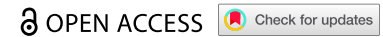



RESEARCH ARTICLE



# Genome-wide identification of genes critical for *in vivo* fitness of multi-drug resistant porcine extraintestinal pathogenic *Escherichia coli* by transposon-directed insertion site sequencing using a mouse infection model

Fan Yin<sup>a</sup>, Yan Hu<sup>b</sup>, Zixuan Bu<sup>a</sup>, Yuying Liu<sup>a</sup>, Hui Zhang<sup>a</sup>, Yawen Hu<sup>a</sup>, Ying Xue<sup>a</sup>, Shaowen Li<sup>a</sup>, Chen Tan<sup>a,c,d</sup>, Xiabing Chen<sup>e</sup>, Lu Li<sup>a,c,d</sup>, Rui Zhou<sup>a,c,d,f</sup>, and Qi Huang <sup>a,c,d</sup>

<sup>a</sup>State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan, China; <sup>b</sup>College of Animal Sciences & Technology, Huazhong Agricultural University, Wuhan, China; <sup>c</sup>Cooperative Innovation Center for Sustainable Pig Production, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan, China; <sup>d</sup>Ministry of Science and Technology, International Research Center for Animal Disease, Wuhan, China; <sup>e</sup>Institute of Animal Husbandry and Veterinary Science, Wuhan Academy of Agricultural Science and Technology, Wuhan, China; <sup>f</sup>The HZAU-HVSEN Institute, Wuhan, China

## ABSTRACT

Extraintestinal pathogenic *Escherichia coli* (ExPEC) is an important zoonotic pathogen. Recently, ExPEC has been reported to be an emerging problem in pig farming. However, the mechanism of pathogenicity of porcine ExPEC remains to be revealed. In this study, we constructed a transposon (Tn) mutagenesis library covering Tn insertion in over 72% of the chromosome-encoded genes of a virulent and multi-drug resistant porcine ExPEC strain PCN033. By using a mouse infection model, a transposon-directed insertion site sequencing (TraDIS) assay was performed to identify *in vivo* fitness factors. By comparing the Tn insertion frequencies between the input Tn library and the recovered library from different organs, 64 genes were identified to be involved in fitness during systemic infection. 15 genes were selected and individual gene deletion mutants were constructed. The *in vivo* fitness was evaluated by using a competitive infection assay. Among them,  $\Delta fimG$  was significantly outcompeted by the WT strain *in vivo* and showed defective adhesion to host cells. *rfa* which was involved in lipopolysaccharide biosynthesis was shown to be critical for *in vivo* fitness which may have resulted from its role in the resistance to serum killing. In addition, several metabolic genes including *lepB*, *sdhC*, *lepG*, *gltS*, *dcuA*, *ccmH*, *ddpD*, *narU*, *glpD*, *malM*, and *yabL* and two regulatory genes *metJ* and *baeS* were shown as important determinants of *in vivo* fitness of porcine ExPEC. Collectively, this study performed a genome-wide screening for *in vivo* fitness factors which will be important for understanding the pathogenicity of porcine ExPEC.

## ARTICLE HISTORY

Received 16 March 2022  
Revised 3 December 2022  
Accepted 11 December 2022

## KEYWORDS

Transposon mutagenesis; transposon-directed insertion site sequencing (TraDIS); genome-wide screening; extraintestinal pathogenic *Escherichia coli*; fitness factor


## Introduction

Extraintestinal pathogenic *Escherichia coli* (ExPEC) is posing an important threat to public health, which is involved in infections or even death of both humans and farm animals [1]. Differing from commensal or intestinal pathogenic *E. coli*, ExPEC causes a wide range of extraintestinal infections which can colonize the intestine asymptotically as the reservoir of infection [2,3]. In human, ExPEC cause diseases including urinary tract infections (UTIs), neonatal meningitis, and septicemia [4]. It is estimated that ExPEC accounts for more than 60% of human UTIs and 27% of documented bacteremia which causes significant public health problems and economic burdens [5–7]. ExPEC is also a common bacterial pathogen in farm animals. Apart from the well-characterized avian pathogenic *E. coli* (APEC) that can cause multiple local and systemic infections in poultry

which is one of the leading causes of morbidity and mortality [8], ExPEC in pigs has recently emerged as a fatal pathogen causing porcine meningitis, pneumonia, and septicemia, resulting in huge economic losses [9–11]. The incidence of ExPEC in pigs is reported to be increasing in recent years [11–13]. In addition, ExPEC of animal origins has been recognized as an important source of foodborne illness [14]. The high level of antimicrobial resistance and high flexibility in genome diversity brings enormous difficulties in controlling and preventing porcine ExPEC infections.

Understanding the mechanism of bacterial pathogenesis is critical for the development of drugs and vaccines. *E. coli* populations have high genetic and phenotypic diversity [15]. Different types of ExPEC carry specific fitness factors and virulence factors (VFs) to adapt to various niches [16]. Uropathogenic

**CONTACT** Rui Zhou  [rzhou@mail.hzau.edu.cn](mailto:rzhou@mail.hzau.edu.cn); Qi Huang  [qhuang@mail.hzau.edu.cn](mailto:qhuang@mail.hzau.edu.cn)

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/21505594.2022.2158708>

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*E. coli* (UPEC) produce P fimbriae to facilitate adherence to renal cells [17,18] and are equipped with a higher amino acid biosynthetic capacity which is important for growth in urine [19]. Neonatal meningitis *E. coli* (NMEC) and sepsis-associated *E. coli* (SEPEC) strains have the ability to survive in the bloodstream in which their important VFs include capsular antigens to protect against phagocytosis, Iss protein for survival in serum, and IbeABC to promote host cells invasion [4,20,21]. APEC carry avian-associated ColV plasmids essential for poultry adaptation that encompass virulence genes such as *iucC*, *iucD*, *iutA*, *cvaA*, *etsA*, *hlyF*, *ompT*, *cvaB*, *cvaC*, and *cvi* [22,23]. Nevertheless, some VFs such as the iron acquisition factors and some adhesins are shared among different pathotypes of ExPEC [4]. Moreover, efforts have been made to analyze the population structure of ExPEC by using comparative genomics approaches. However, the results suggest that it is still difficult to reveal the underlying mechanism of pathogenesis and identify disease-specific functional or genetic coordinates of ExPEC [24,25].

Porcine ExPEC has recently become an emerging problem in pig farming in China [10,26,27]. Increasing incidence and high levels of antimicrobial resistance make porcine ExPEC a major bacterial pathogen in the pig industry [11,28]. However, the mechanisms of pathogenicity are less understood. So far, the type VI secretion system, Rhs proteins, the twin-arginine translocation system, the biosynthesis of nucleotide and extracellular polysaccharides, and the global regulators Fur and FNR have been reported critical for the virulence of porcine ExPEC [9,29–31]. However, the underlying mechanism of pathogenesis is far from comprehensively elaborated. Therefore, further identification of factors involved in pathogenicity is necessary to provide a better understanding of this pathogen.

Transposon-directed insertion site sequencing (TraDIS) has become a powerful approach that combines transposon mutagenesis with high-throughput sequencing to do genome-wide screenings [32–34]. By comparing the transposon insertion frequency between the treated and untreated transposon libraries, TraDIS allows the identification of a large number of functional genes in a single experiment. It has been used to identify essential genes, drug-resistant genes, and *in vivo* fitness genes in several important bacterial pathogens [35–38].

Our previous studies isolated and characterized a number of clinical ExPEC strains from pigs and identified a strain PCN033 which is multi-drug resistant and virulent, belonging to sequence type 597,

serogroup O11, and phylogenetic group D [26,39,40]. Using a pig infection model, it was demonstrated that the PCN033 strain possessed high virulence which can cause systemic infection including neurological symptoms, bacteremia, and even death [27]. The high virulence of the PCN033 strain was also observed in the mouse infection model where the mice infected with this strain also exhibited similar symptoms to those in the pig infection model, suggesting that the mouse infection model can be a surrogate model of pig infection [31,41]. In this study, we constructed a transposon mutagenesis library of ExPEC PCN033 strain covering Tn insertion in over 72% of the chromosome-encoded genes and performed a TraDIS assay using a mouse infection model which identified a variety of genes potentially involved in *in vivo* fitness during infection. By constructing individual gene deletion mutants, we further confirmed that genes involved in fimbriae and LPS biosynthesis, metabolism, and transcription regulation are critical for porcine ExPEC to establish systemic infection. Our genome-wide screening for *in vivo* fitness factors will be important for understanding the pathogenesis of porcine ExPEC.

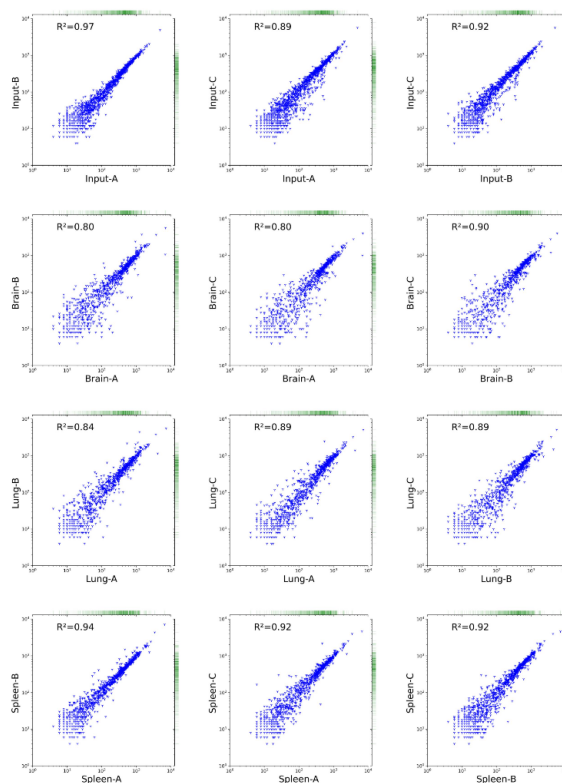
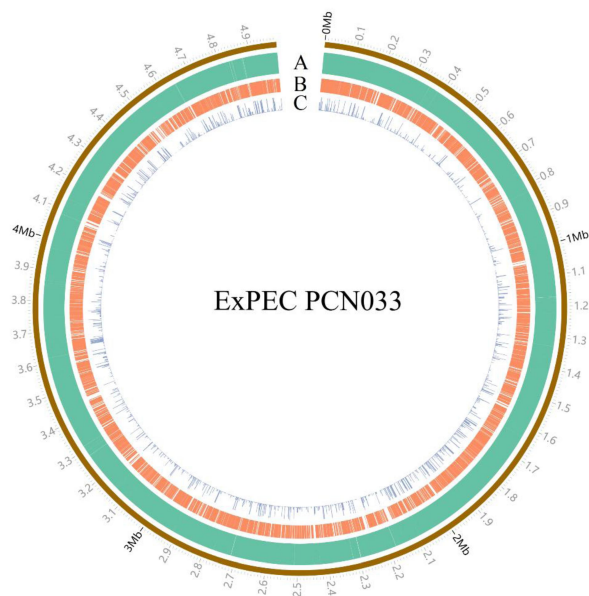
## Results

### Transposon mutagenesis library construction and evaluation

The ExPEC PCN033 strain is a virulent clinical isolate and a multi-drug resistant strain [39]. So, we designed and constructed a plasmid pSAM-Tat-AprA for transposon mutagenesis for this strain (Fig. S1A). Transposon mutagenesis was successfully achieved as verified by PCR amplification and insertion site identification. Randomly selected transposon mutants showed transposition at different sites on the chromosome (Fig. S1B). A transposon library containing approximately 68,000 individual transposon mutants was constructed. By sequencing the library and mapping the transposon insertion sites to the genome, it was shown that a total of 3541 genes contained insertions, comprising about 72% of the total chromosome-encoded genes (Figure 1(a)).

### Identification of genes critical for *in vivo* fitness by TraDIS

To screen for genes critical for *in vivo* fitness, the transposon library was used to infect mice intraperitoneally. The input library and the bacteria recovered from the organs including the brain, spleen, and lung (output libraries) were isolated and sequenced. By



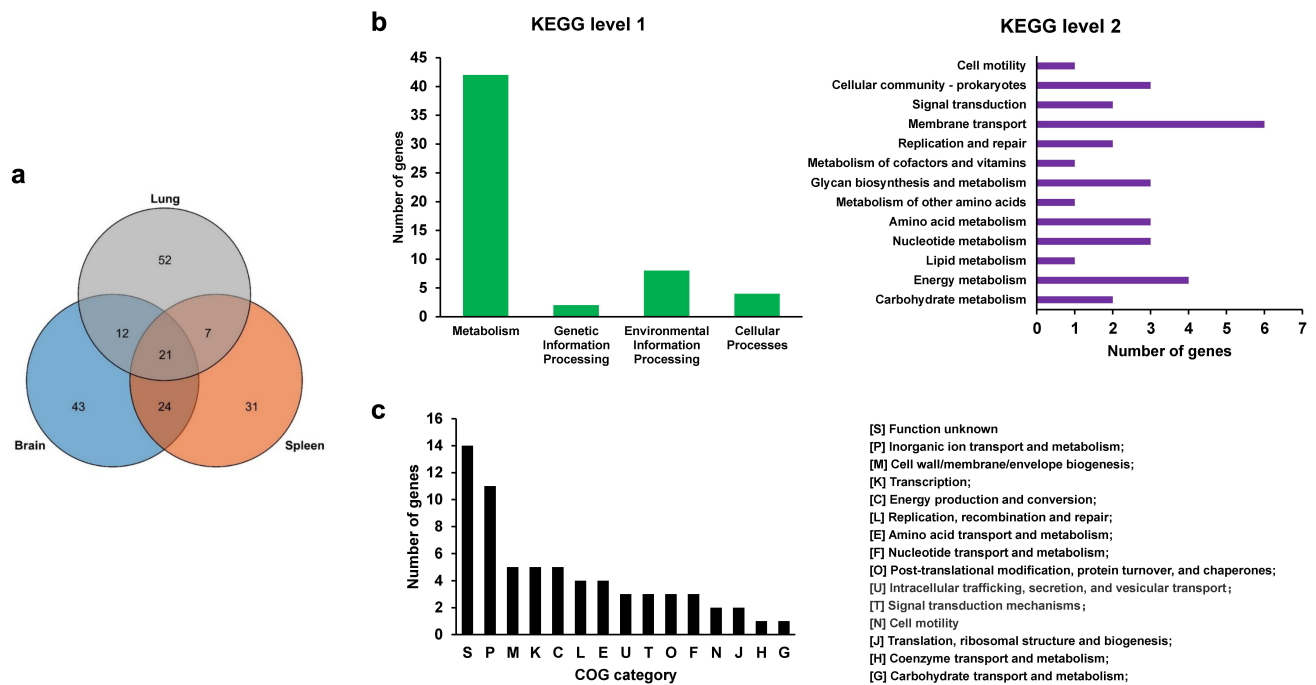
**Figure 1.** Distribution of transposon insertion sites and correlation coefficients. (a) distribution of transposon insertion sites of the input library. the transposon insertion sites were mapped to the genome of ExPEC strain PCN033 and the data was visualized by using circos. Circle a represents the genome of ExPEC PCN033 strain. Circle B shows the transposon insertion sites. Circle C shows the mapped reads for each insertion site. The figure showed the data of one of the input libraries. (b) correlation coefficients between duplicates of input and output libraries. the correlation coefficients were calculated using the StandardScaler submodule of the sklearn module of the python programming language. DESeq2 was used to compare the change in the total number of insertion reads for each gene between the input and output libraries. Fold-change (FC) > 2 and  $p$  value < 0.05 was used as the threshold to identify enriched genes.

calculating the insertion frequency between the input and output libraries, genes with differential transposon insertion were identified. As shown in Figure 1(b), the correlation coefficients of the genes with differential insertion frequency were between 0.8–0.97, indicating consistency between biological replicates. A total of 190 genes were identified with differential insertions (FC > 2,  $p$  value < 0.05) in the output libraries recovered from brain, spleen, and lung tissues compared with the input library, which may be potentially involved in *in vivo* fitness of ExPEC (Figure 2 and Table S3). Among them, 64 genes were identified that were involved in the fitness in at least two organs (Figure 2 and Table S3). KEGG pathway enrichment analysis with these 64 genes revealed that most of the genes are involved in metabolism, among which membrane transport and energy metabolism were the most enriched level 2 KEGG pathways. By COG enrichment, most of the TraDIS-identified genes belonged to the categories including functional unknown [S], inorganic ion transport and metabolism [P], Cell wall/membrane/envelope

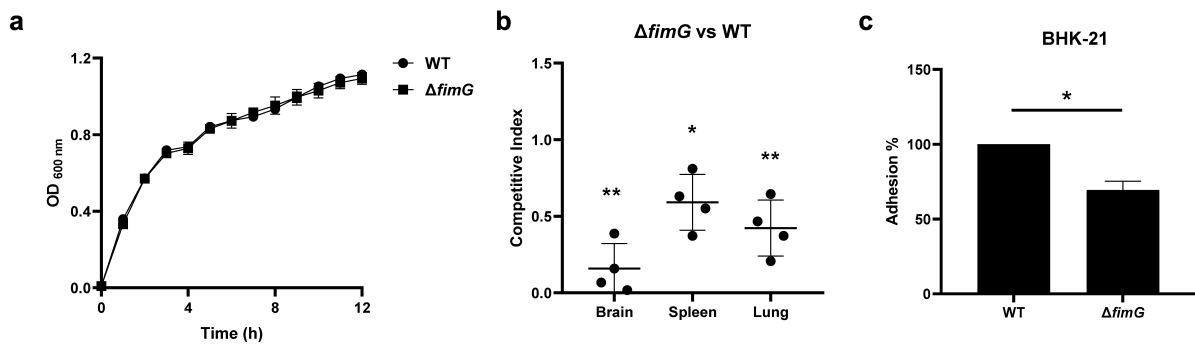
biogenesis [M], Transcription [K], and Energy production and conversion [C].

### ***fimG* is required for adhesion of ExPEC**

*fimG* encodes the minor component of type 1 fimbriae, which has been identified with a decreased transposon insertion in the output libraries recovered from all three organs (Table S3). To further confirm its role in pathogenesis, we constructed the deletion mutant of *fimG*. The growth curve data showed that the mutant did not show growth defects compared with WT strain (Figure 3(a)). Competitive animal infection assay, a sensitive and accurate approach to determine the degree of virulence attenuation [42], was used to evaluate the *in vivo* colonization of the mutant. A control assay using ExPEC PCN033 strain and PCN033-Chl strain which was the PCN033 strain inserted with a chloramphenicol resistance cassette in the intergenic region of the chromosome, to inject



**Figure 2.** Genes essential for *in vivo* fitness. (a) venn diagram of the fitness genes identified in different organs. the diagram was generated by using Tbttools according to Table S3. (b) KEGG pathway enrichment of the fitness genes. the fitness genes were analyzed by using BlastKOALA tool (<https://www.Kegg.jp/blastkoala/>) to assign K numbers which were then used for pathway enrichment by using KEGG mapper (<https://www.Genome.jp/kegg/mapper/search.Html>). The numbers inside the vertical bars are the total number of genes belonging to each level 2 KEGG pathway. (c) COG enrichment of the fitness genes. the fitness genes were analyzed using the EggNOG v5.0 database (<http://eggno5.Embl.de/#/app/home>).

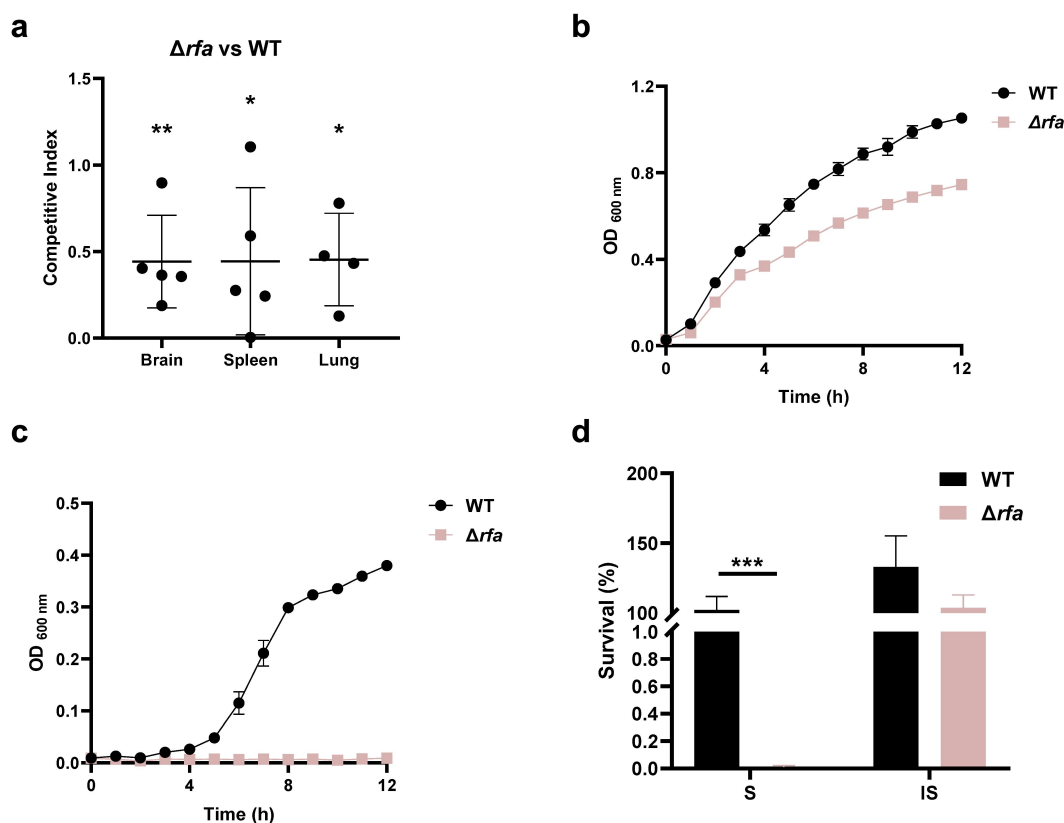


**Figure 3.** *fimG* is required for adhesion. (a) growth assay. Cells of WT and  $\Delta fimG$  strains were inoculated 1:100 from overnight-grown culture to LB. The growth was monitored using an automatic plate reader (bioscreen C, FP-1100-C, oy growth curves AB, USA) at 37°C with shaking. Five replicates were tested for each strain. The data shown are the mean  $\pm$  standard error of the mean (SEM). (b) competitive infection assay. 200  $\mu$ L of cell suspension containing a total of  $6 \times 10^5$  CFU of WT and  $\Delta fimG$  cells with a ratio of 1:1 was used to intraperitoneally inject mice with five mice in each group. At 12 hpi, mice were euthanized, and brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on LB agar plates with and without appropriate antibiotics, respectively, to distinguish the mutant and the wild-type cells. The bacteria were enumerated and the competition index (CI) was calculated as defined as the mutant-to-WT ratio within the recovered sample, divided by the corresponding ratio in the inoculum. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the mean CI value and 1. \* indicates  $p < 0.05$  and \*\* indicates  $p < 0.01$ . (c) cell adhesion assay. the bacterial cells at the mid-log phase were used to infect BHK-21 cells with a multiplicity of infection (MOI) of 10:1 in a six-well plate followed by incubation at 37°C for 1 h. The mixture was washed with sterile PBS and then sterile water was added to the mixture which was incubated at 4°C for 1 h for cell lysis. The lysate was serially diluted and applied to LB agar plates for bacterial counting. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the two groups. \* indicates  $p < 0.05$ .

mice simultaneously. The result showed that the insertion of the resistance cassette did not interfere with the competitive index (Figure S2). By performing a competitive animal infection assay using  $\Delta fimG$  and WT strain, it was shown that the mutant strain displayed significantly reduced colonization in the brain (CI = 0.158), spleen (CI = 0.591), and lung (CI = 0.338) compared with the WT strain (Figure 3(b)). Then, *in vitro* cell adhesion assay was carried out which showed that the  $\Delta fimG$  strain displayed a significantly lower ability to adhere to and be internalized by BHK-21 cells (Figure 3(c)). Therefore, our results showed that *fimG* is critical for host adhesion of ExPEC.

### Lipopolysaccharide biosynthesis is critical for serum survival

*rfaI*, *rfaJ*, and *rfaL* (also known as *waaL*) are within the same operon involved in the biosynthesis of lipopolysaccharide. These three genes have been identified in the TraDIS assay (Table S3). A mutant strain lacking the *rfa* operon was constructed ( $\Delta rfa$ ). The competitive infection assay confirmed that  $\Delta rfa$  strain showed a reduced ability to colonize in mice (Figure 4(a)). The growth assay revealed that  $\Delta rfa$  showed slightly reduced growth in LB medium compared with the WT strain in which the maximum growth rate and



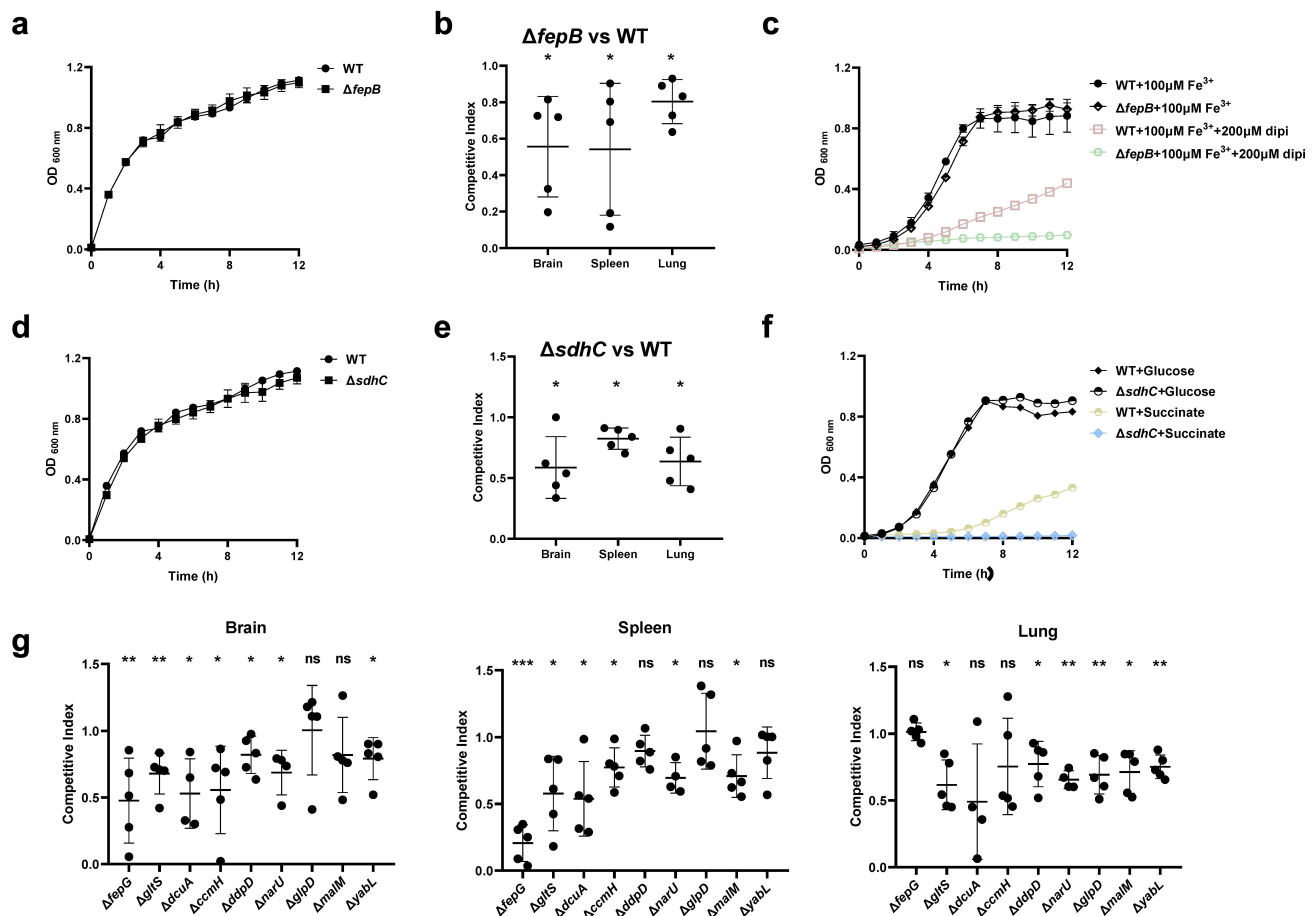
**Figure 4.** *Rfa* is critical for serum survival. (a) competitive infection assay. 200  $\mu$ L of cell suspension containing a total of  $6 \times 10^5$  CFU of WT and  $\Delta rfa$  cells with a ratio of 1:1 was used to intraperitoneally inject mice with five mice in each group. At 12 hpi, mice were euthanized, and brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on LB agar plates with and without appropriate antibiotics, respectively, to distinguish the mutant and the wild-type cells. The bacteria were enumerated and the competition was calculated. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the mean CI value and 1. \* indicates  $p < 0.05$ , and \*\* indicates  $p < 0.01$ . (b and c) growth assay. Cells of WT and  $\Delta rfa$  strains were inoculated 1:100 from overnight-grown culture to LB (A) or supplemented with 75% fresh mouse serum (B). The growth was monitored using an automatic plate reader (bioscreen C, FP-1100-C, oy growth curves AB, USA) at 37°C with shaking. Five replicates were tested for each strain. The data shown are the mean  $\pm$  standard error of the mean (SEM). (d) serum killing assay. Cells of WT and  $\Delta rfa$  strains were subcultured in LB from overnight-grown culture and grown at 37°C to the mid-log phase. The cells were washed with sterile saline and  $5 \times 10^5$  CFU of the cells was incubated with 75% fresh (S) or heat-inactivated serum (IS) for 20 min at 37°C. Afterward, the mixture was serially diluted and plated on LB agar plates for viable cell enumeration. The assay was performed in triplicates and the data shown are the mean  $\pm$  standard error of the mean (SEM). The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the two groups. \*\*\* indicates  $p < 0.001$ .



the total growth yield of the mutant were 68.1% and 70.4%, respectively, compared to that of the WT strain (Figure 4(b)). In contrast, the growth of  $\Delta rfa$  was almost abolished in the presence of 75% fresh mice serum (Figure 4(c)). This was consistent with the serum killing assay results demonstrating that the  $\Delta rfa$  strain was markedly more susceptible to fresh serum (S) but not the heat-inactivated serum (IS) than the WT strain (Figure 4(d)). Therefore, the data suggest that *rfa* operon plays an important role in fitness *in vivo* which could be mediated by its effects on resistance to serum killing.

### Several metabolic genes are involved in colonization in vivo

In the TraDIS results, a large proportion of genes related to bacterial metabolism were identified to be required for fitness *in vivo* (Figure 2(b)). *fepB* is part of the ABC transporter complex FepBDGC involved in ferric enterobactin uptake [43]. To investigate the role of *fepB* in the pathogenesis of ExPEC, we constructed a *fepB* mutant strain. Animal infection assay revealed that the ability of the mutant to colonize in the brain, spleen, and lung tissues of mice was significantly lower than that of the WT strain, indicating that *fepB*



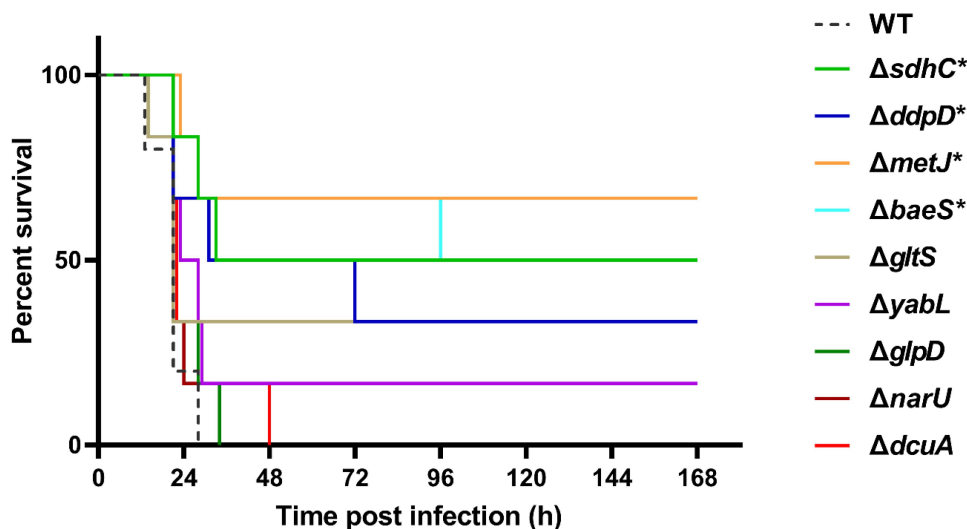
**Figure 5.** Fitness of metabolism-related mutants. (a, c, d, and f) growth assay. Cells of each indicated strain were inoculated 1:100 from overnight-grown culture to LB (a and e), M9 medium or M9 medium supplemented with the additives (D and G). The growth was monitored using an automatic plate reader (bioscreen C, FP-1100-C, oy growth curves AB, USA) at 37°C with shaking. Five replicates were tested for each strain. The data shown are the mean  $\pm$  standard error of the mean (SEM). (b, e, and g) competitive infection assay. 200  $\mu$ L of cell suspension containing a total of  $6 \times 10^5$  CFU of cells of WT and each indicated strain with a ratio of 1:1 was used to intraperitoneally inject mice with five mice in each group. At 12 hpi, mice were euthanized, and brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on LB agar plates with and without appropriate antibiotics, respectively, to distinguish the mutant and the wild-type cells. The bacteria were enumerated and the competition index (CI) was calculated. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the mean CI value and 1. ns indicates no significant difference, \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , \*\*\* indicates  $p < 0.001$ .

significantly improved ExPEC colonization *in vivo* (Figure 5(b)). Growth assay revealed that  $\Delta fepB$  showed a similar growth with WT strain in LB (Figure 5(a)) or in M9 medium supplemented with sufficient iron (Figure 5(c)). However, when iron was depleted by the addition of 2, 2'-bipyridine (dipi), compared with the WT strain that the growth was partially affected, the growth of  $\Delta fepB$  was almost abolished (Figure 5(c)). This suggests that  $\Delta fepB$  is involved in the growth under conditions with limited iron. *sdhC* encodes the membrane-anchored subunit of succinate dehydrogenase, which is involved in the tricarboxylic acid cycle pathway of carbohydrate metabolism, catalyzing succinate to fumarate [44]. It was shown that the growth of the  $\Delta sdhC$  was comparable to that of the WT strain in LB medium (Figure 5(d)). However, the competitive infection assay showed that the deletion of *sdhC* significantly compromised the ability of ExPEC to colonize the brain, spleen, and lung (Figure 5(e)). We further tested the growth of  $\Delta sdhC$  in the M9 medium containing glucose or succinate as the only carbon source. The  $\Delta sdhC$  strain grew at a similar rate to the WT strain in the M9-glucose medium (Figure 5(f)). It was also seen that the WT strain grow slower in the M9-succinate medium than in the M9-glucose medium (Figure 5(f)), which was consistent with the fact that glucose is a preferred carbon source for *E. coli*. However, the growth of  $\Delta sdhC$  strain was almost abolished in the M9-succinate medium while the growth of the WT strain was only partially influenced

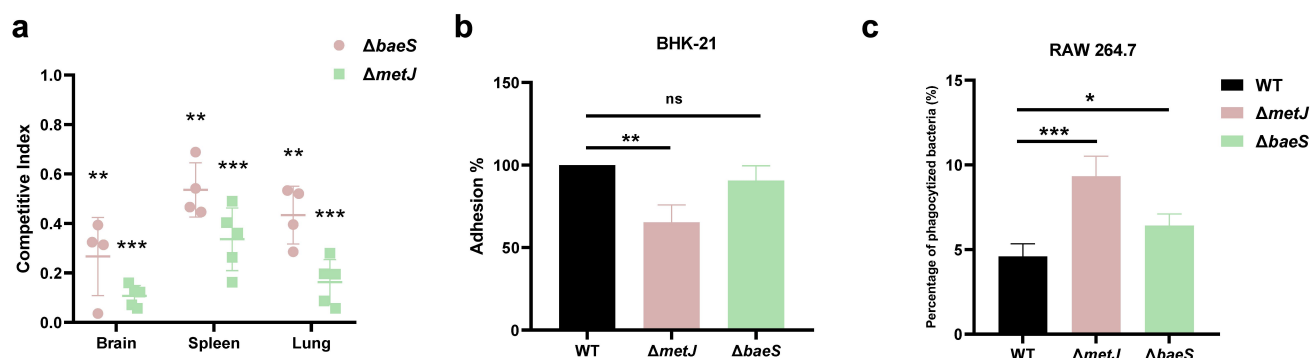
(Figure 5(f)). In addition, we constructed deletion mutants of *fepG*, *gltS*, *dcuA*, *ccmH*, *ddpD*, *narU*, *glpD*, *malM*, and *yabL*, which are involved in ferric enterobactin transport, glutamate transport, C4-dicarboxylate transport, c-type cytochrome biosynthesis, D, D-dipeptide transport, nitrate/nitrite transport, glycerol metabolism, maltose uptake, and thiamine import, respectively. There was no significant difference in growth in LB medium between the mutants and the WT strain (data not shown). However, the competitive infection assay revealed that their *in vivo* fitness was compromised in at least one organ tested (Figure 5(g)). To further determine the contribution of each gene to the pathogenicity of ExPEC, we tested the survival of mice infected with WT,  $\Delta sdhC$ ,  $\Delta gltS$ ,  $\Delta yabL$ ,  $\Delta dcuA$ ,  $\Delta ddpD$ ,  $\Delta narU$ , and  $\Delta glpD$  strain, respectively. As shown in Figure 6, the mice infected with  $\Delta sdhC$  or  $\Delta ddpD$  showed significantly higher survival, while the other strains showed comparable virulence to the WT strain.

### Two regulators are identified critical for *in vivo* fitness and virulence

*metJ* which encodes a transcription regulator [45] and *baeS* which encodes the sensor kinase of the two-component system BaeSR [46] were identified in the TraDIS assay. The deletion mutants of *metJ* and *baeS* were constructed. It was shown that both of the deletion mutants were significantly outcompeted by the



**Figure 6.** Mice survival assay. the indicated ExPEC strains were subcultured from overnight-grown cultures into LB and grown to the mid-log phase at 37°C with shaking. The cells were harvested by centrifugation, washed with sterile saline, and diluted to appropriate concentrations with saline. 0.2 ml of the bacteria suspension containing approximately  $6.9 \times 10^5$  CFU was used to infect 4-week-old Kunming mice intraperitoneally with 6 mice in each group (except the WT group which harbored 5 mice). The survival of mice was recorded. Statistical differences in survival between the mutant group and the WT group were determined using the log-rank test. \* indicates  $p < 0.05$ .



**Figure 7.** Two regulatory genes are critical fitness factors. (a) competitive infection assay. 200  $\mu$ L of cell suspension containing a total of  $6 \times 10^5$  CFU of cells of WT and each indicated strain with a ratio of 1:1 was used to intraperitoneally inject mice with five mice in each group. At 12 hpi, mice were euthanized, and brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on LB agar plates with and without appropriate antibiotics, respectively, to distinguish the mutant and the wild-type cells. The bacteria were enumerated and the competition index (CI) was calculated. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the mean CI value and 1. \*\* indicates  $p < 0.01$ , and \*\*\* indicates  $p < 0.001$ . (b) cell adhesion assay. the bacterial cells at the mid-log phase were used to infect BHK-21 cells with a multiplicity of infection (MOI) of 10:1 in a six-well plate followed by incubation at 37°C for 1 h. The mixture was washed with sterile PBS and then sterile water was added to the mixture which was incubated at 4°C for 1 h for cell lysis. The lysate was serially diluted and applied to LB agar plates for bacterial counting. (c) macrophage phagocytosis assay. the bacterial cells were mixed with RAW 264.7 cells with a multiplicity of infection (MOI) of 10:1 followed by incubation at 37°C for 1 h. The cells were then washed twice with sterile PBS, and incubated with PBS containing chloramphenicol (final concentration 50  $\mu$ g/ml). The macrophage cells were then lysed with sterile water and the lysate was serially diluted and applied to LB agar plates for bacterial counting. The student's t test (two-tail, unpaired) was used to calculate the statistical difference between the two groups. ns indicates no significant difference, \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , \*\*\* indicates  $p < 0.001$ .

WT strain in the brain, spleen, and lung, in which  $\Delta metJ$  strain had a very low competitive index indicating a significant colonization defect compared with the WT strain (Figure 7(a)). To further confirm the roles of *metJ* and *baeS* in the pathogenesis of ExPEC, adhesion assay and phagocytosis by macrophages were carried out. It was shown that the deletion of *metJ* rather than *baeS* significantly compromised the adhesion of ExPEC to BHK-21 cells (Figure 7(b)). Also, both of  $\Delta metJ$  and  $\Delta baeS$  were easier to be phagocytized by RAW 264.7 macrophages than WT cells (Figure 7(c)). The survival assay further showed that the virulence of  $\Delta metJ$  and  $\Delta baeS$  was significantly attenuated compared with the WT strain (Figure 6). These results suggest that *MetJ* and *BaeSR* two-component system are virulence regulators of ExPEC.

## Discussion

Compared to intestinal pathogenic *E. coli*, ExPEC causes extraintestinal infections including meningitis, pneumonia, or septicemia, leading to more severe consequences. ExPEC not only threatens human health but also causes serious problems in animal husbandry. Recently, ExPEC has been reported to be an emerging problem of pig farming in which its incidence in pigs is increasing and highly virulent strains have been

reported [11,27,47]. So far, there are no effective vaccines developed for preventing ExPEC infection in pigs. Moreover, porcine ExPEC isolates have been revealed to possess broad-spectrum antimicrobial resistance, significantly compromising the efficacy of antibiotic treatment [28]. Therefore, there is an urgent need for research to understand the mechanism of the pathogenesis of porcine ExPEC, which is important for the development of novel vaccines and antimicrobials.

The ability to adapt to and survive in adverse environments is critical for bacterial pathogens to establish infection. *In vitro* conditions that mimic the *in vivo* environment have been widely used to study the mechanism of the pathogenesis of bacterial pathogens [48–50]. However, bacteria encounter complicated challenges during infection *in vivo*, especially for ExPEC which causes systemic infection. Therefore, elucidating the mechanism of *in vivo* fitness is important for understanding the pathogenesis of ExPEC. TraDIS has recently become a popular approach that offers the chance to do genome-wide identification of *in vivo* fitness factors in one shot [35,36,51,52]. One of the key steps of performing TraDIS is the construction of the transposon mutagenesis library. Mariner and Tn5 transposons are the most popular ones used for mutagenesis due to convenient manipulations, broad-host range, and more importantly the nature of near-



random insertion [35,53]. Various approaches have been utilized to deliver the transposon into the recipient cells, including the phage delivery systems [51], plasmid delivery systems [54,55], and using *in vitro* assembled transposase [37]. In this study, a clinically isolated and multi-drug resistant strain ExPEC PCN033 was used as the recipient strain to generate the transposon mutagenesis library in which mariner transposon encoded by a suicide plasmid was delivered by transconjugation. A Tn library consisting of over 68,000 mutants was generated, which covered 72% of the genes encoded by the genome as revealed by library sequencing and genome mapping (Figure 1(a)), indicating a near-saturating library.

Mouse with intraperitoneal injection was then used to model the infection of ExPEC to identify the *in vivo* fitness genes of the ExPEC PCN033 strain. Several other routes of infection have also been reported for establishing the mouse infection model, including the intravenous, intranasal routes, and oral administration, depending on the type of pathogen [56]. For pathogens that cause systemic infection, such as ExPEC, intraperitoneal injection is a commonly used route of infection when using the mouse model of infection [57,58]. However, it should be noted that ExPEC enters extraintestinal sites and cause systemic infection from the intestine or through the urinary tract system [59]. Therefore, using the natural host pigs as the model of infection will be of value for future pathogenesis studies of porcine ExPEC despite the limitations of pig models including variations in genetic background, health conditions, and high expenses.

In this study, we constructed a Tn library and for the first time performed a TraDIS assay to do genome-wide identification of *in vivo* fitness factors with porcine ExPEC using a mouse infection model. A total of 64 genes were identified as essential fitness factors during systemic infection. Among these genes, *fimG*, encoding the minor component of type 1 fimbriae, was shown to be involved in fitness [60]. By performing cell adhesion assay, we confirmed that *fimG* plays an important role in adhesion to host cells. Type 1 fimbriae have also been reported as an important virulence factor in UPEC [60–62]. *fepG* and *fepB* encode the components of a transporter involved in the uptake of ferric enterobactin, a siderophore important for the virulence of ExPEC [63,64]. They were also identified in our TraDIS assay to be required for *in vivo* fitness of porcine ExPEC. Our *in vitro* assay further demonstrated that *fepB* is required for ferric iron uptake. The Fep system serves as a siderophore import system and allows the uptake of iron through both the enterobactin and salmochelin siderophores [65,66]. Some bacteria employ

other siderophores to scavenge iron that do not depend on the Fep system, such as aerobactin [67] and yersiniabactin [68]. It should be noted that ExPEC PCN033 strain does not produce yersiniabactin but contains an aerobactin-encoding gene on the plasmid. Therefore, it remains to be investigated the role of the plasmid-encoded aerobactin synthesis and import system. In addition, we identified several genes involved in the biosynthesis of lipopolysaccharide core (*rfa*) in our TraDIS assay. It was further shown that  $\Delta rfa$  displayed remarkable growth defects in fresh serum but not in heat-inactivated serum, indicating that lipopolysaccharides play important role in protecting ExPEC from being killed by heat-sensitive antimicrobial substances in serum, such as complements. The critical role of LPS biosynthesis in the resistance of serum killing has also been reported by other groups [9,69–71]. This could be attributed to the critical role of the LPS O-antigen in resistance to complement killing and hydrolysis of lysozyme [71,72].

Bacterial metabolism has been recently recognized to play an important role in pathogenesis. By KEGG pathway enrichment analysis, among the 64 genes identified by TraDIS, 42 genes are metabolism-related (Figure 2(b)). Mouse competitive infection assays further validated the involvement of *fepB*, *sdhC*, *fepG*, *gltS*, *dcuA*, *ccmH*, *ddpD*, *narU*, *glpD*, *malM*, and *yabL* in *in vivo* fitness. Consistently, a previous TraDIS study with an ExPEC strain of avian origin also shows that metabolism genes comprise the most enriched group that is essential for *in vivo* fitness [58]. In UPEC, a significant number of proteins exhibited differential expression in response to urine treatment. Among them, metabolic genes including those involved in peptide uptake, gluconeogenesis, and the tricarboxylic acid cycle were demonstrated to be required for fitness during urinary tract infection [73]. Among these genes, *sdhC* which is a subunit of the succinate dehydrogenase catalyzing the production of fumarate from succinate has also been reported to be important for the virulence of other bacterial pathogens, such as enterohemorrhagic *E. coli* [74] and *Salmonella* [75]. Recently, metabolic crosstalk between bacteria and host cells is becoming an emerging research hotspot [76]. Among the critical metabolites, succinate has been shown to modulate host inflammatory response [77]. Whether the increased succinate level in the  $\Delta sdhC$  mutant has an impact on the host innate immune response to facilitate bacteria clearance remains to be revealed. *ddpD* is an uncharacterized putative D, D-dipeptide transporter in *E. coli* in addition to the other reported transporters such as DppA [78], DtpA [79], DtpB [79], DtpC [80], and DtpD [81]. The mouse survival assay further

confirmed the critical role of *ddpD* for the virulence of ExPEC (Figure 6). Therefore, its role in the pathogenesis of ExPEC is worth further investigation. It should be noted that although *dcuA*, *gltS*, *narU*, *glpD*, and *yabL* were outcompeted by the WT strain *in vivo*, mouse assay with each individual mutant showed that they had similar overall virulence in terms of mice survival (Figure 6). This could be attributed to the existence of other functional redundant genes. For example, *dcuA* encodes a transporter for C4-dicarboxylate, such as succinate, L-malate, L-tartrate, and L-aspartate. Apart from DcuA, another C4-dicarboxylate transporter DcuB exists in *E. coli*. These two transporters have preferred substrates, and DcuA is constitutively expressed while DcuB is expressed only anaerobically [82,83]. *gltS* encodes a glutamate transporter in *E. coli*. Moreover, there are three other glutamate transporters in *E. coli*, among which Glts is responsible for around 25% of the total glutamate transport [84]. In addition to NarU, *E. coli* produce another transporter NarK for nitrate uptake [85]. But, it has been shown that NarU is more important during nutrient starvation or very slow growth [86]. Therefore, disrupting these genes may affect their capability of *in vivo* competition and fitness while may not necessarily decrease the overall virulence.

In addition, a regulatory gene *metJ*, encoding a transcriptional repressor involved in methionine biosynthesis, was identified as critical for *in vivo* fitness of ExPEC. Deleting *metJ* may result in the increase of methionine biosynthesis in ExPEC. It has been reported that increasing exogenous or endogenous methylthioadenosine (MTA), a methionine-derived metabolite, or deletion of *metJ* could suppress the virulence of *Salmonella* [87]. MetJ is also reported to be required for the virulence of other bacteria, including *Vibrio*

*cholera* [88] and *Pectobacterium atrosepticum* [89]. This indicates a critical role of bacterial methionine metabolism for virulence. As methionine metabolism is intimately related to methylation processes, Bourgeois *et al.* investigated the influence of *metJ* deletion on the DNA methylome of *Salmonella* [90]. However, the results show that the methylome is remarkably static even in the *metJ* deletion strain. Therefore, the exact role of *metJ* in the pathogenesis of ExPEC needs further investigation.

A total of 15 genes out of the 64 genes identified from the TraDIS were selected for validation of which the results of the competitive infection assay were summarized in Table 1. They showed different fitness abilities in different organs. Moreover, it was noted that in the TraDIS assay, some genes showed fitness defects in only one of the organs, in which 43 genes were identified only in the brain, 52 genes only in the lung, and 31 genes only in the spleen (Figure 2(a) and Table S3). This may be because different organs have their specific microenvironments as well as immune defense mechanisms [91]. For example,  $\Delta glpD$  only shows colonization defects in the lung but not in the brain and spleen (Table 1), this could be due to the aerobic condition in the lung, and *glpD* is required for glycerol metabolism in aerobic growth [92]. It will be interesting to identify genes critical for local infections in specific organs in the future.

It should be noted that there are several reported virulence-related factors of porcine ExPEC that have not been identified in our TraDIS assay, such as the twin-arginine translocation system [31] and the type VI secretion system (T6SS) [93]. One reason could be that our Tn library did not cover all the genes. For example, the genes encoding the T6SS were without any Tn insertions (Table S4). Also, it has been reported that

**Table 1.** Summary of competitive index.

Strain	Organ								
	Brain			Spleen			Lung		
	CI mean	<i>p</i> value	significance	CI mean	<i>p</i> value	significance	CI mean	<i>p</i> value	significance
$\Delta fimG$	0.158	0.002	***	0.591	0.021	*	0.338	0.008	**
$\Delta metJ$	0.163	0.000	***	0.499	0.000	***	0.249	0.000	***
$\Delta baeS$	0.267	0.003	***	0.536	0.003	***	0.434	0.002	***
$\Delta fepG$	0.477	0.021	*	0.206	0.000	***	1.014	0.359	ns
$\Delta dcuA$	0.530	0.037	*	0.538	0.021	*	0.773	0.100	ns
$\Delta fepB$	0.556	0.023	*	0.541	0.047	*	0.803	0.045	*
$\Delta ccmH$	0.556	0.039	*	0.773	0.026	*	0.755	0.203	ns
$\Delta sdhC$	0.587	0.022	*	0.824	0.011	*	0.637	0.015	*
$\Delta gltS$	0.679	0.010	**	0.577	0.028	*	0.615	0.020	*
$\Delta narU$	0.687	0.033	*	0.695	0.013	*	0.656	0.002	**
$\Delta yabL$	0.791	0.042	*	0.883	0.247	ns	0.752	0.003	**
$\Delta malM$	0.819	0.223	ns	0.709	0.015	*	0.713	0.016	*
$\Delta ddpD$	0.819	0.045	*	0.896	0.122	ns	0.773	0.040	*
$\Delta rfa$	0.846	0.009	**	0.719	0.043	*	0.454	0.026	*
$\Delta glpD$	1.004	0.980	ns	1.045	0.743	ns	0.693	0.009	**

ns indicates no significant difference, \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , and \*\*\* indicates  $p < 0.001$ .

TraDIS analysis using such a large mixed Tn library with an animal infection model may have experimental bottlenecks which results in inaccurate assignments of candidate fitness genes [94].

In conclusion, by TraDIS analysis with a Tn library covering insertions in 72% of the genome-encoded genes of a virulent and multi-drug resistant porcine ExPEC strain PCN033 using a mouse infection model, we identified 64 genes that may be involved in *in vivo* colonization. Genes involved in fimbriae and LPS biosynthesis, metabolism, and transcription regulation are further verified to be critical for porcine ExPEC to establish systemic infection. Our genome-wide screening for *in vivo* fitness factors will be important for understanding the pathogenesis of porcine ExPEC.

## Materials and methods

### Strains and plasmids constructions

The strains and plasmids used in this study were listed in Table S1. The primer sequences used in this study were listed in Table S2. ExPEC PCN033 strain is a multi-drug resistant clinical isolate that showed high virulence in pigs as well as in mice [27,39]. *E. coli* DH5 $\alpha$  was used as the host strain for regular cloning. *E. coli* DH5 $\alpha$  *pir* was used as the host strain for the propagation of pRE112 derived plasmids [95]. *E. coli*  $\chi$ 7213 is a diaminopimelic acid autotrophic *E. coli* strain used for transconjugation [96]. The recombinant plasmids were constructed by using a seamless cloning method with the ClonExpress MultiS One Step Cloning Kit (Vazyme, China) as previously described [31]. The gene deletion strains were constructed as previously described [31]. Briefly, the suicide plasmid containing the upstream and downstream flanking region of the target gene was delivered into ExPEC PCN033 strain via transconjugation using *E. coli*  $\chi$ 7213 strain as the donor strain. The single exchange strain was obtained and verified by PCR which was then screened on lysogeny broth (LB) agar plate containing 10% sucrose for the double exchanged mutant.

### Construction of transposon mutagenesis library

Transposon mutagenesis in ExPEC PCN033 strain was achieved using the transposon-containing suicide plasmid pSAM-Tat-AprA (Fig. S1A) delivered by transconjugation. Briefly, the auxotrophic *E. coli*  $\chi$ 7213 strain (nalidixic acid and chloramphenicol sensitive) carrying pSAM-Tat-AprA plasmid was used as the donor strain and ExPEC PCN033 as the recipient (nalidixic acid resistant and chloramphenicol sensitive). The donor

cells and the recipient cells were grown to the mid-log phase, washed three times with sterile saline, and mixed with a ratio of 1:3. The mixture was dropped onto a sterile filter membrane and placed on pre-warmed LB agar plate containing 2, 6-diaminopimelic acid (DAP) (final concentration 50  $\mu$ g/ml). After incubation for 3 h at 37°C, the filter membrane was taken and the bacteria were washed off with LB and plated on LB agar plate containing nalidixic acid (final concentration 20  $\mu$ g/ml) and chloramphenicol (final concentration 50  $\mu$ g/ml) followed by incubation overnight at 37°C. It was found that after transconjugation the colonies that grew on the chloramphenicol-containing LB plates were not all transposon mutants but also contained some donor cells although they should not grow. Therefore, a second-round screening was performed by patching individual colonies to chloramphenicol-containing LB plates followed by incubation at 37°C overnight. The colonies that can grow again on the chloramphenicol-containing plates were further verified by PCR amplification with primer pair *kpsM*-F/R which targets the *kpsM* gene, a specific gene present in PCN033 strain but absent in  $\chi$ 7213 strain, Transposon-Cmr-F/R which targets the chloramphenicol resistance cassette, and R6K-F/R which target the backbone of pSAM-Tat-AprA plasmid. The transposon mutant library was generated by mixing the individual transposon mutants from different batches which was then stored at -80°C until use.

### Experimental animals and ethics statement

The specific pathogen-free (SPF) Kunming mice were purchased from the Experimental Animal Center, Huazhong Agricultural University, Wuhan, China. All animal experiments were approved by the Laboratory Animal Monitoring Committee of Huazhong Agricultural University and performed according to the recommendations in the Guide for the Care and Use of Laboratory Animals of Hubei Province, China (approval number HZAUMO-2021-0095).

### Animal infection with transposon library

The transposon mutant library stored at -80°C was reactivated by two-time consecutive passages in LB at 37°C. Then, the bacterial culture was grown to the mid-log phase, washed twice with sterile saline, and 100  $\mu$ l of the bacterial suspension containing  $5.6 \times 10^7$  CFU of cells was taken as the input library to intraperitoneally inject each mouse. At 12 h post-infection (hpi), the mice were euthanized, and the brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on chloramphenicol-containing plates. After overnight

incubation at 37°C, the colonies were scraped out from the plates, serving as the output libraries (output<sub>brain</sub>, output<sub>spleen</sub>, output<sub>lung</sub>, respectively). A total of 9 mice were used in the above infection assay with bacteria collected from 3 of the mice were regarded as a replicate.

### ***In vitro cell and invasion adhesion assay and macrophage phagocytosis assay***

Baby hamster kidney cells (BHK-21) BHK-21, and murine monocyte/macrophage cell line RAW 264.7 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum in a 37°C incubator with 5% CO<sub>2</sub>. To do *in vitro* cell adhesion and invasion assay, bacterial cells were subcultured in LB from overnight-grown culture and grown at 37°C to the mid-log phase. The bacterial cells were washed twice with sterile saline and mixed with BHK-21 cells with a multiplicity of infection (MOI) of 10:1 in a six-well plate followed by incubation at 37°C for 1 h. The cell culture was then washed twice with sterile PBS and sterile water was then added which was incubated at 4°C for 1 h for cell lysis. The lysate was serially diluted and applied to LB agar plates for bacterial counting. For macrophage phagocytosis, the bacterial cells were mixed with RAW264.7 cells with a multiplicity of infection (MOI) of 10:1 followed by incubation at 37°C for 1 h. Then, the cells were washed twice with sterile PBS, and the unphagocytized bacterial cells were killed by incubation with PBS containing chloramphenicol (final concentration 50 µg/ml). The macrophage cells were then lysed with sterile water and the lysate was serially diluted and applied to LB agar plates for bacterial counting.

### ***Serum killing assay***

Bacterial cells were subcultured in LB from overnight-grown culture and grown at 37°C. Bacterial cells were subcultured in LB from 37°C to the mid-log phase. Cells were washed with sterile saline and 5 × 10<sup>5</sup> CFU of the cells were incubated with 75% fresh or heat-inactivated serum for 20 min at 37°C. Afterward, the mixture was serially diluted and plated on LB agar plates for viable cell enumeration.

### ***Transposon library sequencing***

The insertion site of transposon on the chromosome of an individual mutant was determined using an arbitrary PCR method as previously described [97]. In brief, the first round of PCR was performed using the

genomic DNA of the transposon mutant as the template with the abridged primer pair CmR-in-1 and Adapter primer1, then the PCR product was used as a template for the second round of amplification using primer pair CmR-in-2 and Adapter primer 2. The PCR product was then sequenced to accurately identify the insertion site on the chromosome. The Tn insertion sites of the Tn library were determined by enriching the chromosomal regions flanking the transposon followed by high throughput sequencing. Briefly, bacterial genomic DNA was extracted from the input and the output libraries using E.Z.N.A.® Bacterial DNA Kit (Omega Bio-Tek, USA). Then, TIANSeq Fragment/Repair/Tailing Module kit (Cat# NG301, TIANGEN Biotech, China) was used for DNA fragmentation, end repair, and 3' end dA-tailing. The Illumina universal adaptors were added using the TIANSeq Quick Connect Module (Cat# NG303, TIANGEN Biotech, China) which was then purified using Magic DNA Select Beads (Cat# M3022, Magic Bio, China). Primer pair F-/R- was used in the first round PCR amplification for the enrichment of the DNA region flanking the transposon. The PCR product was purified and used as the template for the second round amplification using P5/7 primer pair, and the amplification product was purified which was used to construct Illumina sequencing library by using Magic DNA Select Beads (Cat# M3022, Magic Bio, China). The DNA library was sequenced using an Illumina Novaseq sequencer. Sequence data have been submitted to NCBI GEO database with the accession number GSE197084 (The data are accessible via the link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197084> by using the secure token: sbobaoowhdynjor).

### ***TraDIS data analysis***

The above obtained raw sequencing reads were subjected to process by using Trimmomatic to remove sequencing adapter sequences and low-quality reads [98]. The resulting clean reads were aligned to the ExPEC PCN033 (GenBank: CP006632.1) genome using BWA (mem algorithm) [99]. The resulting bam files were sorted by using Sambamba (default parameter) to exclude redundant sequences [100]. Correlation coefficients between duplicates within groups were calculated using the StandardScaler submodule of the sklearn module of the python programming language. DEseq2 was used to compare the change in the total number of insertion reads for each gene between the input and output libraries [101]. Fold-change (FC) > 2 and *p* value <0.05 was



used as the threshold to identify differentially enriched genes.

### Growth assay

Cells of the indicated strain were subcultured from overnight-grown cultures into LB, M9 medium, or M9 medium supplemented with the additives in a 96-well plate to give an initial OD<sub>600 nm</sub> value of 0.01. The growth was monitored using an automatic plate reader (Bioscreen C, FP-1100-C, Oy Growth Curves AB, USA) at 37°C with shaking. Five replicates were tested for each strain.

### Competitive infection assay

A competitive infection assay was used to compare the *in vivo* fitness of each mutant strain and the wild-type PCN033 strain (WT) as described previously [31]. Briefly, the overnight-grown cultures of the mutant strain and WT strain were subcultured into LB and grown to the mid-log phase at 37°C with shaking. The two strains to be tested were mixed with a ratio of 1:1 which contained a total of  $6 \times 10^5$  CFU of cells. The bacterial suspension was used to intraperitoneally inject mice with five mice in each group. At 12 hpi, mice were euthanized, and the brain, spleen, and lung tissues were taken, homogenized, diluted, and plated on LB agar plates with and without appropriate antibiotics, respectively, to distinguish the mutant and the wild-type cells. The bacteria were enumerated and the competition index (CI) was calculated as the mutant-to-WT ratio within the recovered sample divided by the corresponding ratio in the inoculum.

### Mouse survival assay

ExPEC strains were subcultured from overnight-grown cultures into LB and grown to the mid-log phase at 37°C with shaking. The cells were harvested by centrifugation, washed with sterile saline, and diluted to appropriate concentrations with saline. 0.2 ml of the bacteria suspension containing approximately  $6.9 \times 10^5$  CFU was used to infect 4-week-old Kunming mice intraperitoneally with 6 mice in each group. Serial dilutions of each inoculum were plated to confirm the dose. The survival of mice was recorded. Statistical differences in survival between the mutant group and the WT group were determined using the log-rank test.

### Statistical analysis

Statistical analysis was performed using GraphPad Prism (version 8) software. The Student's t-test

(unpaired, two-tail) was used to calculate the statistical difference between the two groups as well as the statistical difference of the competitive indices. Significant differences are indicated by “\*”“\*\*”“\*\*\*”, where  $p > 0.05$ , ns;  $0.05 > p > 0.01$ , \*;  $0.01 > p > 0.001$ , \*\*;  $p < 0.001$ , \*\*\*. Error bars in the graphs represent the standard deviations of the means.

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### Funding

This work was supported by the National Key Research and Development Program of China [2022YFD1800902 and 2021YFD1800401] and the National Natural Science Foundation of China [31802211].

### Data availability statement

The Tn-seq data are accessible via the link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197084> by the using secure token: sbobaoowhdynjor.

### ORCID

Qi Huang  <http://orcid.org/0000-0002-1600-1169>

### References

- [1] Smith JL, Fratamico PM, Gunther NW. Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog Dis.* 2007;4(2):134–163.
- [2] Köhler CD, Dobrindt U. What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol.* 2011;301(8):642–647.
- [3] Leimbach A, Hacker J, Dobrindt UE. *Coli* as an all-rounder: the thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol.* 2013; 358:3–32.
- [4] Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, et al. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathog.* 2019;11(1):10.
- [5] Minardi D, D’Anzeo G, Cantoro D, et al. Urinary tract infections in women: etiology and treatment options. *Int J Gen Med.* 2011;4:333–343.
- [6] Guglietta A. Recurrent urinary tract infections in women: risk factors, etiology, pathogenesis and prophylaxis. *Future Microbiol.* 2017;12(3):239–246.
- [7] Bonten M, Johnson JR, van den Biggelaar AHJ, et al. Epidemiology of *Escherichia coli* bacteremia: a systematic literature review. *Clin Infect Dis.* 2021;72(7):1211–1219.
- [8] Kathayat D, Lokesh D, Ranjit S, et al. Avian pathogenic *Escherichia coli* (APEC): an overview of virulence and



- pathogenesis factors, zoonotic potential, and control strategies. *Pathogens*. 2021;10(4):10.
- [9] Ma J, Cheng Z, Bai Q, et al. Screening virulence factors of porcine extraintestinal pathogenic *Escherichia coli* (an emerging pathotype) required for optimal growth in swine blood. *Transbound Emerg Dis*. 2021;68(4):2005–2016.
- [10] Kong LC, Guo X, Wang Z, et al. Whole genome sequencing of an ExPEC that caused fatal pneumonia at a pig farm in Changchun, China. *BMC Vet Res*. 2017;13(1):169.
- [11] Zhu Y, Dong W, Ma J, et al. Characterization and virulence clustering analysis of extraintestinal pathogenic *Escherichia coli* isolated from swine in China. *BMC Vet Res*. 2017;13(1):94.
- [12] Ma J, Pan X, Zhong X, et al. Preferential use of carbon central metabolism and anaerobic respiratory chains in porcine extraintestinal pathogenic *Escherichia coli* during bloodstream infection. *Vet Microbiol*. 2020;249:108830.
- [13] Chen X, Gao S, Jiao X, et al. Prevalence of serogroups and virulence factors of *Escherichia coli* strains isolated from pigs with postweaning diarrhoea in eastern China. *Vet Microbiol*. 2004;103(1–2):13–20.
- [14] Jakobsen L, Spangholm DJ, Pedersen K, et al. Broiler chickens, broiler chicken meat, pigs and pork as sources of ExPEC related virulence genes and resistance in *Escherichia coli* isolates from community-dwelling humans and UTI patients. *Int J Food Microbiol*. 2010;142(1–2):264–272.
- [15] Touchon M, Hoede C, Tenaillon O, et al. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet*. 2009;5(1):e1000344.
- [16] Johnson JR, Russo TA. Extraintestinal pathogenic *Escherichia coli*: “the other bad *E coli*”. *J Lab Clin Med*. 2002;139(3):155–162.
- [17] Korhonen TK, Väisänen V, Saxén H, et al. P-antigen-recognizing fimbriae from human uropathogenic *Escherichia coli* strains. *Infect Immun*. 1982;37(1):286–291.
- [18] Lane MC, Mobley HL. Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney Int*. 2007;72(1):19–25.
- [19] Hull RA, Hull SI. Nutritional requirements for growth of uropathogenic *Escherichia coli* in human urine. *Infect Immun*. 1997;65(5):1960–1961.
- [20] Wijetunge DS, Gongati S, DebRoy C, et al. Characterizing the pathotype of neonatal meningitis causing *Escherichia coli* (NMEC). *BMC Microbiol*. 2015;15(1):211.
- [21] Bonacorsi S, Bingen E. Molecular epidemiology of *Escherichia coli* causing neonatal meningitis. *Int J Med Microbiol*. 2005;295(6–7):373–381.
- [22] Waters VL, Crosa JH. Colicin V virulence plasmids. *Microbiol Rev*. 1991;55(3):437–450.
- [23] Huja S, Oren Y, Trost E, et al. Genomic avenue to avian colisepticemia. *MBio*. 2015;6(1):6.
- [24] Ranjan A, Shaik S, Nandanwar N, et al. Comparative genomics of *Escherichia coli* isolated from skin and soft tissue and other extraintestinal infections. *MBio*. 2017;8:4.
- [25] Tourret J, Denamur E, Mulvey MA. Population phylogenomics of extraintestinal pathogenic *Escherichia coli*. *Microbiol Spectr*. 2016;4:1.
- [26] Ding Y, Tang X, Lu P, et al. Clonal analysis and virulent traits of pathogenic extraintestinal *Escherichia coli* isolates from swine in China. *BMC Vet Res*. 2012;8(1):140.
- [27] Liu C, Zheng H, Yang M, et al. Genome analysis and in vivo virulence of porcine extraintestinal pathogenic *Escherichia coli* strain PCN033. *BMC Genomics*. 2015;16(1):717.
- [28] Tang X, Tan C, Zhang X, et al. Antimicrobial resistances of extraintestinal pathogenic *Escherichia coli* isolates from swine in China. *Microb Pathog*. 2011;50(5):207–212.
- [29] Zong B, Zhang Y, Wang X, et al. Characterization of multiple type-VI secretion system (T6SS) VgrG proteins in the pathogenicity and antibacterial activity of porcine extra-intestinal pathogenic *Escherichia coli*. *Virulence*. 2019;10(1):118–132.
- [30] Lu W, Tan J, Lu H, et al. Function of Rhs proteins in porcine extraintestinal pathogenic *Escherichia coli* PCN033. *J Microbiol*. 2021;59(9):854–860.
- [31] Liu J, Yin F, Liu T, et al. The tat system and its dependent cell division proteins are critical for virulence of extra-intestinal pathogenic *Escherichia coli*. *Virulence*. 2020;11(1):1279–1292.
- [32] van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods*. 2009;6(10):767–772.
- [33] Cain AK, Barquist L, Goodman AL, et al. A decade of advances in transposon-insertion sequencing. *Nat Rev Genet*. 2020;21(9):526–540.
- [34] Luan SL, Chaudhuri RR, Peters SE, et al. Generation of a Tn5 transposon library in *Haemophilus parasuis* and analysis by transposon-directed insertion-site sequencing (TraDIS). *Vet Microbiol*. 2013;166(3–4):558–566.
- [35] Barquist L, Boinett CJ, Cain AK. Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol*. 2013;10(7):1161–1169.
- [36] Bachman MA, Breen P, Deornellas V, et al. Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung infection. *MBio*. 2015;6(3):e00775.
- [37] Goodall ECA, Robinson A, Johnston IG, et al. The essential genome of *Escherichia coli* K-12. *MBio*. 2018;9:1.
- [38] Coe KA, Lee W, Stone MC, et al. Multi-strain Tn-seq reveals common daptomycin resistance determinants in *Staphylococcus aureus*. *PLOS Pathog*. 2019;15(11):e1007862.
- [39] Tan C, Xu Z, Zheng H, et al. Genome sequence of a porcine extraintestinal pathogenic *Escherichia coli* strain. *J bacteriol*. 2011;193(18):5038.
- [40] Tan C, Tang X, Zhang X, et al. Serotypes and virulence genes of extraintestinal pathogenic *Escherichia coli* isolates from diseased pigs in China. *Vet j*. 2012;192(3):483–488.
- [41] Yang R, Liu W, Miao L, et al. Induction of VEGFA and Snail-1 by meningitic *Escherichia coli* mediates disruption of the blood-brain barrier. *Oncotarget*. 2016;7(39):63839–63855.

- [42] Beuzón CR, Holden DW. Use of mixed infections with salmonella strains to study virulence genes and their interactions in vivo. *Microbes Infect.* 2001;3(14–15):1345–1352.
- [43] Pierce JR, Earhart CF. Escherichia coli K-12 envelope proteins specifically required for ferrienterobactin uptake. *J Bacteriol.* 1986;166(3):930–936.
- [44] Yang X, Yu L, He D, et al. The quinone-binding site in succinate-ubiquinone reductase from Escherichia coli. quinone-binding domain and amino acid residues involved in quinone binding. *J Biol Chem.* 1998;273(48):31916–31923.
- [45] Augustus AM, Spicer LD. The MetJ regulon in gammaproteobacteria determined by comparative genomics methods. *BMC Genomics.* 2011;12(1):558.
- [46] Nishino K, Honda T, Yamaguchi A. Genome-wide analyses of Escherichia coli gene expression responsive to the BaeSR two-component regulatory system. *J Bacteriol.* 2005;187(5):1763–1772.
- [47] Zhang B, Ku X, Yu X, et al. Prevalence and antimicrobial susceptibilities of bacterial pathogens in Chinese pig farms from 2013 to 2017. *Sci Rep.* 2019;9(1):9908.
- [48] Mandal RK, Jiang T, Kwon YM. Genetic determinants in salmonella enterica serotype typhimurium required for overcoming in vitro stressors in the mimicking host environment. *Microbiol Spectr.* 2021;9(3):e0015521.
- [49] Crabbe A, Ledesma MA, Nickerson CA. Mimicking the host and its microenvironment in vitro for studying mucosal infections by pseudomonas aeruginosa. *Pathog Dis.* 2014;71(1):1–19.
- [50] Krismer B, Liebeke M, Janek D, et al. Nutrient limitation governs staphylococcus aureus metabolism and niche adaptation in the human nose. *PLOS Pathog.* 2014;10(1):e1003862.
- [51] Santiago M, Matano LM, Moussa SH, et al. A new platform for ultra-high density staphylococcus aureus transposon libraries. *BMC Genomics.* 2015;16(1):252.
- [52] McCarthy AJ, Stabler RA, Taylor PW. Genome-wide identification by transposon insertion sequencing of Escherichia coli K1 genes essential for in vitro growth, gastrointestinal colonizing capacity, and survival in serum. *J Bacteriol.* 2018;200(7):200.
- [53] Kwon YM, Ricke SC, Mandal RK. Transposon sequencing: methods and expanding applications. *Appl Microbiol Biotechnol.* 2016;100(1):31–43.
- [54] Goodman AL, McNulty NP, Zhao Y, et al. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe.* 2009;6(3):279–289.
- [55] Li M, Rigby K, Lai Y, et al. Staphylococcus aureus mutant screen reveals interaction of the human antimicrobial peptide dermcidin with membrane phospholipids. *Antimicrob Agents Chemother.* 2009;53(10):4200–4210.
- [56] Segura M, Fittipaldi N, Calzas C, et al. Critical streptococcus suis virulence factors: are they all really critical? *Trends Microbiol.* 2017;25(7):585–599.
- [57] He XH, Ouyang DY, Xu LH. Injection of Escherichia coli to induce sepsis. *Methods Mol Biol.* 2021; 2321 43–51.
- [58] Zhang H, Chen X, Nolan LK, et al. Identification of host adaptation genes in extraintestinal pathogenic Escherichia coli during infection in different hosts. *Infect Immun.* 2019;87:12.
- [59] Riley LW. Extraintestinal foodborne pathogens. *Annu Rev Food Sci Technol.* 2020;11(1):11.
- [60] Buckles EL, Luterbach CL, Wang X, et al. Signature-tagged mutagenesis and co-infection studies demonstrate the importance of P fimbriae in a murine model of urinary tract infection. *Pathog Dis.* 2015;73(4):ftv014.
- [61] Chen SL, Hung CS, Pinkner JS, et al. Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *Proc Natl Acad Sci U S A.* 2009;106:22439–22444.
- [62] Staerk K, Gronnemoose RB, Nielsen TK, et al. Escherichia coli type-1 fimbriae are critical to overcome initial bottlenecks of infection upon low-dose inoculation in a porcine model of cystitis. *Microbiol (Reading).* 2021;167(10): 001101.
- [63] Caza M, Lepine F, Dozois CM. Secretion, but not overall synthesis, of catecholate siderophores contributes to virulence of extraintestinal pathogenic Escherichia coli. *Mol Microbiol.* 2011;80:266–282.
- [64] Mu X, Gao R, Xiao W, et al. EntE, EntS and TolC synergistically contributed to the pathogenesis of APEC strain E058. *Microb Pathog.* 2020;141:103990.
- [65] Rabsch W, Voigt W, Reissbrodt R, et al. Salmonella typhimurium IroN and FepA proteins mediate uptake of enterobactin but differ in their specificity for other siderophores. *J Bacteriol.* 1999;181:3610–3612.
- [66] Nagy TA, Moreland SM, Andrews-Polymenis H, et al. The ferric enterobactin transporter Fep is required for persistent salmonella enterica serovar typhimurium infection. *Infect Immun.* 2013;81:4063–4070.
- [67] Perez-Casal JF, Crosa JH. Aerobactin iron uptake sequences in plasmid ColV-K30 are flanked by inverted IS1-like elements and replication regions. *J Bacteriol.* 1984;160:256–265.
- [68] Koh EI, Hung CS, Henderson JP. The yersiniabactin-associated ATP binding cassette proteins YbtP and YbtQ enhance Escherichia coli fitness during high-titer cystitis. *Infect Immun.* 2016;84:1312–1319.
- [69] Tomas JM, Benedi VJ, Ciurana B, et al. Role of capsule and O antigen in resistance of Klebsiella pneumoniae to serum bactericidal activity. *Infect Immun.* 1986;54:85–89.
- [70] Bravo D, Silva C, Carter JA, et al. Growth-phase regulation of lipopolysaccharide O-antigen chain length influences serum resistance in serovars of salmonella. *J Med Microbiol.* 2008;57:938–946.
- [71] Bao Y, Zhang H, Huang X, et al. O-specific polysaccharide confers lysozyme resistance to extraintestinal pathogenic Escherichia coli. *Virulence.* 2018;9:666–680.
- [72] Kintz E, Heiss C, Black I, et al. Salmonella enterica serovar typhi lipopolysaccharide O-antigen modification impact on serum resistance and antibody recognition. *Infect Immun.* 2017;85(4): e01021–16.
- [73] Alteri CJ, Smith SN, Mobley HL. Fitness of Escherichia coli during urinary tract infection requires

- gluconeogenesis and the TCA cycle. *PLOS Pathog.* **2009**;5:e1000448.
- [74] Kuo CJ, Wang ST, Lin CM, et al. A multi-omic analysis reveals the role of fumarate in regulating the virulence of enterohemorrhagic *Escherichia coli*. *Cell Death Dis.* **2018**;9:381.
- [75] Tchawa Yimma M, Leatham MP, Allen JH, et al. Role of gluconeogenesis and the tricarboxylic acid cycle in the virulence of salmonella enterica serovar typhimurium in BALB/c mice. *Infect Immun.* **2006**;74:1130–1140.
- [76] Olive AJ, Sassetti CM. Metabolic crosstalk between host and pathogen: sensing, adapting and competing. *Nat Rev Microbiol.* **2016**;14:221–234.
- [77] Tannahill GM, Curtis AM, Adamik J, et al. Succinate is an inflammatory signal that induces IL-1 $\beta$  through HIF-1 $\alpha$ . *Nature.* **2013**;496:238–242.
- [78] Olson ER, Dunyak DS, Jurss LM, et al. Identification and characterization of dppA, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. *J Bacteriol.* **1991**;173:234–244.
- [79] Harder D, Stolz J, Casagrande F, et al. DtpB (YhiP) and DtpA (TppB, YdgR) are prototypical proton-dependent peptide transporters of *Escherichia coli*. *FEBS J.* **2008**;275:3290–3298.
- [80] Ernst HA, Pham A, Hald H, et al. Ligand binding analyses of the putative peptide transporter YjdL from *E. coli* display a significant selectivity towards dipeptides. *Biochem Biophys Res Commun.* **2009**;389:112–116.
- [81] Casagrande F, Harder D, Schenk A, et al. Projection structure of DtpD (YbgH), a prokaryotic member of the peptide transporter family. *J Mol Biol.* **2009**;394:708–717.
- [82] Golby P, Kelly DJ, Guest JR, et al. Transcriptional regulation and organization of the dcuA and dcuB genes, encoding homologous anaerobic C4-dicarboxylate transporters in *Escherichia coli*. *J Bacteriol.* **1998**;180:6586–6596.
- [83] Uden G, Strecker A, Kleefeld A, et al. C4-dicarboxylate utilization in aerobic and anaerobic growth. *EcoSal Plus.* **2016**;7(1).
- [84] Schellenberg GD, Furlong CE. Resolution of the multiplicity of the glutamate and aspartate transport systems of *Escherichia coli*. *J Biol Chem.* **1977**;252:9055–9064.
- [85] Clegg S, Yu F, Griffiths L, et al. The roles of the polytopic membrane proteins NarK, NarU and NirC in *Escherichia coli* K-12: two nitrate and three nitrite transporters. *Mol Microbiol.* **2002**;44:143–155.
- [86] Clegg SJ, Jia W, Cole JA. Role of the *Escherichia coli* nitrate transport protein, NarU, in survival during severe nutrient starvation and slow growth. *Microbiol (Reading).* **2006**;152:2091–2100.
- [87] Bourgeois JS, Zhou D, Thurston TLM, et al. Methylthioadenosine suppresses salmonella virulence. *Infect Immun.* **2018**;86(9): e00429–18.
- [88] Bogard RW, Davies BW, Mekalanos JJ. MetR-regulated vibrio cholerae metabolism is required for virulence. *MBio.* **2012**;3(5): e00236–12.
- [89] Cubitt MF, Hedley PE, Williamson NR, et al. A metabolic regulator modulates virulence and quorum sensing signal production in *Pectobacterium atrosepticum*. *Mol Plant Microbe Interact.* **2013**;26:356–366.
- [90] Bourgeois JS, Anderson CE, Wang L, et al. Integration of the salmonella typhimurium methylome and transcriptome reveals that DNA methylation and transcriptional regulation are largely decoupled under virulence-related conditions. *MBio.* **2022**;13:e0346421.
- [91] Subramanian K, Henriques-Normark B, Normark S. Emerging concepts in the pathogenesis of the streptococcus pneumoniae: from nasopharyngeal colonizer to intracellular pathogen. *Cell Microbiol.* **2019**;21: e13077.
- [92] Iuchi S, Cole ST, Lin EC. Multiple regulatory elements for the glpA operon encoding anaerobic glycerol-3-phosphate dehydrogenase and the glpD operon encoding aerobic glycerol-3-phosphate dehydrogenase in *Escherichia coli*: further characterization of respiratory control. *J Bacteriol.* **1990**;172:179–184.
- [93] Peng Y, Wang X, Shou J, et al. Roles of Hcp family proteins in the pathogenesis of the porcine extraintestinal pathogenic *Escherichia coli* type VI secretion system. *Sci Rep.* **2016**;6:26816.
- [94] Chao MC, Abel S, Davis BM, et al. The design and analysis of transposon insertion sequencing experiments. *Nat Rev Microbiol.* **2016**;14:119–128.
- [95] Platt R, Drescher C, Park SK, et al. Genetic system for reversible integration of DNA constructs and lacZ gene fusions into the *Escherichia coli* chromosome. *Plasmid.* **2000**;43:12–23.
- [96] Roland K, Curtiss R, Sizemore D. Construction and evaluation of a delta cya delta crp salmonella typhimurium strain expressing avian pathogenic *Escherichia coli* O78 LPS as a vaccine to prevent airsacculitis in chickens. *Avian Dis.* **1999**;43:429–441.
- [97] Chun KT, Edenberg HJ, Kelley MR, et al. Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. *Yeast.* **1997**;13:233–240.
- [98] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinf.* **2014**;30:2114–2120.
- [99] Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinf.* **2009**;25: 1754–1760.
- [100] Tarasov A, Vilella AJ, Cuppen E, et al. Sambamba: fast processing of NGS alignment formats. *Bioinf.* **2015**;31:2032–2034.
- [101] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**;15:550.