Identification of novel biofilm genes in avian pathogenic Escherichia coli by Tn5 transposon mutant library

Jiangang Hu
Anhui Agricultural University
Yi Gu
Anhui Agricultural University
Huiqi Lu
Anhui Agricultural University
Muhammad Akmal Raheem
Anhui Agricultural University
Fangheng Yu
Anhui Agricultural University
Xiangpeng Niu
Anhui Agricultural University
Jiakun Zuo
Shanghai Veterinary Research Institute Chinese Academy of Agricultural Sciences
Huifang Yin
Longyan University
Cuiqin Huang
Longyan University
Xiangjun Song
Anhui Agricultural University
Jian Tu
Anhui Agricultural University
Wen Zhou
Shanghai Veterinary Research Institute Chinese Academy of Agricultural Sciences
Wei Jiang
Shanghai Veterinary Research Institute Chinese Academy of Agricultural Sciences
Zhaoguo Chen
Shanghai Veterinary Research Institute Chinese Academy of Agricultural Sciences
Xiangan Han (✉ hanxgan@163.com)
Shanghai Veterinary Research Institute https://orcid.org/0000-0001-5625-637X
Kezong Qi

Anhui Agricultural University

Research Article

Keywords: Avian pathogenic Escherichia coli, Biofilm-related genes, Tn5, Gene deletion, yehE

Posted Date: March 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1373401/v1

Abstract

Avian pathogenic *Escherichia coli* (APEC) is the main pathogens that inflict the poultry industry. Biofilm as the pathogenic factors of APEC, which can enhance the anti-host immune system of APEC and improve its survival in the environment. In order to screen new genes related to APEC biofilm. The APEC strain APEC81 was used to construct a mutant library by Tn5 insertion mutagenesis. Moreover the 28 mutant strains with severely weakened biofilm were successfully screened from 1500 mutant strains by crystal violet staining, in which 17 genes were obtained by high-efficiency thermal asymmetric interlaced PCR (HiTAIL-PCR). The reported genes including 3 flagella genes (*fliS, fliD, fliR*), 4 curli fimbriae genes (*csgD, csgA*, *csgF*, *csgG*) and 3 type 1 fimbriae genes (*fimA*, *fimD*, *fimC*). The novel genes including 3 coenzyme genes *(gltA, bglX, mltF)* and 4 putative protein genes (*yheE, 07045, 11735, 11255*). To investigate whether these 17 genes co-regulate the biofilm, the 17 identified genes were deleted on APEC strain APEC81. The result shown that except for the *11735* and *11255* genes, the deletion of 15 genes significantly reduced the biofilm formation ability of APEC81 ($P < 0.05$). The result of rdar (red, dry and rough) colony morphology showed that curli fimbriae genes (*csgD*, *csgA*, *csgF*, *csgG*) and other functional genes (*fimC*, *glxK*, *yehE*, *07045*, *11255*) affected the colony morphology. Particularly, the hypothetical protein YehE had the greatest influence on the biofilm. It was predicted to have the same structure as type 1 fimbria protein. When *yehE* was deleted, the *fimE* transcription was up-regulated, *fimA* and *fimB* transcription were down-regulated, resulting in a decrease in type 1 fimbriae. Hence the *yehE* mutant significantly reduced the biofilm and the adhesion and invasion ability to cells ($P < 0.05$). Altogether, this study had identified 5 novel genes (*gltA, bglX, mltF, yheE, 07045*) related to biofilm formation and confirmed that *yehE* affects biofilm formation by type 1 fimbriae, which will benefit for further study mechanism of biofilm regulation in APEC.

Introduction

Avian Pathogenic *Escherichia coli* (APEC) is a highly pathogenic bacteria that can cause local or systemic diseases in poultry and brings huge economic losses to the poultry industry (Hu et al., 2019). Biofilm are heterogeneous congregation of surface associated microorganisms encapsulated consisting of vast exopolysaccharides, secreted proteins, and DNA that can help bacteria enhance their resistance and resistance to external environmental pressure (Filloux and Vallet, 2003). Biofilm renders the penetration of conventional antibiotics hard and makes the cells less susceptible to the antibiotics (Ito et al., 2009; Mittal et al., 2015). Moreover, the biofilm are responsible for causing a broad range of chronic diseases, and due to the emergence of antibiotic resistance in bacteria it has really become difficult to treat them with efficacy (Roy et al., 2018). The transposon is randomly inserted into the entire bacterial genome, and random mutations can be performed without obtaining the gene sequence to generate large-scale mutants, and the phenotype changes of the mutants are screened to reversely identify the transposon inserted gene (Pang et al., 2017). There is currently no research on screening new biofilm-related genes by the Tn5 transposon mutation library in APEC.

In this study, the Tn5 transposon mutation library was constructed to screen biofilm-related genes on APEC81. The selected genes were proved to be involved in biofilm formation by constructing gene deletions in wild-type (WT) strain APEC81. The 17 biofilm-related gene sites were identified not only confirmed the importance of the biofilm-related genes have been reported, but more significant is that 5 novel biofilm-related genes were identified.

Materials And Methods

Strains, growth conditions and plasmid

The *E. coli* strains and plasmids used in this study were listed in Table 1. The bacteria were grown at 37°C or 25℃ in Luria-Bertani (LB) broth or on LB plates containing 1.5% agar. Antibiotics were used at the given concentration when needed, ampicillin (100 μg/mL), chloramphenicol (30 μg/mL), nalidixic acid (50 μg/mL).

Table 1
Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant genotype and property	Source or reference
Escherichia coli strains		
S17-1	Random transposon donor cells	Biomedal
APEC81	Wild type *E. coli*	(Xue et al., 2020)
APEC81Δ*fliS*	APEC81 *fliS* mutant strain	This study
APEC81Δ*fliD*	APEC81 *fliD* mutant strain	This study
APEC81Δ*fliR*	APEC81 *fliR* mutant strain	This study
APEC81Δ*csgA*	APEC81 *csgA* mutant strain	This study
APEC81Δ*csgD*	APEC81 *csgD* mutant strain	This study
APEC81Δ*csgF*	APEC81 *csgF* mutant strain	This study
APEC81Δ*csgG*	APEC81 *csgG* mutant strain	This study
APEC81Δ*fimA*	APEC81 *fimA* mutant strain	This study
APEC81Δ*fimD*	APEC81 *fimD* mutant strain	This study
APEC81Δ*fimC*	APEC81 *fimC* mutant strain	This study
APEC81Δ*gltA*	APEC81 *gltA* mutant strain	This study
APEC81Δ*bglX*	APEC81 *bglX* mutant strain	This study
APEC81Δ*mltF*	APEC81 *mltF* mutant strain	This study
APEC81Δ*yehE*	APEC81 *yehE* mutant strain	This study
APEC81Δ*07045*	APEC81 *07045* mutant strain	This study
APEC81Δ*11735*	APEC81 *11735* mutant strain	This study
APEC81Δ*11255*	APEC81 *11255* mutant strain	This study
Plasmids		
pUTmini-Tn5-Cm	Provide the transposon Tn5 to insert into the chromosome of the target strain	Biomedal
pKD46	The vector containing arabinose-inducible phage λ Red recombinase (Amp^R)	(Datsenko and Wanner, 2000)
pCP20	A helper plasmid encoding the FLP recombinase (Amp^R)	(Datsenko and Wanner, 2000)

Strain or plasmid	Relevant genotype and property	Source or reference
pKD3	Chloramphenicol resistance (Cm^R) cloning vector	(Datsenko and Wanner, 2000)
pSTV28	Cm^R Cloning vector	TaKaRa
pSTV28-yehE	Covering the *yehE* gene	This study

Random transposon mutagenesis

The Tn5 was introduced into APEC strain APEC81 by conjugation from *E. coli* S17-1 as described previously with some modifications (X. Wang et al., 2015). Briefly, the recipient and donor cells were cultured to the mid-log phase, mixed at a ratio of 4:5 (based on optical density at 600nm, OD_{600}), and filtered onto a 0.22 μm membrane (EMD Millipore, MA, USA). The filter was placed on a TSA plate and incubated at 30°C for 12 h. After incubation, the cells were scraped off the filter, resuspended in 10 mM $MgSO_4$, and plated onto TSA supplemented with chloramphenicol (30 μg/ml) and nalidixic acid (50 μg/ml) to select for potential transconjugants. The strains which amplified Tn5 fragments identified by PCR were identified as the Tn5-disrupted mutant (primer sequences were shown in supplementary material Table S1).

Screening of random mutants

An assay involving crystal violet staining was used to quantify biofilm formation of APEC as previously described (Zuo et al., 2019a). APEC81 and Tn5 transposon random insertion strains were cultured until they reached the logarithmic growth phase. Then centrifugation at 5500 rpm at 4°C for 10 min to collect bacteria. Bacteria were washed by LB three times and adjusted to OD_{600} of 1. Bacteria are diluted 50 times and 200 μL cultures were added to 96 wells plate (Corning Inc., NY, USA) three groups of bacteria were added in parallel, the LB media was added to 96 well microtiter plates as negative control. The previous study found that APEC81 strain had a relatively strong biofilm ability 25℃, so the plates were incubated static incubator at 25℃ for 48 h. The plates were washed by PBS three times and then dried, dyed with 0.1% crystal violet for about 20 minutes, poured out the crystal violet, washed by PBS three times and then dry. The biofilm was dissolved in 95% alcohol (200 μL per well), and OD_{595} was detected by a microplate reader (BioTek, Winooski, VT, USA).

Identification of insertion position of Tn5 transposon

The sites of transposon insertions in the mutants were determined by cloning the Tn5-disrupted genes by genomic walking. Genomic walking was performed with a genome walking kit (TaKaRa Bio Inc, Dalian, China) according to the manufacturer's instructions. Three specific primers (SP1, SP2, SP3, see the supplemental material Table S1) with the same direction and higher annealing temperature were designed and used for thermal asymmetric PCR with four of them provided in the kit AP1, AP2, AP3, AP4, which were specially designed and annealed at a lower temperature. Using four pairs of primers, AP1 and

SP1, AP2 and SP1, AP3 and SP1, AP4 and SP1 and the mutant strain DNA as the template for primary TAIL-PCR. The product was diluted and used as the template for secondary TAIL-PCR, four pairs of primers was AP1 and SP2, AP2 and SP2, AP3 and SP2, AP4 and SP2. The product was diluted and used as the template for the third TAIL-PCR, four pairs of primers were AP1 and SP3, AP2 and SP3, AP3 and SP3, AP4 and SP3. After purifying the PCR products from agarose gels with the GeneJET gel extraction kit (Thermo Fisher Scientific, MA, USA), the PCR products were sequenced by Sangon Biotech Co., Ltd. (Sangon, Shanghai, China). The flanking sequences were then BLAST searched against the APEC O_{78} genome (NCBI reference sequence: NC_020163.1) and other *E. coli* sequences in the NCBI database to determine which genes were interrupted by the transposon (http://www.ncbi.nlm.nih.gov/BLAST/).

Construction of deletion strain and complementary strain

The Tn5-disrupted genes were deleted by red homologous recombination technology. Meanwhile, we constructed deletions from wild strain APEC81. The primer sequences were shown in supplementary material Table S1. Gene-cm-F/gene-cm-R as primers and pKD3 were used as template, the chloramphenicol fragment containing homologous arm was amplified and transformed into APEC81 competent cells containing pKD46 plasmid for gene deletion. The positive strains were screened and detected by gene-out-F/ gene-out-R to detect whether the target gene was replaced by chloramphenicol fragment, and the pCP20 plasmid was used for antibody elimination, and the results of antibody elimination were verified by PCR with gene-out-F/gene-out-R and gene-in-F/gene-in-R primers. We constructed 17 mutant strains. For functional complementation, the DNA sequence covering the *yehE* open reading frame was amplified from the chromosomal DNA of APEC81 with the primers C-yehE-F/C-yehE-R and cloned into pSTV28 to yield the plasmid pSTV28-yehE. The pSTV28-yehE was transformed into the APEC81Δ*yehE*, and grown on LB agar plates containing chloromycetin (30 μg/mL). The resulting *yehE* complementation strain was named cAPEC81Δ*yehE*.

Biofilm forming ability of mutant strains

These deletion strains have similar growth ability to APEC81 by measuring the growth curve (see the supplemental material Fig. S1). The method in 2.3 was used to detect the biofilm forming ability of the mutant strains, and used GraphPad Prism 6 software and SPSS V19.0 software to analyze the difference between the mutant strains and the wild strain.

Rdar colony morphology

The assay of rdar morphotype was performed with some modifications(Wu et al., 2017). In order to see the expression of curli fimbriae and cellulose macroscopically, strains were grown to the logarithmic phase (OD_{600} = 1.0), and then the cells were centrifuged for 10 min at 5500 rpm and 4°C, then the bacteria were washed twice with PBS and the OD_{600} was adjusted to 3, 2 μL of bacterial suspension was dropped onto a salt free LB medium containing 40 μg/ml Congo red (Sangon, Shanghai, China) and 20 μg/ml coomassie brilliant blue (Sangon, Shanghai, China). The mutants were cultured at 25℃ on the same plate.

Prediction of protein function and regulation

The decrease of biofilm was most significant after the deletion of hypothetical protein YehE gene in 4 putative proteins. To determine the function of the hypothetical protein YehE, the protein-related networks and protein domains were predicted by String 11.0 (https://string-db.org/) and Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html).

Mutant of hypothetical proteins yehE adherence and invasion assays

After prediction analysis, the hypothetical protein YehE has the same structure as type 1 fimbria protein. Therefore, we want to further study the influence of hypothesized proteins YehE on the adhesion and invasion to host cells. The Hela cells were used for bacterial adhesion tests, HD11 chicken macrophages cells were used for bacterial invasion tests as described previously (Crepin et al., 2017; Zuo et al., 2019a). Hela or HD11 cells were seeded in 24-well tissue culture plates in DMEM containing 15% fetal bovine serum. APEC81, APEC81Δ*yehE* and c APEC81Δ*yehE* strains were grown in LB medium at 37°C to the mid-log phase of growth (OD_{600} = 1.0). The cells were rinsed with PBS and infected with both APEC81 and APEC81Δ*yehE* at a multiplicity of infection (MOI) of 100 after incubation at 37°C with 5% CO_2 for 24 h. After adhesion for 2 h, cells were washed three times with PBS, then the cells were lysed with 0.5% TritonX-100. The cell lysate was 10-fold diluted and spread onto LB plates for bacterial counting. For the invasion assay, DMEM medium containing 100 μg/mL gentamicin was added to the wells for one additional hour to kill the extracellular bacteria. Bacterial cells were then counted as described above. All assays were performed in triplicate.

RNA extraction and quantification of gene expression

To study the effect of hypothetical proteins YehE on biofilm and pathogenicity. The genes affected by the hypothetical protein YehE were analyzed, including type 1 fimbriae genes *fimA*, *fimB* and *fimE*, which expression was evaluated by quantitative real time-PCR, as described previously with some modifications (primer sequences were shown in Table S1) (Zuo et al., 2019b). Briefly, bacteria were grown in LB medium at 37°C to the mid-log phase of growth (OD_{600} = 1.0) and total RNA was extracted by TRIzol reagent (Invitrogen Corporation, USA) according to the manufacturer's protocol. The RNA was treated with PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa Bio Inc, Dalian, China) to remove genomic DNA, and then reverse transcribed into cDNA (complementary DNA). The SYBR Green based two-step qRT-PCR (TB Green® Premix Ex Taq™ II, TaKaRa Bio Inc, Dalian, China) amplification was performed. Relative changes in gene expression level were assessed using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The *dnaE* gene was used as an internal control. The final qRT-PCR data was presented as the means of three separate experiments.

Statistical analysis

Statistical analyses were conducted using SPSS V19.0 software (SPSS Inc., Chicago, IL, USA). The student's t-test was used to analyze the data and P values less than 0.05 were considered significant.

Results

Large scale isolation of biofilm-defective strains

The biofilm formation ability of 1500 transposon mutant strains were detected by crystal violet staining. The results indicated that the biofilm formation ability of 12 mutants were increased (accounted for 0.8%) ($P < 0.05$), 1460 mutants (accounted for 97.3%) was no significant difference ($P > 0.05$) and 28 mutants (accounted for 1.9%) significantly decreased biofilm formation was observed with mutants compared with the WT strain ($P < 0.05$) (Fig. 1 and Fig. 2).

Location of Tn5 inserted sites

The transposon insertion sites were located by HiTAIL-PCR of the genomic DNAs extracted from the 28 mutants, which significantly decreased biofilm formation. The 17 different insertion sites were identified from the 28 mutants. The proteins of insertion site genes are classified into six classes, flagellum (flagellar biosynthesis and structure), type 1 fimbriae, curli fimbriae, metabolic pathways (Coenzymes), extracellular matrix and other functions (hypothetical proteins) (shown in Table 2). This result reveals that different mutants have the same transposon insertion sites, *fimD* (four mutants), *fimA* (three mutants), *fimC* (three mutants), *csgD* (two mutants), *mltF* (two mutants), *fliR* (two mutants), hypothetical protein *07045* (two mutants). The 17 transposon targeted genes were randomly distributed throughout the strain APEC81 genome. The locations of these genes in APEC81 were shown in Fig. 3.

Table 2
Identification of Tn5 transposon insertion sites

Function group	Strains	Gene name	Locus tag	Protein function	Decrease rate of biofilm
Flagellum	Tn-135	*fliS*	APECO78_13415	Flagellar protein FliS	76.4%
	Tn-189	*fliD*	APECO78_13410	Flagellar capping protein FliD	80.9%
	Tn-521 Tn-576	*fliR*	APECO78_13540	Flagellar biosynthesis protein FliR	74.5%
Curli fimbriae	Tn-422	*csgA*	APECO78_09180	Curli major subunit CsgA	66.1%
	Tn-371 Tn-798	*csgD*	APECO78_09170	DNA-binding transcriptional regulator CsgD	72.4%
	Tn-356	*csgF*	APECO78_09160	Curli assembly protein CsgF	54.4%
	Tn-874	*csgG*	APECO78_09155	Curli production assembly/transport protein CsgG	52.9%
Type 1 fimbriae	Tn-107 Tn-454 Tn-1492	*fimA*	APECO78_02710	Major type 1 fimbrial subunit FimA	83.2%
	Tn-74 Tn-238 Tn-1373 Tn-1414	*fimD*	APECO78_02725	Type 1 fimbrial outer membrane usher protein FimD	86.0%
	Tn-757 Tn-788 Tn-1476	*fimC*	APECO78_02720	Chaperone protein FimC	83.8%
Coenzyme	Tn-535	*gltA*	APECO78_06065	Glycerate kinase	64.0%
	Tn-054	*bglX*	APECO78_14375	Beta-D-glucoside glucohydrolase	48.0%

Function group	Strains	Gene name	Locus tag	Protein function	Decrease rate of biofilm
	Tn-695 Tn-960	*mltF*	APECO78_16520	Membrane-bound lytic murein transglycosylase F	68.0%
Hypothetical protein	Tn-311 Tn-685	None	APECO78_07045	Hypothetical protein	61.5%
	Tn-080		APECO78_14275	Uncharacterized protein YehE	35.10%
	Tn-842		APECO78_11735	Dimethyl sulfoxide reductase subunit A	51.3%
	Tn-410		APECO78_11255	Putative fimbrial-like adhesin exported protein	48.5%

Biofilm formation ability of deletion strains

To confirm the importance of our identified genes in APEC biofilm formation, the deletion strains were constructed by the red homologous recombination method in APEC81. We successfully constructed 17 genes deletion strains in APEC81 (Table 1). These deletion strains were identified by PCR, and the identification image data was not shown. The biofilm forming ability of the deletion strains were detected at 25℃ by the crystal violet staining method. In addition to the *11735* and *11255* genes, the deletion of other genes significantly reduced the biofilm formation ability of APEC81 ($P < 0.05$, shown in Fig. 4).

The selected genes affected the biofilm to different degrees (as shown in Table 3). The biofilm weakening rate was from 4.1–91.0%. Among these deletion strains, compared with the strain APEC81, the deletion of type 1 fimbriae genes had the greatest effect on the biofilm (from 82.0–91.0%), the flagellum genes caused the rate of biofilm decline from 68.0–81.0%, the curli fimbriae genes caused biofilm decline in the range from 60.9–83.4%, the hypothetical protein genes caused the rate of biofilm decline from 4.1–34.1%. Among the four hypothetical proteins, APEC81Δ*yehE*had the greatest influence on the biofilm, reducing by 29.6%.

Table 3
Decrease rate of biofilm at different mutants

Mutant strains	Decrease rate of biofilm
APEC81Δ*fliS*	77.1%
APEC81Δ*fliD*	73.5%
APEC81Δ*fliR*	72.5%
APEC81Δ*csgA*	74.9%
APEC81Δ*csgD*	83.4%
APEC81Δ*csgF*	76.8%
APEC81Δ*csgG*	60.9%
APEC81Δ*fimA*	87.9%
APEC81Δ*fimD*	88.9%
APEC81Δ*fimC*	86.2%
APEC81Δ*gltA*	58.3%
APEC81Δ*bglX*	56.2%
APEC81Δ*mltF*	43.3%
APEC81Δ*yehE*	29.6%
APEC81Δ*07045*	25.3%
APEC81Δ*11735*	7.5%
APEC81Δ*11255*	5.5%

Rdar colony morphology of mutant strains

The APEC81 stains rdar morphological was investigated on congo red and coomassie brilliant blue agar plates. The rdar morphology of mutant strains were shown in Fig. 5. The APEC81 was a rdar (red, dry and rough) colony morphology on the plates, which proved that it could product curli and cellulose production at 25°C. Compared with APEC81, the strain APEC81△*csgA* displayed a pink and smooth colony morphology on the plate, which proved that it lack the expression of curli fimbriae and the expression of cellulose were weak, APEC81△*csgF*, APEC81△*csgG*, APEC81△*csgD* displayed smooth and white colony morphology on the plate, which proved that the deletion stains lack the expression of curli fimbriae and cellulose. The APEC81△*fimC*, APEC81△*glxK*, APEC81△*yehE*, APEC81△*07045*, APEC81△*11255* was a bdar (brown, dry and rough) colony morphology on the plates, which proved that the deletion stains sole expression of curli and lack the expression of cellulose. The other deletion genes had no effect on the rdar morphology, it does not affect the synthesis of curli fimbriae and cellulose.

Hypothetical protein function prediction

The String and Phyre2 protein prediction software was used to estimate the function and interaction network of the hypothetical protein YehE. It shown that the protein encoded by *yehE* has the same structure as type 1 fimbria protein with 56.6% confidence (Fig. 6a). Moreover, YehE protein also has an interactive relationship with the fiber chaperone molecule YehC with 40% confidence (Fig. 6b).

Mutant of hypothetical proteins YehE decreased bacterial adherence and invasion

As shown in Fig. 6, the mutant of *yehE* decreased the adherence and invasion of cells by APEC. The APEC81 strain adhered to the Hela cells at 15.7%, the mutant of *yehE* significantly decreased the adherence abilities to Hela cells at 10.5% ($P < 0.05$, Fig. 7a). In addition, *yehE* deletion significantly decreased invasion to HD11 cells from 2.9–1.3% ($P < 0.01$, Fig. 7b).

Transcription of the type 1 fimbriae major subunit is decreased in the yehE mutant

To test whether the hypothesized protein gene *yehE* could affect the expression of type 1 fimbriae, the expression of the type 1 fimbriae was analyzed by qRT-PCR, including type 1 fimbriae major subunit *fimA*, *fimB and fimE* recombinases. Compared to that for the WT strain, the *fimA* and *fimB* transcription of *yehE* mutant were down-regulated by 2.1-fold and 1.6-fold, respectively. The *fimE* transcription of *yehE* mutant was up-regulated by 1.6-fold (Fig. 8).

Discussion

The selected genes were deleted in APEC81, and the deletion of these genes significantly down-regulated the biofilm formation ability of APEC81, and the growth ability shows that their down-regulation of biofilm is not caused by reduced growth. Therefore, the reason for the biofilm down-regulation may be that these genes are involved in *E. coli* different systems, which leads to down-regulate biofilm. Although there are continuous studies to explore the genes that affect the *E. coli* biofilm, there is no study available to compare the effects of different systems involved in these genes on the *E. coli* biofilm. Among the gene sites that have been identified in this study, 10 reported genes are associated with biofilm, including three flagella genes, four curli fimbriae genes, and three type 1 fimbriae genes. The results emphasizes that the type 1 fimbriae system had the greatest influence on the biofilm, followed by flagella and curli fimbriae, and other genes screened had less influence on the biofilm than these three systems.

The type 1 fimbriae is encoded by the chromosomally located *fim* gene cluster which encodes the major structural subunit (FimA) and several minor components: two adaptor proteins (FimG and FimF), the adhesion (FimH), two chaperon proteins (FimC and FimD), two site-specific recombinases (FimB and FimE) and the regulator FimI (Blomfield, 2001). The type 1 fimbriae has been proved was necessary for *E. coli* to adhere to non-biological surfaces, and the lack of *fimA* will reduce the adhesion of *E. coli* to non-biological surfaces, the ability of the chaperone FimC and the usher FimD has been proven to be necessary for the surface localization of type 1 fimbriae in *E. coli*, which is the main reason why fimbriae

genes weaken the biofilm (Per Klemm and Christiansen, 1990; Liyun et al., 2018). The decrease of biofilm was most significant after the deletion of hypothetical protein *yehE* gene in four putative proteins at 37℃. To study the function, the *yehE* was predicted about their domains and functions by software. Interestingly, the *yehE* was considered to be related to type 1 fimbria. The type 1 fimbria operon expression is governed by a phase-variable promoter (fimS) which is located on a 314-bp invertible element flanked by two 9-bp inverted repeats (Abraham et al., 1985). Phase-variable expression of type 1 fimbriae is mediated by the inversion of fimS to and from the “on” or “off” orientation. Orientation of fimS is controlled by the FimB and FimE recombinases (P. Klemm, 1986). By this means, *fimE* promotes switching of fimS to the off orientation (from the phase “on” to phase “off”), while FimB mediates switching in both directions (from the phase “off” to phase “on” and phase “on” to phase “off”), where the on orientation is favored (Gally et al., 1996; P. Klemm, 1986). When *yehE* was deleted in WT strain, the *fimE* transcription was up-regulated and *fimB* transcription was down-regulated. This differential expression promotes inversion of the fimS promoter to the off position. Therefore, *fimA* transcription is down-regulated, resulting in weakening of the bacterial biofilm and weakening of the adhesion and invasion ability of cells. Moreover, the YehE protein also has an interactive relationship with the fiber chaperone molecule YehC with 40% confidence. The *yeh* operons (including *yehA, yehB, yehC, yehD*) encode cryptic but functional fimbriae adhesins which observed that deletion of the *yeh* chaperone-usher fimbriae operons did not result in any major effect on *E. coli* capacity to form biofilm compared with the WT *E. coli* MG1655 strain. The poorly expressed *yeh* putative fimbriae operons don't contribute to *E. coli* adhesion under classical laboratory conditions (Korea et al., 2010). This may partially explain why those potential fimbriae were not previously identified as adhesion factors. Therefore, it is possible that the effect of protein YehE on biofilm and adhesion is mediated by type 1 fimbriae.

The flagellum is the moving organ of bacteria, which plays an important role in the biofilm of gram-negative bacteria, and the flagella-mediated movement on biofilm formation in *E. coli* has been studied, and it is mainly involved in contact and adhesion with solid surface (Wood et al., 2006). As the important structure of flagella T3SS, three genes (*fliS*, *fliD*, *fliR*) are necessary for flagellin output and flagellum-mediated movement (Minamino, 2014). Curli fimbriae includs 6 proteins CsgA, CsgB, CsgC, CsgE, CsgF, CsgG (Robinson et al., 2006). Curli fimbriae play an important role in the early stage of biofilm formation, also attests that curli participates in cell-cell adhesion and cell-solid surface adhesion in *E. coli*, and promotes biofilm formation. C-di-GMP is a major regulator of bacterial behavior and physiology affecting motility and biofilm formation in bacteria (Castiblanco and Sundin, 2018). The c-di-GMP signaling major target is the transcription regulator *csgD*, which can activate the expression of amyloid curli fimbriae and the exopolysaccharide cellulose, and thus participate in biofilm formation (Li et al., 2019). Therefore, both the flagella system and the fimbriae system are very important for the adhesion and contact with non-biological surfaces, which may be the reason for down-regulation of biofilm formation. Hence, the results not only enhanced the importance of these genes in the biofilm regulatory network, but also confirms the accuracy of our screening of biofilm-related genes.

The identified biofilm-related genes have sites related to metabolism. *E. coli* produces ATP through anaerobic glycolysis of sugar and tricarboxylic acid cycle (TCA cycle) to provide energy for life activities,

such as flagella movement and extracellular protein secretion (Folsom et al., 2014). In *Staphylococcus*, the presence of ATP synthase aureus is necessary for biofilm (Bosch et al., 2020). In *S.epidermidis*, the exopolysaccharide facilitating bacterial adherence in a biofilm is polysaccharide intercellular adhesin (PIA), environmental changes that inhibited TCA-cycle activity also resulted in a massive derepression of PIA biosynthetic genes and increased PIA production (Sadykov et al., 2008). The *gltA* gene encodes citrate synthase, as the first key enzyme in the TCA cycle. Citrate synthase has been considered to be an important control point for the determination of cellular metabolic rate, the *gltA* gene affects the TCA cycle and thus affects the ability of biofilm formation (De Backer et al., 2018). The periplasm is unique to gram-negative bacteria, it's a complex structure located between the outer and inner membranes, and plays an important role in regulating the entry and exit of substances (Scheurwater and Clarke, 2008). The periplasm contains large amounts of glucose-derived polymers, and these osmotic regulating sugars have been found to play an important role in the biofilm formation of *E. coli* and *Pseudomonas aeruginosa* (Mahasenan et al., 2020; Malinverni and Silhavy, 2009). Two coenzymes membrane-bound lytic transglycosylase F (*mltF*) and beta-D-glucoside glucohydrolase (*bglX*) were identified, which related to glucose derivatives in the periplasm. In *Pseudomonas aeruginosa*, the substrate pattern of BglX may be consistent with PSL exopolysaccharides, and both are related to the biofilm formation, and the inactivation of *bglX* reduces the biofilm formation (Mahasenan et al., 2020). The MltF protein as a newly named glycosyltransferase, had a similar effect to glycosyl hydrolase and was found to be homologous to the periplasmic substrate-binding protein of the ABC transporter in *E. coli* (Madoori and Thunnissen, 2010). In *Pseudomonas aeruginosa*, the inactivation of the ABC transporter substrate *egg* gene *dppA1* reduces biofilm formation and delays cell aggregation (Lee et al., 2018). Although these genes have not been reported to be associated with biofilm formation in *E. coli*, it is not difficult to find that these proteins are involved in the exchange of substances inside and outside the cell membrane, which demonstrates the importance of membrane surface substances for biofilm formation.

Among the biofilm-related genes identified in this study, except for the 10 flagellar system and fimbriae system genes that have been studied, those other biofilm-related genes have not been reported in *E. coli*. Therefore, these genes can be called the novel gene associated with biofilm. This study also had important implications for the formation and elimination of biofilm in *E. coli.* However, we have not completely screened out the systems related to biofilm formation in this study. For example, the cyclic di-GMP and quorum sensing genes have not been screened, although they play an important role in biofilm formation (L. Wang et al., 2005; Pruss et al., 2006). Moreover, the biofilm formation-related genes were screened by 96-well polyethylene plate in this study. The biofilm mechanism of *E. coli* on the surface of non-biological materials is determined by different genes, which may result in different gene was screened by different screening methods. For instance, different temperatures and different times may cause different genes to be screened. Our screening strategy was to screen strains with weakened biofilm, but we also screen strains with enhanced biofilm in our study. This may be the direction of our next research.

Declarations

Acknowledgment This work was supported by the fund of the National Natural Science Foundation of China (31872483, 31772707 and 32072829), the Shanghai Natural Science Foundation of China (No.22ZR1475800), Natural Science Foundation of Fujian Province of China (No. 2021J011099) and the Research Foundation for Advanced Talents of Longyan University (Grant No. 2021ZN001).

Author Contributions JH, YG, KZ, and XN conceived and designed the study. Material preparation, data collection and analysis were performed by JH, YG, HL, FY, XN, MAR, JZ, HY, CH. XS and JT provided the suggestion of the research work. The manuscript was drafted and revised for important intellectual content by JH, YG, WZ, WJ, ZC, KZ and XH and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability All data generated or analyzed during this study are included in this manuscript.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This study does not contain any experiments with human participants or animals.

References

1. Abraham JM, Freitag CS, Clements JR, Eisenstein BI (1985) An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc Natl Acad Sci U S A 82(17):5724–5727. doi:10.1073/pnas.82.17.5724
2. Blomfield IC (2001) The regulation of pap and type 1 fimbriation in *Escherichia coli*. Adv Microb Physiol 45:1–49. doi:10.1016/s0065-2911(01)45001-6
3. Bosch ME, Bertrand BP, Heim CE, Alqarzaee AA, Kielian T (2020) *Staphylococcus aureus* ATP Synthase Promotes Biofilm Persistence by Influencing Innate Immunity. mBio 11(5)
4. Castiblanco LF, Sundin GW (2018) Cellulose production, activated by cyclic di-GMP through BcsA and BcsZ, is a virulence factor and an essential determinant of the three-dimensional architectures of biofilms formed by *Erwinia amylovora* Ea1189. Mol Plant Pathol 19(1):90–103. doi:10.1111/mpp.12501
5. Crepin S, Porcheron G, Houle S, Harel J, Dozois CM (2017) Altered Regulation of the Diguanylate Cyclase YaiC Reduces Production of Type 1 Fimbriae in a Pst Mutant of Uropathogenic *Escherichia coli* CFT073. J Bacteriol 199(24). doi:10.1128/JB.00168-17
6. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640–6645. doi:10.1073/pnas.120163297
7. De Backer S, Sabirova J, De Pauw I, De Greve H, Hernalsteens JP, Goossens H et al (2018) Enzymes Catalyzing the TCA- and Urea Cycle Influence the Matrix Composition of Biofilms Formed by Methicillin-Resistant *Staphylococcus aureus* USA300. Microorganisms 6(4). doi:10.3390/microorganisms6040113

8. Filloux A, Vallet I (2003) [Biofilm: set-up and organization of a bacterial community]. Med Sci (Paris) 19(1):77–83. doi:10.1051/medsci/200319177
9. Folsom JP, Parker AE, Carlson RP (2014) Physiological and Proteomic Analysis of *Escherichia coli* Iron-Limited Chemostat Growth. J Bacteriol 196(15):2748–2761
10. Gally DL, Leathart J, Blomfield IC (1996) Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. Mol Microbiol 21(4):725–738. doi:10.1046/j.1365-2958.1996.311388.x
11. Hu J, Zuo J, Chen Z, Fu L, Lv X, Hu S et al (2019) Use of a modified bacterial ghost lysis system for the construction of an inactivated avian pathogenic *Escherichia coli* vaccine candidate. Vet Microbiol 229:48–58. doi:10.1016/j.vetmic.2018.12.020
12. Ito A, Taniuchi A, May T, Kawata K, Okabe S (2009) Increased antibiotic resistance of *Escherichia coli* in mature biofilms. Appl Environ Microbiol 75(12):4093–4100. doi:10.1128/AEM.02949-08
13. Klemm P (1986) Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in *Escherichia coli*. EMBO J 5(6):1389–1393. http://www.ncbi.nlm.nih.gov/pubmed/2874022
14. Klemm P, Christiansen G (1990) The fimD gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. Mol Gen Genet Mgg 220(2):334–338
15. Korea CG, Badouraly R, Prevost MC, Ghigo JM, Beloin C (2010) *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. Environ Microbiol 12(7): 1957-77. doi:10.1111/j.1462-2920.2010.02202.x
16. Lee Y, Song S, Sheng L, Zhu L, Kim JS, Wood TK (2018) Substrate Binding Protein DppA1 of ABC Transporter DppBCDF Increases Biofilm Formation in *Pseudomonas aeruginosa* by Inhibiting Pf5 Prophage Lysis. Front Microbiol 9:30. doi:10.3389/fmicb.2018.00030
17. Li F, Cimdins A, Rohde M, Jansch L, Kaever V, Nimtz M et al (2019) DncV Synthesizes Cyclic GMP-AMP and Regulates Biofilm Formation and Motility in *Escherichia coli* ECOR31. mBio 10(2). doi:10.1128/mBio.02492-18
18. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402–408. doi:10.1006/meth.2001.1262
19. Liyun W, Robert K, Qi Z, A, W. J., E, B. C., L, R. A., et al (2018) Influence of Type I Fimbriae and Fluid Shear Stress on Bacterial Behavior and Multicellular Architecture of Early *Escherichia coli* Biofilms at Single-Cell Resolution. Applied and environmental microbiology 84(6)
20. Madoori PK, Thunnissen AM (2010) Purification, crystallization and preliminary X-ray diffraction analysis of the lytic transglycosylase MltF from *Escherichia coli*. Acta Crystallogr Sect F Struct Biol Cryst Commun 66(Pt 5):534–538. doi:10.1107/S1744309110010596
21. Mahasenan KV, Batuecas MT, De Benedetti S, Kim C, Rana N, Lee M et al (2020) Catalytic Cycle of Glycoside Hydrolase BglX from *Pseudomonas aeruginosa* and Its Implications for Biofilm Formation. ACS Chem Biol 15(1):189–196. doi:10.1021/acschembio.9b00754

22. Malinverni JC, Silhavy TJ (2009) An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. Proc Natl Acad Sci U S A 106(19):8009–8014. doi:10.1073/pnas.0903229106
23. Minamino T (2014) Protein export through the bacterial flagellar type III export pathway. Biochim Biophys Acta 1843(8):1642–1648. doi:10.1016/j.bbamcr.2013.09.005
24. Mittal S, Sharma M, Chaudhary U (2015) Biofilm and multidrug resistance in uropathogenic *Escherichia coli*. Pathog Glob Health 109(1):26–29. doi:10.1179/2047773215Y.0000000001
25. Pang M, Xie X, Dong Y, Du H, Wang N, Lu C et al (2017) Identification of novel virulence-related genes in Aeromonas hydrophila by screening transposon mutants in a Tetrahymena infection model. Vet Microbiol 199:36–46. doi:10.1016/j.vetmic.2016.12.021
26. Pruss BM, Besemann C, Denton A, Wolfe AJ (2006) A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. J Bacteriol 188(11):3731–3739. doi:10.1128/JB.01780-05
27. Robinson LS, Ashman EM, Hultgren SJ, Chapman MR (2006) Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol 59(3):870–881. doi:10.1111/j.1365-2958.2005.04997.x
28. Roy R, Tiwari M, Donelli G, Tiwari V (2018) Strategies for combating bacterial biofilms: A focus on anti-biofilm agents and their mechanisms of action. Virulence 9(1):522–554. doi:10.1080/21505594.2017.1313372
29. Sadykov MR, Olson ME, Halouska S, Zhu Y, Fey PD, Powers R et al (2008) Tricarboxylic acid cycle-dependent regulation of *Staphylococcus epidermidis* polysaccharide intercellular adhesin synthesis. J Bacteriol 190(23):7621–7632. doi:10.1128/JB.00806-08
30. Scheurwater EM, Clarke AJ (2008) The C-terminal domain of *Escherichia coli* YfhD functions as a lytic transglycosylase. J Biol Chem 283(13):8363–8373. doi:10.1074/jbc.M710135200
31. Wang L, Li J, March JC, Valdes JJ, Bentley WE (2005) luxS-dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling. J Bacteriol 187(24): 8350-60. doi:10.1128/JB.187.24.8350-8360.2005
32. Wang X, Ding C, Wang S, Han X, Yu S (2015) Whole-Genome Sequence Analysis and Genome-Wide Virulence Gene Identification of *Riemerella anatipestifer* Strain Yb2. Appl Environ Microbiol 81(15):5093–5102. doi:10.1128/AEM.00828-15
33. Wood TK, Barrios G, Herzberg AF, Lee J (2006) Motility influences biofilm architecture in *Escherichia coli*. Appl Microbiol Biotechnol 72(2):361–367. doi:10.1007/s00253-005-0263-8
34. Wu X, Lv X, Lu J, Yu S, Jin Y, Hu J et al (2017) The role of the ptsI gene on AI-2 internalization and pathogenesis of avian pathogenic *Escherichia coli*. Microb Pathog 113:321–329. doi:10.1016/j.micpath.2017.10.048
35. Xue M, Xiao Y, Fu D, Raheem MA, Shao Y, Song X et al (2020) Transcriptional Regulator YqeI, Locating at ETT2 Locus, Affects the Pathogenicity of Avian Pathogenic *Escherichia coli*. Animals (Basel) 10(9). doi:10.3390/ani10091658

36. Zuo J, Tu C, Wang Y, Qi K, Hu J, Wang Z et al (2019a) The role of the wzy gene in lipopolysaccharide biosynthesis and pathogenesis of avian pathogenic *Escherichia coli*. Microb Pathog 127:296–303. doi:10.1016/j.micpath.2018.12.021
37. Zuo J, Yin H, Hu J, Miao J, Chen Z, Qi K et al (2019b) Lsr operon is associated with AI-2 transfer and pathogenicity in avian pathogenic *Escherichia coli*. Vet Res 50(1):109. doi:10.1186/s13567-019-0725-0

Figures

Figure 1

Analysis of biofilm formation ability by crystal violet staining. It was significantly decreased in the mutant strains, as compared to that of the strains APEC81 ($P < 0.05$). All assays were performed in triplicate and repeated three times. Differences between mean values were assessed by an unpaired two-tailed Student's t test (ns, not significant; *, $P < 0.05$ represents a significant decrease in biofilm formation ability compared with APEC81).

Figure 2

Classification of the Tn5 random insertion mutants based on biofilm formation ability. Group 1 contains 12 strains, accounting for 0.8%, the biofilm formation ability of mutants were increased compared with the strain APEC81. Group 2 contains 1460 strains, accounting for 97.3%, there was no significant difference between the biofilm and the strain APEC81. Group 3 contains 28 strains, accounting for 1.9%, the mutants significantly reduced the biofilm formation ability compared with the strain APEC81.

Figure 3

The insertion sites were shown on the APEC81 genome (the sequence moves clockwise). The analysis of insertion sites showed that the insertion hotspots affecting biofilm were located in the type 1 fimbriae operon, flagella operon, and curli fimbriae operon (red circle). From the outermost track to the innermost track, the gene positions and sizes are respectively (red arrow indicates the sequence direction), and GC skew (green represents GC skew+, purple represents GC skew-), GC content (yellow).

Figure 4

Biofilm formation of deletion strains was determined by crystal violet staining. Biofilm formation on the 96-well plate polystyrene surface after 48 h of incubation at 25°C under steady-state conditions in LB medium. Bars represent the mean values, with error bars representing SD. Differences between mean values were assessed by an unpaired two-tailed Student's t test (ns, not significant; *, $P < 0.05$ represents a significant decrease in biofilm formation ability compared with APEC81).

Figure 5

Colony morphology of the deletion strain on the congo red coomassie brilliant blue plate. The colony sequence: a. APEC81, b. 81△*fliS*, c. 81△*fliD*, d. 81△*fliR*, e. 81△*csgA*, f. 81△*csgF*, g. 81△*csgG*, h. 81△*csgD*, i. 81△*fimA*, j. 81△*fimD*, k. 81△*fimC*, l. 81△*glxK*, m. 81△*bglX*, n. 81△*mltF*, o. 81△*07045*, p. 81△*14275*, q. 81△*11735*, r. 81△*11255*.

Figure 6

Protein prediction results. (a) The prediction result of the protein domain encoded by yehE. The image is colored in rainbow color order from the N-terminal to the C-terminal**.** (b) Prediction of related pathways between the protein encoded by gene *yehE* and other proteins. The yellow line represents text mining, the purple line represents experimentally determined, the green line represents gene neighborhood.

Figure 7

Deletion of yehE decreased the adherent and invasive abilities of APEC. (a) Adherence of strain APEC81, APEC81Δ*yehE* and cAPEC81Δ*yehE* to Hela cells. Adherence to Hela cells by APEC81Δ*yehE* was significantly reduced compared with APEC81 ($P < 0.05$). (b) The invasion of HD11 cells by APEC81Δ*yehE* was also significantly decreased ($P < 0.01$). The columns represent the means and standard deviations of three experiments. *P* values less than 0.05 was considered significant.

Figure 8

The effect of hypothesized protein gene yehE on transcriptional profiles of fimA, fimB and fimE. Compared to that for the APEC81 strain, the mRNA levels of two genes (*fimA and fimB*) decreased in the *yehE* mutant strain. The mRNA levels of *fimE* increased in the *yehE* mutant strain. Average (fold change, normalized to WT) for each transcript (X-axis) are plotted using data from at least three independent experiments.

Supplementary Files

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

- Supplementalmaterial.pdf