological process was significantly enriched in six GO terms (“negative regulation of protein kinase activity,” “regulation of neuron death,” “negative regulation of cell cycle,” “ribosome biogenesis,” “chromatin organization,” and “organophosphate biosynthetic process”), and the molecular function was significantly enriched in two GO terms (“myosin binding” and “lipase activity”). A total of 14 GO terms were enriched in the CB region (Fig. 6b), in which the cellular components were significantly enriched in two GO terms, the biological process was significantly enriched in nine GO terms, and the molecular function was significantly enriched in three GO terms. In the CC region, the top five significantly enriched GO terms were “axon,” “associative learning,” “myosin binding,” “negative regulation of protein kinase activity,” and “regulation of neuron death” (Fig. 6c). In the CB region, the top five significantly enriched GO terms were “negative regulation of transmembrane transport,” “translation factor activity, RNA binding,” “histone lysine methylation,” “response to glucose,” and “dioxygenase activity” (Fig. 6d).

KEGG enrichment analysis

KEGG pathway analysis of DEGs in the HP and ST regions of the model group helped identify the biological pathways related to DEGs. The $P < 0.05$ pathway was defined as the DEG enrichment pathway. A total of 179 and 521 DEGs in the HP and ST regions were enriched in 122 and 213 KEGG pathways, respectively. Table 1 shows the pathways and related genes that were significantly enriched in the HP

region. In the ST region, the top 10 significantly enriched signaling pathways and related genes are shown in Table 2. There were three signaling pathways (“dopaminergic synapse,” “glutamatergic synapse,” and “adrenergic signaling in cardiomyocytes”), which was consistent with the study by Fu et al. (2019). However, only a few KEGG pathways were enriched in the CC and CB regions. Only the “focal adhesion” and “regulation of actin cytoskeleton” pathways were significantly enriched in the CC region (Table 3), and only the “adrenergic signaling in cardiomyocytes” and “RNA transport” pathways were significantly enriched in the CB region (Table 3).

PPI networks

There were 427 proteins in the PPI network and the nodes represented proteins (Fig. 7a). The most interacting proteins in the PPI network were LRRK2, DA receptor D2 (DRD2), protein kinase A catalytic subunits (PRKACA), PPP2R5C, insulin-like growth factor 1 (IGF-1), PIK3R1, guanine nucleotide binding protein alpha inhibiting 1 (GNAI1), and guanine nucleotide binding protein alpha inhibiting 3 (GNAI3). A total of 15 proteins that may be associated with PD were extracted from the PPI network and the protein network interactions were reconstructed (Fig. 7b).

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

The P value was corrected by FDR, and when FDR $\leq$ 0.01, this was regarded as significant enrichment.

In the knockout of candidate genes, data analysis was performed using T-test analysis.

6. Abbreviations

ATCC	American Type Culture Collection
cDNA	Complementary DNA
CFU	Colony forming unit
CKG	Candidate key gene
DEG	Differentially expressed gene
FDR	False discovery rate
FPKM	Fragments per kilobase of exon model per million reads mapped
GO	Gene ontology
i.p.	Intraperitoneal
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria–Bertani
NF-κB	Nuclear factor-kappa B
NO	Nitric oxide
PBS	Phosphate-buffered solution
pdxJ	Pyridoxol 5-phosphate synthase
RIN	RNA Integrity Number
RNA-seq	RNA sequencing
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription quantitative chain polymerase reaction
SCVs	*Salmonella*-containing vacuoles
SPI2	*Salmonella* Pathogenicity Island 2
T3SS	Type III secretion system
WHO	World Health Organization
WT	Wild-type

7. Declarations

Acknowledgement

We thank the Beijing Genomics Institute (BGI, Shenzhen, China) for the RNA sequencing and its related analysis. We thank Kate Fox, DPhil, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

Authors' contributions

FW, YS. and SS conceived and designed the experiments and analyzed the data. NS, YS, CL, YD, and LY performed the experiments and contributed the reagents. YS, ML, and FW wrote the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported by the Fund of Shandong Agricultural Major Application Technology Innovation (SD2019XM009), National Natural Science Foundation of China (No. 31201929), Natural Science Foundation of Shandong (ZR2016HL44), and the Funds of Shandong "Double Tops" Program (SYL2017YSTD11). These funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. The datasets are also available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The study protocol and all mice experiments were approved by the Animal Care and Use Committee of Shandong Agricultural University, Tai'an, China (SCUC permission no. SDAUA-2015-060), and were performed in strict accordance with experimental animal regulation ordinances defined by China National Science and Technology Commission.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

8. References

1. Gal-Mor O, Boyle EC, Grassl GA. Same species, different diseases: how and why typhoidal and non-typhoidal Salmonella enterica serovars differ. Front Microbiol. 2014; 5:391.
2. Ansari N, Yazdian-Robati R, Shahdordizadeh M, Wang Z, Ghazvini K. Aptasensors for quantitative detection of Salmonella Typhimurium. Anal Biochem. 2017; 533:18-25.
3. Braukmann M, Methner U, Berndt A. Immune reaction and survivability of salmonella typhimurium and salmonella infantis after infection of primary avian macrophages. PLoS One. 2015;10: e0122540.
4. Ricchi M, Mazzarelli A, Piscini A, Di Caro A, Cannas A, Leo S, et al. Exploring MALDI-TOF MS approach for a rapid identification of Mycobacterium avium ssp. paratuberculosis field isolates. J Appl Microbiol. 2017; 122:568-577.
5. Yin J, Chen Y, Xie X, Xia J, Li Q, Geng S, et al. Influence of Salmonella enterica serovar Pullorum pathogenicity island 2 on type III secretion system effector gene expression in chicken macrophage HD11 cells. Avian Pathol. 2017; 46:209-214.
6. Dandekar T, Fieselmann A, Fischer E, Popp J, Hensel M, Noster J. Salmonella-how a metabolic generalist adopts an intracellular lifestyle during infection. Front Cell Infect Microbiol. 2015; 4:191.
7. Shen S, Fang FC. Integrated stress responses in Salmonella. Int J Food Microbiol. 2012; 152:75-81.
8. Chen L, Hu H, Chavda KD, Zhao S, Liu R, Liang H, et al. Complete sequence of a KPC-producing IncN multidrug-resistant plasmid from an epidemic Escherichia coli sequence type 131 strain in China.

Antimicrob Agents Chemother. 2014; 58:2422-2425.

9. Ellis TC, Jain S, Linowski AK, Rike K, Bestor A, Rosa PA, et al. In vivo expression technology identifies a novel virulence factor critical for Borrelia burgdorferi persistence in mice. PLoS Pathog. 2013; 9: e1003567.
10. Hu Y, Huang J, Li Q, Shang Y, Ren F, Jiao Y, et al. Use of in vivo-induced antigen technology to identify in vivo-expressed genes of Campylobacter jejuni during human infection. J Microbiol Biotechnol. 2014; 24:363-370.
11. Casselli T, Bankhead T. Use of in vivo Expression Technology for the Identification of Putative Host Adaptation Factors of the Lyme Disease Spirochete. J Mol Microbiol Biotechnol. 2015; 25:349-361.
12. Ron M, Gorelick-Ashkenazi A, Levisohn S, Nir-Paz R, Geary SJ, Tulman E, et al. Mycoplasma gallisepticum in vivo induced antigens expressed during infection in chickens. Vet Microbiol. 2015; 175:265-274.
13. Hu Y, Cong Y, Li S, Rao X, Wang G, Hu F. Identification of in vivo induced protein antigens of Salmonella enterica serovar Typhi during human infection. Sci China C Life Sci. 2009; 52:942-948.
14. Alam MM, Tsai LL, Rollins SM, Sheikh A, Khanam F, Bufano MK, et al. Identification of in vivo-induced bacterial proteins during human infection with Salmonella enterica serotype Paratyphi A. Clin Vaccine Immunol. 2013; 20:712-719.
15. Geng S, Liu Z, Lin Z, Barrow P, Pan Z, Li Q, et al. Identification of in vivo-induced genes during infection of chickens with Salmonella enterica serovar Enteritidis. Res Vet Sci. 2015; 100:1-7.
16. Geng S, Wang Y, Xue Y, Wang H, Cai Y, Zhang J, et al. The SseL protein inhibits the intracellular NF-kappaB pathway to enhance the virulence of Salmonella Pullorum in a chicken model. Microb Pathog. 2019; 129:1-6.
17. Muller C, Cacaci M, Sauvageot N, Sanguinetti M, Rattei T, Eder T, et al. The Intraperitoneal Transcriptome of the Opportunistic Pathogen Enterococcus faecalis in Mice. PLoS One. 2015;10: e0126143.
18. Coburn B, Sekirov I, Finlay BB. Type III secretion systems and disease. Clin Microbiol Rev. 2007; 20:535-549.
19. Schleker S, Sun J, Raghavan B, Srnec M, Muller N, Koepfinger M, et al. The current Salmonella-host interactome. Proteomics Clin Appl. 2012; 6:117-133.
20. Pinaud L, Sansonetti PJ, Phalipon A. Host Cell Targeting by Enteropathogenic Bacteria T3SS Effectors. Trends Microbiol. 2018; 26:266-283.
21. Jennings E, Thurston TLM, Holden DW. Salmonella SPI-2 Type III Secretion System Effectors: Molecular Mechanisms And Physiological Consequences. Cell Host Microbe. 2017; 22:217-231.
22. Shao F. Mechanisms of Interactions between Bacterial Pathogen and Host Innate Immune Defense System. China Basic Science. 2012; 1:7-13.
23. Giacomodonato MN, Noto Llana M, Aya Castañeda MdR, Buzzola FR, Sarnacki SH, Cerquetti MC. AvrA effector protein of Salmonella enterica serovar Enteritidis is expressed and translocated in

mesenteric lymph nodes at late stages of infection in mice. Microbiology. 2014; 160:1191-1199.

24. Jones RM, Wu H, Wentworth C, Luo L, Collier-Hyams L, Neish AS. Salmonella AvrA Coordinates Suppression of Host Immune and Apoptotic Defenses via JNK Pathway Blockade. Cell Host Microbe. 2008; 3:233-244.
25. Freeman JA, Rappl C, Kuhle V, Hensel M, Miller SI. SpiC is required for translocation of Salmonella pathogenicity island 2 effectors and secretion of translocon proteins SseB and SseC. J Bacteriol. 2002; 184:4971-4980.
26. Nunez-Hernandez C, Tierrez A, Ortega AD, Pucciarelli MG, Godoy M, Eisman B, et al. Genome expression analysis of nonproliferating intracellular Salmonella enterica serovar Typhimurium unravels an acid pH-dependent PhoP-PhoQ response essential for dormancy. Infect Immun. 2013; 81:154-165.
27. Freeman JA, Ohl ME, Miller SI. The Salmonella enterica serovar typhimurium translocated effectors SseJ and SifB are targeted to the Salmonella-containing vacuole. Infect Immun. 2003; 71:418-427.
28. Knodler LA, Vallance BA, Hensel M, Jackel D, Finlay BB, Steele-Mortimer O. Salmonella type III effectors PipB and PipB2 are targeted to detergent-resistant microdomains on internal host cell membranes. Mol Microbiol. 2003; 49:685-704.
29. Zurawski DV, Stein MA. SseA acts as the chaperone for the SseB component of the Salmonella Pathogenicity Island 2 translocon. Mol Microbiol. 2003; 47:1341-1351.
30. Klein JR, Jones BD. Salmonella pathogenicity island 2-encoded proteins SseC and SseD are essential for virulence and are substrates of the type III secretion system. Infect Immun. 2001; 69:737-743.
31. Yu XJ, Liu M, Holden DW. SsaM and SpiC interact and regulate secretion of Salmonella pathogenicity island 2 type III secretion system effectors and translocators. Mol Microbiol. 2004; 54:604-619.
32. Yu XJ, Ruiz-Albert J, Unsworth KE, Garvis S, Liu M, Holden DW. SpiC is required for secretion of Salmonella Pathogenicity Island 2 type III secretion system proteins. Cell Microbiol. 2002; 4:531-540.
33. Dong H, Peng D, Jiao X, Zhang X, Geng S, Liu X. Roles of the spiA gene from Salmonella enteritidis in biofilm formation and virulence. Microbiology. 2011;157: 1798-1805.
34. Fan S, Li DF, Wang DC, Fleming J, Zhang H, Zhou Y, et al. Structure and function of Mycobacterium smegmatis 7-keto-8-aminopelargonic acid (KAPA) synthase. Int J Biochem Cell Biol. 2015; 58:71-80.
35. Shi L, Ansong C, Smallwood H, Rommereim L, McDermott JE, Brewer HM, et al. Proteome of Salmonella Enterica Serotype Typhimurium Grown in a Low Mg/pH Medium. J Proteomics Bioinform. 2009; 2:388-97.
36. Kubota T, Izumi Y. Detection and characterization of a thermophilic biotin biosynthetic enzyme, 7-keto-8-aminopelargonic acid synthase, from various thermophiles. Biosci Biotechnol Biochem. 2012; 76:685-690.
37. Fan S, Li D, Fleming J, Hong Y, Chen T, Zhou L, et al. Purification and X-ray crystallographic analysis of 7-keto-8-aminopelargonic acid (KAPA) synthase from Mycobacterium smegmatis. Acta Crystallogr F Struct Biol Commun. 2014; 70:1372-1375.

38. Galan JE, Ginocchio C, Costeas P. Molecular and functional characterization of the Salmonella invasion gene invA: homology of InvA to members of a new protein family. J Bacteriol. 1992; 174:4338-4349.
39. Pucciarelli MG, Garcia-Del Portillo F. Salmonella Intracellular Lifestyles and Their Impact on Host-to-Host Transmission. Microbiol Spectr. 2017;5.
40. Rowley G, Hensen D, Felgate H, Arkenberg A, Appia-Ayme C, Prior K, et al. Resolving the contributions of the membrane-bound and periplasmic nitrate reductase systems to nitric oxide and nitrous oxide production in Salmonella enterica serovar Typhimurium. Biochem J. 2012; 441:755-762.
41. Helaine S, Holden DW. Heterogeneity of intracellular replication of bacterial pathogens. Curr Opin Microbiol. 2013; 16:184-191.
42. Aguilar-Barajas E, Diaz-Perez C, Ramirez-Diaz MI, Riveros-Rosas H, Cervantes C. Bacterial transport of sulfate, molybdate, and related oxyanions. Biometals. 2011; 24:687-707.
43. Zolotarev AS, Unnikrishnan M, Shmukler BE, Clark JS, Vandorpe DH, Grigorieff N, et al. Increased sulfate uptake by E. coli overexpressing the SLC26-related SulP protein Rv1739c from Mycobacterium tuberculosis. Comp Biochem Physiol A Mol Integr Physiol. 2008; 149:255-266.
44. Senaratne RH, De Silva AD, Williams SJ, Mougous JD, Reader JR, Zhang T, et al. 5'-Adenosinephosphosulphate reductase (CysH) protects Mycobacterium tuberculosis against free radicals during chronic infection phase in mice. Mol Microbiol. 2006; 59:1744-1753.
45. Richardson AR, Payne EC, Younger N, Karlinsey JE, Thomas VC, Becker LA, et al. Multiple targets of nitric oxide in the tricarboxylic acid cycle of Salmonella enterica serovar typhimurium. Cell Host Microbe. 2011; 10:33-43.
46. Valdivia RH, Falkow S. Fluorescence-based isolation of bacterial genes expressed within host cells. Science. 1997; 277:2007-2011.
47. Detweiler CS, Monack DM, Brodsky IE, Mathew H, Falkow S. virK, somA and rcsC are important for systemic Salmonella enterica serovar Typhimurium infection and cationic peptide resistance. Mol Microbiol. 2003; 48:385-400.
48. Goto R, Miki T, Nakamura N, Fujimoto M, Okada N. Salmonella Typhimurium PagP- and UgtL-dependent resistance to antimicrobial peptides contributes to the gut colonization. PLoS One. 2017;12: e0190095.
49. Gall AR, Hegarty AE, Datsenko KA, Westerman RP, SanMiguel P, Csonka LN. High-level, constitutive expression of the mgtC gene confers increased thermotolerance on Salmonella enterica serovar Typhimurium. Mol Microbiol. 2018; 109:327-344.
50. Ibarra JA, Steele-Mortimer O. Salmonella the ultimate insider. Salmonella virulence factors that modulate intracellular survival. Cell Microbiol. 2009; 11:1579-1586.
51. Wong TY, Hall JM, Nowak ES, Boehm DT, Gonyar LA, Hewlett EL, et al. Analysis of the In Vivo Transcriptome of Bordetella pertussis during Infection of Mice. mSphere. 2019;4.
52. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006; 7:3.

53. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015; 12:357-360.
54. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9:357-359.
55. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011; 12:323.
56. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015; 33:290-295.
57. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11: 1650-1667.
58. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics. 2010; 26:136-138.
59. Rao X, Huang X, Zhou Z, Lin X. An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. Biostat Bioinforma Biomath. 2013; 3:71-85.
60. Curtiss R 3rd, Wanda SY, Gunn BM, Zhang X, Tinge SA, Ananthnarayan V, et al. Salmonella enterica serovar typhimurium strains with regulated delayed attenuation in vivo. Infect Immun. 2009; 77:1071-1082.
61. Browne SH, Lesnick ML, Guiney DG. Genetic requirements for salmonella-induced cytopathology in human monocyte-derived macrophages. Infect Immun. 2002; 70:7126-7135.
62. Yu XJ, Grabe GJ, Liu M, Mota LJ, Holden DW. SsaV Interacts with SsaL to Control the Translocon-to-Effector Switch in the Salmonella SPI-2 Type Three Secretion System. MBio. 2018;9.
63. Yoshida Y, Miki T, Ono S, Haneda T, Ito M, Okada N. Functional characterization of the type III secretion ATPase SsaN encoded by Salmonella pathogenicity island 2. PLoS One. 2014;9: e94347.
64. Ge S, He Q, Granfors K. HLA-B27 modulates intracellular growth of Salmonella pathogenicity island 2 mutants and production of cytokines in infected monocytic U937 cells. PLoS One. 2012;7: e34093.
65. Cooper CA, Mulder DT, Allison SE, Pilar AV, Coombes BK. The SseC translocon component in Salmonella enterica serovar Typhimurium is chaperoned by SscA. BMC Microbiol. 2013; 13:221.
66. Dai S, Zhou D. Secretion and Function of Salmonella SPI-2 Effector SseF Require Its Chaperone, SscB. Journal of Bacteriology. 2004; 186:5078-5086.
67. Guy RL, Gonias LA, Stein MA. Aggregation of host endosomes by Salmonella requires SPI2 translocation of SseFG and involves SpvR and the fms-aroE intragenic region. Mol Microbiol. 2000; 37:1417-1435.
68. Hensel M, Shea JE, Waterman SR, Mundy R, Nikolaus T, Banks G, et al. Genes encoding putative effector proteins of the type III secretion system of Salmonella pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. Mol Microbiol. 1998; 30:163-174.
69. Bhowmick PP, Devegowda D, Ruwandeepika HA, Karunasagar I, Karunasagar I. Presence of Salmonella pathogenicity island 2 genes in seafood-associated Salmonella serovars and the role of

the sseC gene in survival of Salmonella enterica serovar Weltevreden in epithelial cells. Microbiology. 2011; 157:160-168.

70. Maserati A, Fink RC, Lourenco A, Julius ML, Diez-Gonzalez F. General response of Salmonella enterica serovar Typhimurium to desiccation: A new role for the virulence factors sopD and sseD in survival. PLoS One. 2017;12: e0187692.
71. Guy RL, Gonias LA, Stein MA. Aggregation of host endosomes by Salmonella requires SPI2 translocation of SseFG and involves SpvR and the fms-aroE intragenic region. Mol Microbiol. 2000; 37:1417-1435.
72. Feng ZZ, Jiang AJ, Mao AW, Feng Y, Wang W, Li J, et al. The Salmonella effectors SseF and SseG inhibit Rab1A-mediated autophagy to facilitate intracellular bacterial survival and replication. J Biol Chem. 2018; 293:9662-9673.
73. Rytkönen A, Poh J, Garmendia J, Boyle C, Thompson A, Liu M, et al. SseL, a Salmonella deubiquitinase required for macrophage killing and virulence. PNAS. 2007; 104:3502–3507.
74. Choi J, Groisman EA. Activation of master virulence regulator PhoP in acidic pH requires the Salmonella-specific protein UgtL. Sci Signal. 2017;10.
75. Mouslim C, Groisman EA. Control of the Salmonella ugd gene by three two-component regulatory systems. Mol Microbiol. 2003; 47:335-344.
76. Quezada CM, Hicks SW, Galan JE, Stebbins CE. A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. Proc Natl Acad Sci U S A. 2009; 106:4864-4869.
77. Baker SJ, Gunn JS, Morona R. The Salmonella typhi melittin resistance gene pqaB affects intracellular growth in PMA-differentiated U937 cells, polymyxin B resistance and lipopolysaccharide. Microbiology. 1999; 145:367-378.
78. Detweiler CS, Monack DM, Brodsky IE, Mathew H, Falkow S. virK, somA and rcsC are important for systemic Salmonella enterica serovar Typhimurium infection and cationic peptide resistance. Molecular Microbiology 2003; 48:385–400.
79. Brodsky IE, Ghori N, Falkow S, Monack D. Mig-14 is an inner membrane-associated protein that promotes Salmonella typhimurium resistance to CRAMP, survival within activated macrophages and persistent infection. Mol Microbiol. 2005; 55:954-972.
80. Park M, Nam D, Kweon DH, Shin D. ATP reduction by MgtC and Mg(2+) homeostasis by MgtA and MgtB enables Salmonella to accumulate RpoS upon low cytoplasmic Mg(2+) stress. Mol Microbiol. 2018;110: 283-295.
81. Tan Z, Chekabab SM, Yu H, Yin X, Diarra MS, Yang C, et al. Growth and Virulence of Salmonella Typhimurium Mutants Deficient in Iron Uptake. ACS Omega. 2019; 4:13218-13230.
82. Zhang Z, Du W, Wang M, Li Y, Su S, Wu T, et al. Contribution of the colicin receptor CirA to biofilm formation, antibotic resistance, and pathogenicity of Salmonella Enteritidis. J Basic Microbiol. 2019; doi: 10.1002/jobm.201900418.

83. Schroeder N, Henry T, de Chastellier C, Zhao W, Guilhon AA, Gorvel JP, et al. The virulence protein SopD2 regulates membrane dynamics of Salmonella-containing vacuoles. PLoS Pathog. 2010;6: e1001002.
84. Knodler LA, Steele-Mortimer O. The Salmonella effector PipB2 affects late endosome/lysosome distribution to mediate Sif extension. Mol Biol Cell. 2005; 16:4108-4123.

Tables

Table 1 Filtered reads quality statistics of the *in vivo* and *in vitro* libraries

Sample	Total Raw Reads (M)	Total Clean Reads (M)	Total Clean Bases (Gb)	Clean Reads Q20(%)	Clean Reads Q30(%)	Clean Reads Ratio(%)
TN6_2A	21.85	21.85	1.09	98.66	91.79	100
TN6_3A	21.85	21.85	1.09	98.66	91.8	100
TN6_4A	21.84	21.84	1.09	98.53	91.34	100
TW6_2A	21.74	21.74	1.09	98.88	92.36	100
TW6_3A	21.68	21.68	1.08	98.93	92.55	100
TW6_4A	22.74	22.74	1.14	98.91	92.44	100

Note: TN6_2A,TN6_3A,TN6_4A represents *in vivo* group 1, group 2 and group 3, respectively. TW6_2A,TW6_3A,TW6_4A represents *in vitro* group 1, group 2 and group 3, respectively.

Table 2 The classification of upregulated CKGs

Gene description	Number	Gene symbol
Type III secretion system associated protein	17	*Ssa* (**D**, **H**, **I**, J[61], L[62], M[31], N[63], **P**, S[64], **T**, **U**, and V [61, 62]，*Ssc*A[64, 65], *Ssc*B[66, 67], *sseA* [29, 68]), ***yscR***, *spiA* [33];
Pathogenicity island2 secreted effector protein	7	*spiC* [25, 32], ***sseB[68]***, *sseC*[69], *sseD*[70], *sseE* [68, 71], *sseF*[68, 72], ***STM1410;***
Hypothetical protein	7	***SteC***, *SseL*[16, 73], ***STM2138***, ***yncJ***, ***SsaK***, ***PagO***, ***pagC***
Enzyme	12	***BioF***, *bioB* [35], ***bioA***, ***bioD***, *ugtl* [48, 74], *ugd* [75]，***STM1952***, ***cysH***, *sspH2* [76], ***entC***, ***Fe-S***, *pqaB*[77];
Antimicrobial resistance protein	3	*virK*[78], *mig-14*[79], *pagP* [48];
Transporter	5	*mgtC* [49], *mgtB*[80], ***rbsA***, ***cysA***, ***fruB***;
Receptor protein	3	*fepA*[81], *cirA*[82], *iroN* [81]
Cold shock protein	1	***cspH;***
Effector protein	3	***sifB***, *pipB2* [28, 83], *pipB*[84];

Note: Attached to the literature are the existing studies that have proved to be virulence factors. Bold were CKGs.

Figures

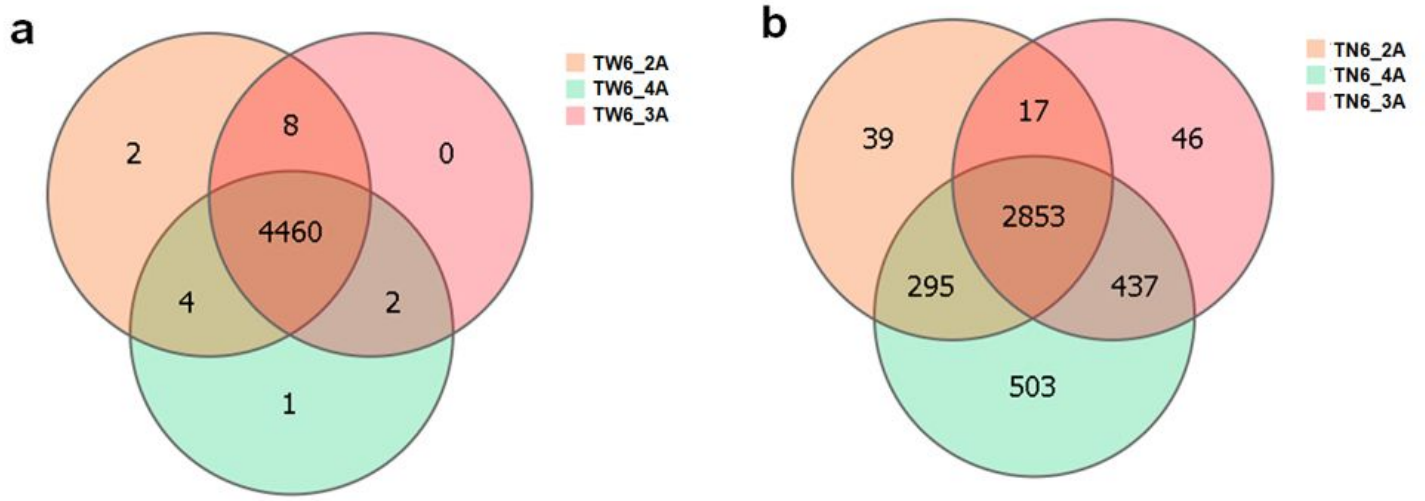

Figure 1

Venn diagrams showing gene expression profiles Venn diagrams were created to quantify the overlap between the three biological replicates. (a) 4,477 transcripts were detected the in vitro group, and 4,460 genes were identified among the three biological replicates. (b) 3,737 transcripts were detected in the in vivo group, of which 2,853 genes were identified among the three biological replicates.

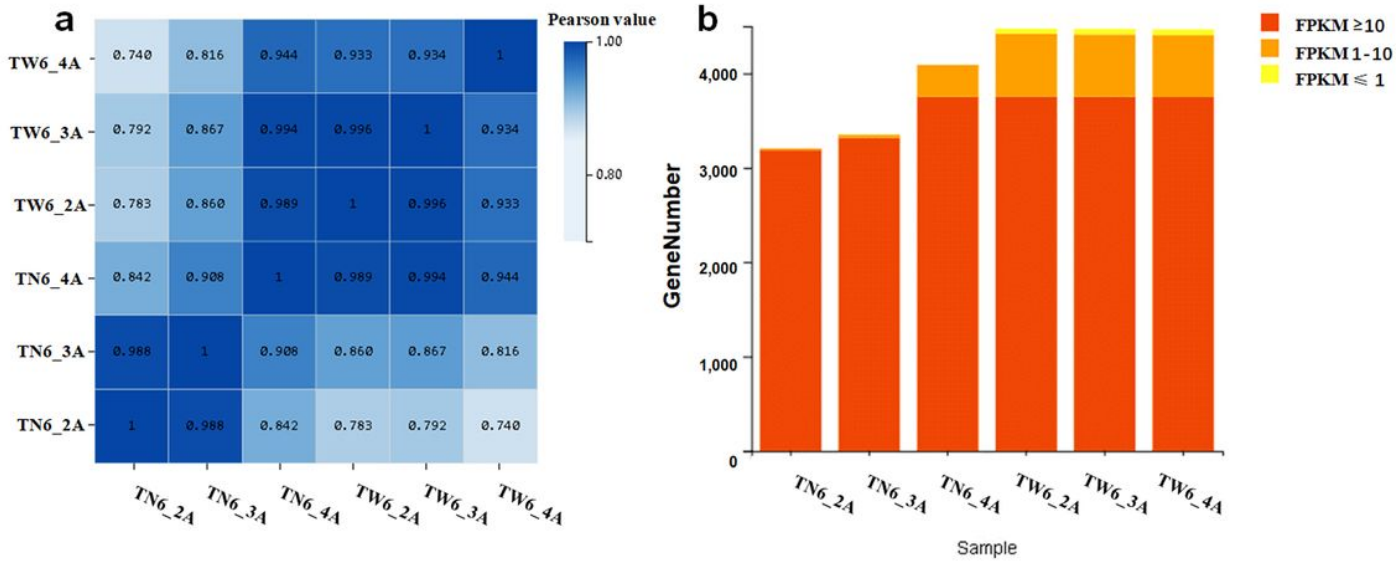

Figure 2

Correlation heatmap and expression quantity accumulation diagram (a) The X and Y axes represent each sample. The color represents the correlation coefficient (the darker the color, the higher the correlation; the lighter the color, the lower the correlation). The correlation coefficient between the in vivo samples was more than 84.2%, and between the in vitro samples was more than 93.3%. (b) X axis represents the name of the sample, Y axis represents the number of genes. The depth of color represents different expression

levels: genes with a very low expression level of FPKM ≤ 1, genes with a low expression level of FPKM between 1 and 10, and genes with a medium or high expression level of FPKM ≥ 10.

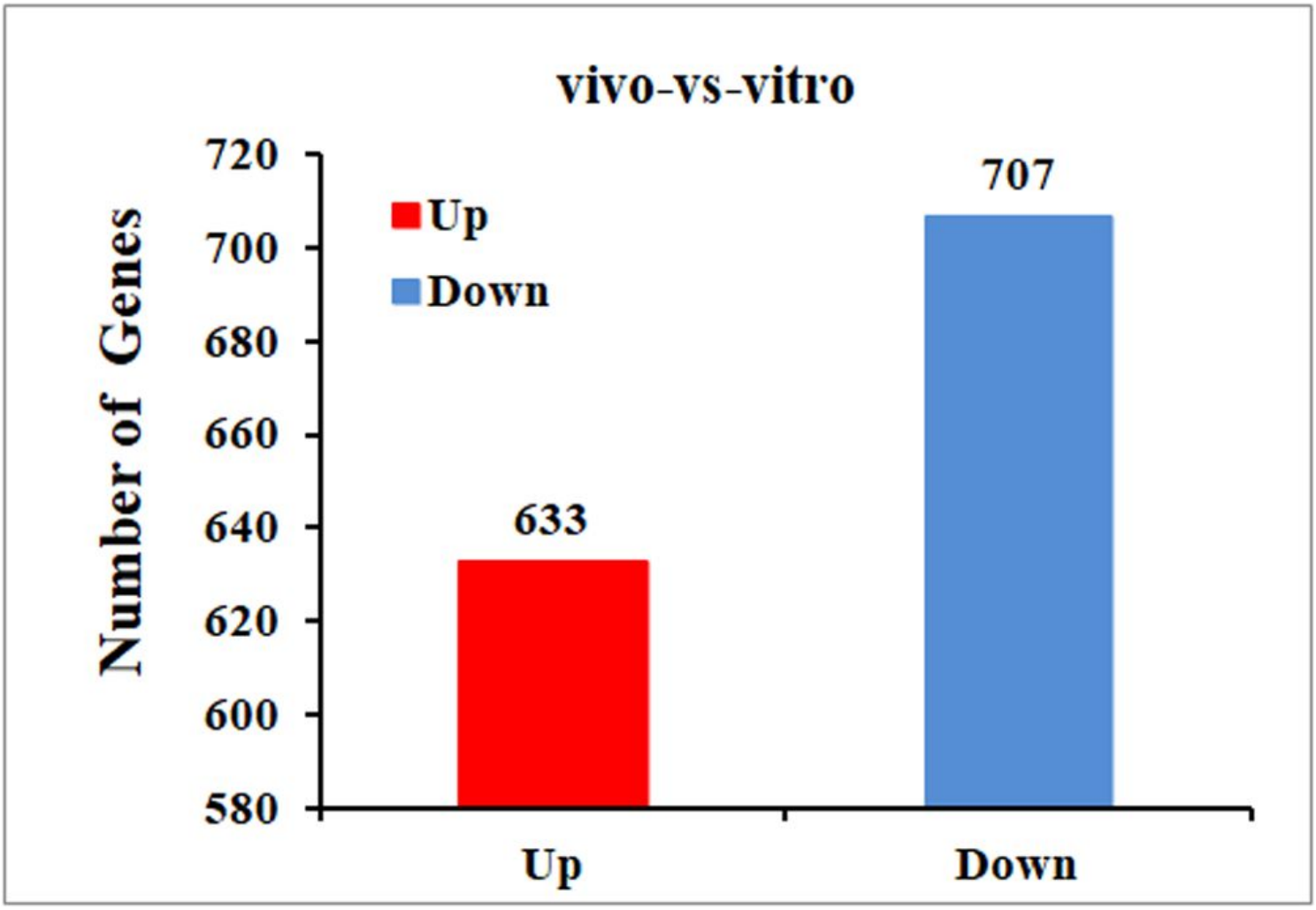

Figure 3

Differentially expressed genes between the in vivo and in vitro groups Comparative transcriptomic analysis (in vivo/in vitro) showed a total of 1,340 DEGs, of which 633 were upregulated in expression and 707 were downregulated in expression.

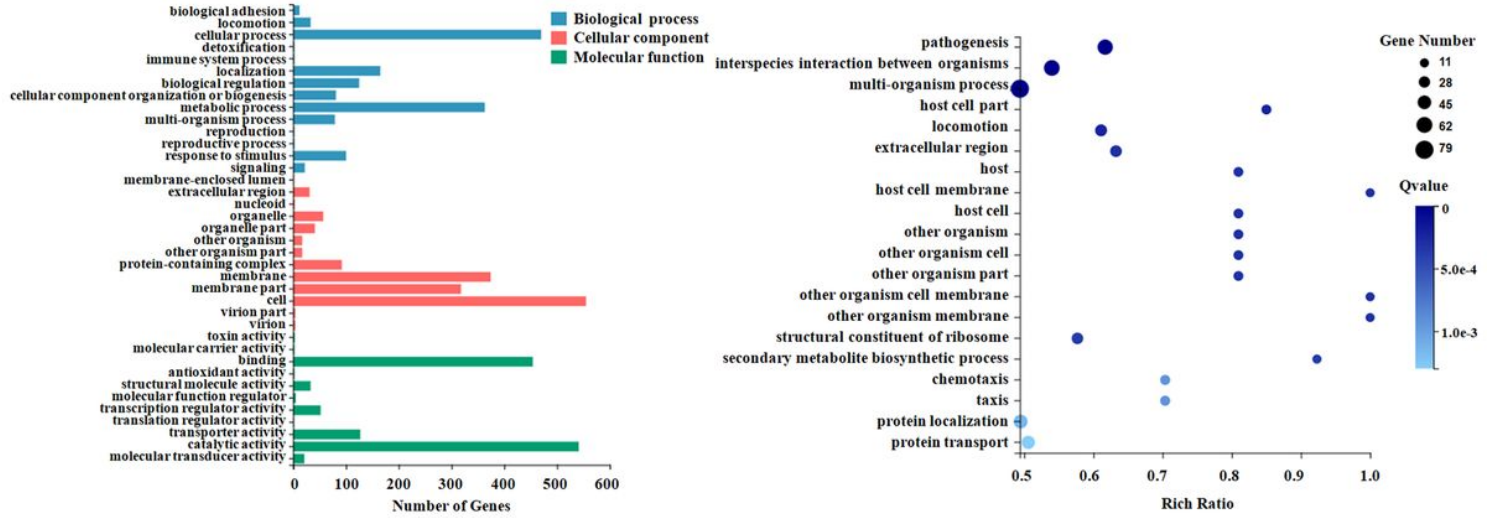

Figure 4

GO classification and enrichment analysis of the DEGs (a) The X axis represents the number of genes annotated for the GO entry, and the Y axis represents the GO functional classification. (b) The X axis represents the enrichment ratio, and the Y axis represents the GO term. The size of the bubble indicates the number of differential genes annotated to a certain GO term. The color represents the enriched Qvalue value. The darker the color, the smaller the Qvalue value.

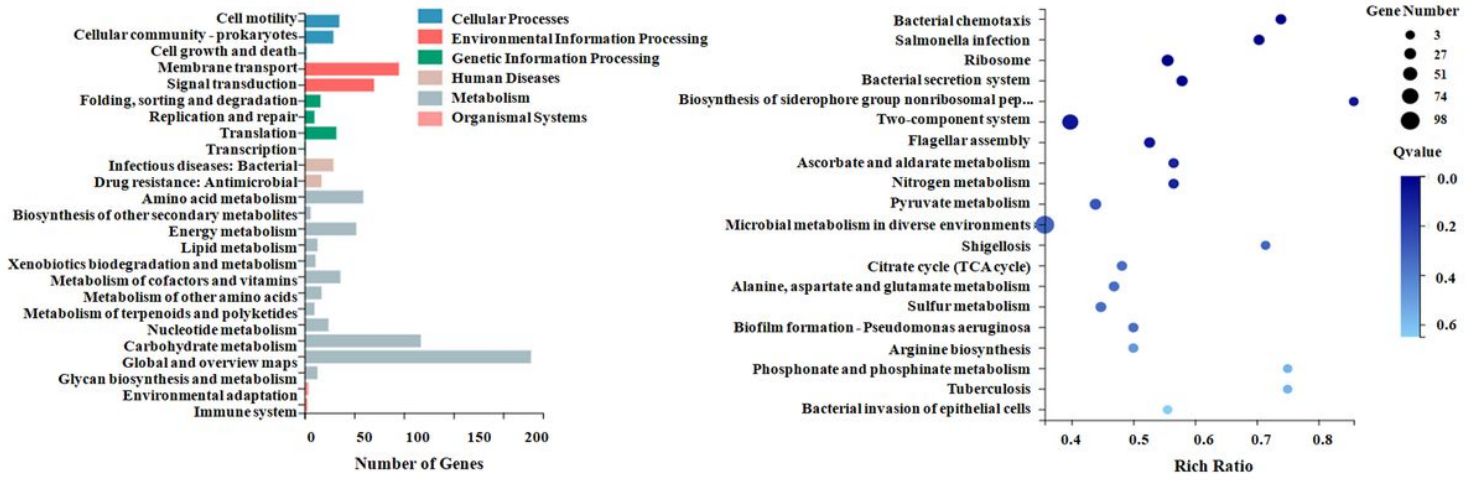

Figure 5

KEGG pathway categories and enrichment analysis (a) The gene-involved KEGG metabolic pathway is divided into seven branches. The X axis represents the number of genes annotated in a KEGG pathway category, and the Y axis represents the KEGG pathway category. (b) The length of the column on the bottom of the X axis represents the size of the Qvalue (−log10(Qvalue)), and the value of the points on the broken line on the top X is the number of DEGs annotated on the KEGG pathway.

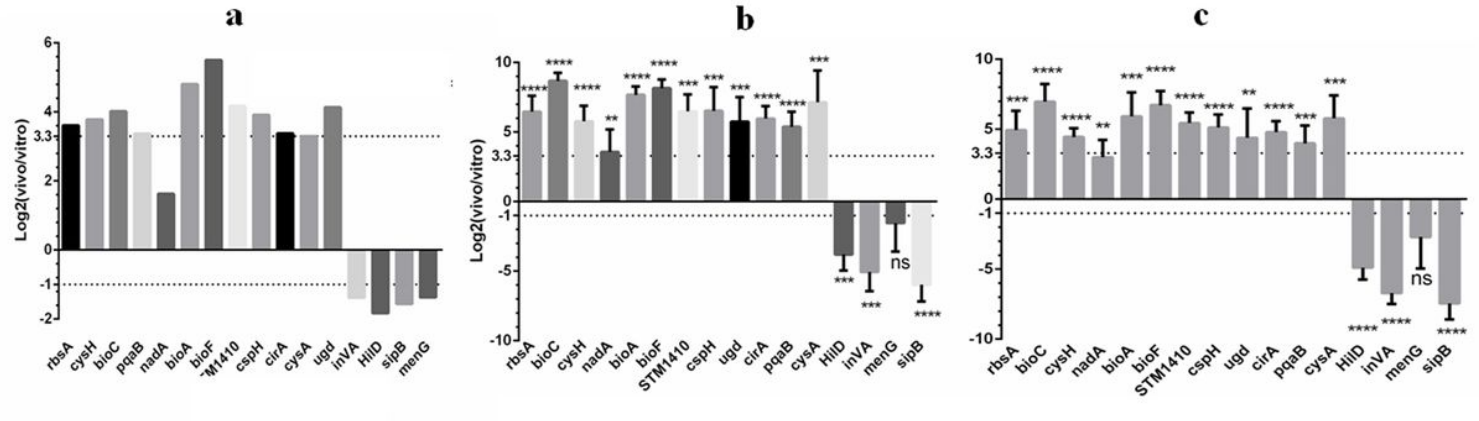

Figure 6

6 qPCR validation of the RNA-seq data (a) Results are expressed as the target ratio of each gene to the log2 (in vivo/in vitro) (b) Relative qPCR was used to measure changes in target gene expression in vivo relative to the in vitro expression levels. Results are expressed as the target ratio of each gene to the log2 (in vivo/in vitro) compared with the reference gene 16S. (c) Relative qPCR was used to measure changes in target gene expression in vivo relative to in vitro expression levels. Results are expressed as the target ratio of each gene to the log2 (in vivo/in vitro) to the reference gene pdxJ. Note: *represents a significant difference (P < 0.05), ** and more represent an extremely significant difference (P < 0.01).

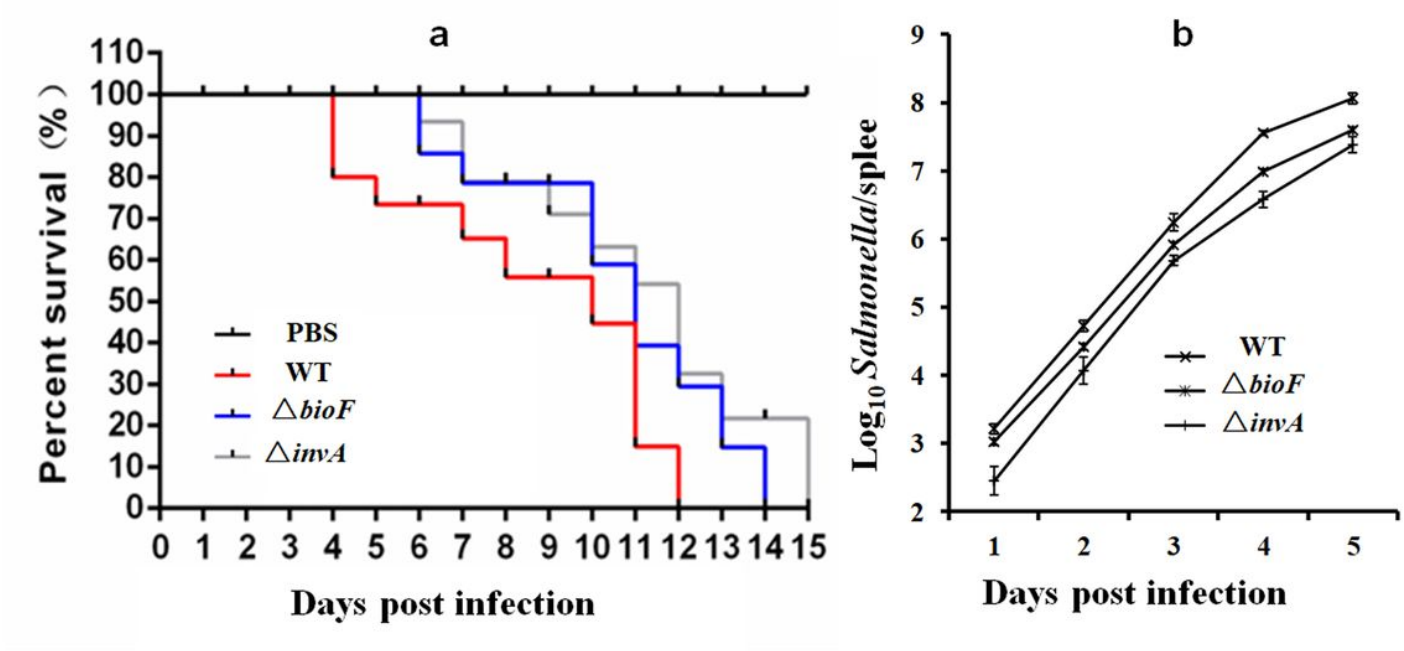

Figure 7

Virulence of bioF mutant strains in mice (a) The curve of percent survival. Regarding the survival rate of mice, the deletion strain was consistent with the positive control in which the virulence gene invA was knocked out. After knocking out the candidate gene bioF, the mortality of the bacteria was significantly reduced (P < 0.05). (b) The S. typhimurium load in the spleen. The bacterial load in the ΔinvA group was significantly lower than in the bioF group and the wild-type group (P < 0.05), and the bacterial load in the bioF group was also lower than in the reference strain group, but this difference was not significant.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS4DetailedDEGsdatafortop20GOenriched.xlsx
- TableS5KEGGpathwaysclassificationforDEGs.xlsx
- TableS1TheprimersusedforrealtimeqPCRanalysis.docx
- TableS6Top20KEGGpathwaysenrichedintheDEGs.xlsx

- TableS2Theprimersusedformutantstrainconstruction.docx
- TableS3DetailedexpressiondataforDEGs.xlsx