



Genetic diversity and features analysis of type VI secretion systems loci in avian pathogenic *Escherichia coli* by wide genomic scanning



Jiale Ma, Min Sun, Yinli Bao, ZiHao Pan, Wei Zhang, Chengping Lu, Huochun Yao*

Key Lab of Animal Bacteriology, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, China

ARTICLE INFO

Article history:

Received 6 August 2013

Received in revised form 24 September 2013

Accepted 30 September 2013

Available online 8 October 2013

Keywords:

Avian pathogenic *Escherichia coli*

Type VI secretion system

Prevalence

Diversity

Genomic

VgrG protein

ABSTRACT

Avian pathogenic *Escherichia coli* (APEC) strains frequently cause extra-intestinal infections and significant economic losses. Recent studies revealed that the type VI secretion system (T6SS) is involved in APEC pathogenesis. Here we provide the first evidence of three distinguishable and conserved T6SS loci in APEC genomes. In addition, we present the prevalence and comparative genomic analysis of these three T6SS loci in 472 APEC isolates. The prevalence of T6SS1, T6SS2 and T6SS3 loci were 14.62% (69/472), 2.33% (11/472) and 0.85% (4/472) positive in the APEC collections, respectively, and revealed that >85% of the strains contained T6SS loci which consisted of the virulent phylogenetic groups D and B2. Comprehensive analysis showed prominent characteristics of T6SS1 locus, including wildly prevalence, rich sequence diversity, versatile VgrG islands and excellent expression competence in various *E. coli* pathotypes. Whereas the T6SS2 locus infatuated with ECOR group B2 and sequence conservation, of which are only expressed in meningitis *E. coli*. Regrettably, the T6SS3 locus was encoded in negligible APEC isolates and lacked several key genes. An in-depth analysis about VgrG proteins indicated that their COG4253 and gp27 domain were involved in the transport of putative effector islands and recognition of host cells respectively, which revealed that VgrG proteins played an important role in functions formation of T6SS.

© 2013 The Authors. Published by Elsevier B.V. Open access under CC BY license.

1. Introduction

Avian pathogenic *Escherichia coli* (APEC) is an important member of the extra-intestinal pathogenic *E. coli* (ExPEC), and systemic infections caused by APEC are economically devastating to poultry industries (Dho-Moulin and Fairbrother, 1999; Ewers et al., 2003). APEC enters and colonizes the avian respiratory tract, leads to localized infections, such as airsacculitis and pneumonia. In certain cases, it leads to acute septicemia, which commonly results in sudden death (Dho-Moulin and Fairbrother, 1999; Ewers et al., 2003; Rodriguez-Siek et al., 2005). APEC strains are closely related to human ExPEC strains. Therefore, some closely related clones could be involved in extra-intestinal infections in both humans and poultry, suggesting that these isolates are not host specific (Rodriguez-Siek et al., 2005; Moulin-Schouleur et al., 2006, 2007; Johnson et al., 2007; Mora et al., 2009).

Recent researches have described a new secretion system, called the type VI secretion system (T6SS), in several bacterial spe-

cies including APEC, representing a new paradigm in protein secretions (Filloux et al., 2008). de Pace et al. (2010) found that the mutants of T6SS core genes (ClpV and Hcp) of APEC strain SEPT362 displayed decreased adherence and actin rearrangement on epithelial cells. A similar result was observed from the Hcp mutant of the neonatal meningitis *E. coli* (NMEC) strain RS218 (Zhou et al., 2012). Therefore, the role that T6SS plays in contributing to the virulence of APEC warrants further investigations.

Some studies have revealed that T6SS gene clusters are assembled from at least 13 core proteins, called 'core components' (Parsons and Heffron, 2005; Bonemann et al., 2009), which comprise the minimal machinery necessary for the functionality of T6SS. However, most T6SS gene clusters actually encode additional proteins, the function of which remains unknown. VgrG protein, one of the core components performing diverse functions, has been demonstrated to be secreted by the T6SS. The N-terminal of VgrG may serve in assembly of the T6SS machinery, while the C-terminal portion functions as effector (Pukatzki et al., 2007; Ma et al., 2009). Former researches have reported the function of evolved C-terminal portion that carries unknown proteins encoded in the VgrG islands in different bacterial species (De Maayer et al., 2011; Sarris et al., 2011). The N-terminal of VgrG protein contains a gp27 domain and a gp5 domain, which leads to the VgrG trimeric structure organized in a similar way as the (gp27)₃–(gp5)₃ complex does of bacteriophage T4 (Leiman et al., 2009). The lysozyme activity of

* Corresponding author. Tel./fax: +86 25 84395328.

E-mail address: yaohch@njau.edu.cn (H. Yao).

bacteriophage gp5 protein is capable of penetrating the outer cell membrane and locally dissolving the periplasmic cell wall (Arisaka et al., 2003). The gp27 protein acts as tail fiber protein in bacteriophage which confers a high degree of host specificity by binding to the specific receptor of sensitive bacteria (Zhang et al., 2009; Xu et al., 2013).

In this study we present the results in silico analyses of putative T6SS-related genes found in the genome of APEC strain ED205. According to the result of prevalence analysis in 472 APEC isolates, all identified T6SS loci were sequenced. We aimed to perform an extensive comparative genomic analysis among abundant APEC isolates to further explore the properties characteristics of T6SS loci in APEC.

2. Materials and methods

2.1. Bacterial strains

A total of 472 APEC strains were isolated from the brains of ducks with septicaemic and neurological symptoms as previously described (Wang et al., 2011) (Table S3). The identity of each strain was confirmed as *E. coli* by using the VITEK 2 system (BioMérieux Vitek, Inc., Hazelwood, MO, USA). Some APEC isolates used as important representative strains in this study are marked in red in Fig. 3. The APEC ED205 strain was chosen for genome sequencing.

In addition, the *E. coli* reference strains (EHEC O157, ETEC PE114, EPEC PE059, UPEC 536, NMEC RS218 and APEC_O1) used in this study were purchased from the Control Institute of Veterinary Bioproducts and Pharmaceuticals of China, and identified to encode the T6SS1 and T6SS2 loci. All *E. coli* strains were grown in Luria–Bertani (LB) medium at 37 °C with aeration.

2.2. In silico identification of the T6SS loci in APEC

APEC ED205 strain genome sequencing was performed by using a Roche 454 genome sequencer FLX system. The draft genome sequence was uploaded to the bioinformatics analysis platform (<http://www.chinasscontrol.com/biosystem/index.php>) of Nanjing Agricultural University. Nucleotide sequences from enteroaggregative *E. coli* Sci-I and Sci-III 042 strain and NMEC T6SS locus RS218 strain were used for BLASTN and reverse BLAST against the ED205 draft genome in the bioinformatics analysis platform. Clusters containing at least five genes encoding proteins with similarity to known T6SS core proteins from NMEC or EAEC were considered as part of a putative T6SS locus in APEC ED205 genomes. These putative genomic regions were then extended by examining 10 kb up-stream and down-stream for putative conserved genes associated with T6SS by reverse blast analysis against the EAEC 042 genome.

2.3. Sequencing of T6SS loci

T6SS loci were detected by multiplex PCR using specific oligonucleotide primers for amplification of 3 conserved T6SS primary core gene: IcmF (membrane-bound protein), ClpV (ATP bind protein) and VgrG (spike protein). T6SS loci were defined by the presence of at least two of the three genes, and used primers were shown in Supplementary Table S1. All of the APEC T6SS loci detected in this study were amplified and sequenced, being performed by following the structure feature of gene clusters to reduce the workload. This process was carried out by Shanghai Sunny Biotechnologe Co., Ltd., and each cluster was assembled as one containing the entire T6SS locus. Phylogenetic groups were determined for the 472 APEC isolates by using the triplex polymer-

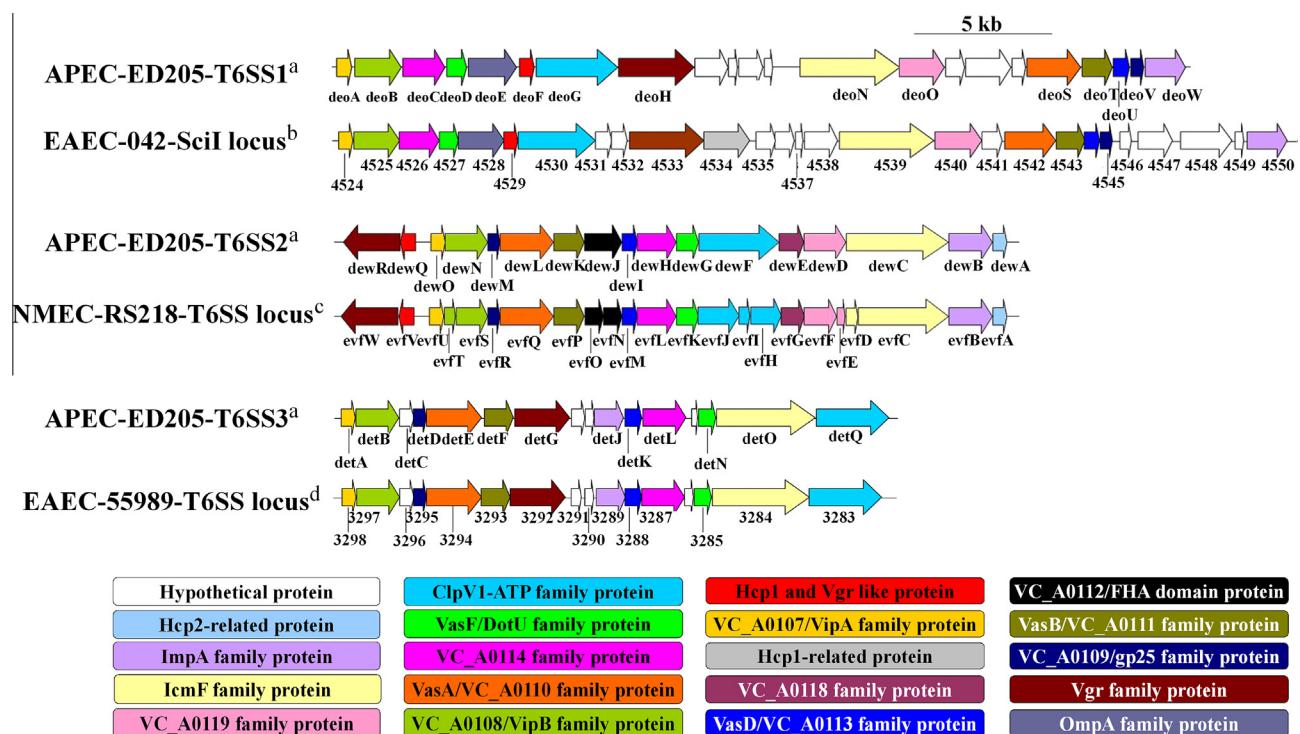


Fig. 1. Comparative genome alignments of APEC ED205 and reference strains. Genes encoding conserved domain proteins were represented by the same colors. White arrows indicate other genes encoded in the T6SS loci which were not identified as part of the conserved core described by Boyer et al. (2009). The direction of the arrows indicates the direction of transcription. The color keys for the functional classes of genes in the T6SS loci are shown at the bottom. (a) The ED205 draft genome sequenced by our lab. (b) The NCBI reference number of EAEC-042 genome is NC_017626. (c) The NCBI reference number of NMEC-RS218 T6SS cluster is JN837480. (d) The NCBI reference number of EAEC-55989 genome is NC_011748.

ase chain reaction (PCR) as previously described (Clermont et al., 2000).

2.4. Phylogenetic analyses

Phylogenetic analyses were performed following the procedures outlined by Bingle et al. (2008). A ClustalW alignment with default parameters was used with the entire T6SS loci nucleic acid sequences and VgrG amino acid sequences. Then, the T6SS clusters phylogenetic tree was constructed with the MEGA (v.5.0.3) software package using the neighbor-joining method, with *P*-distance, complete gap deletion and bootstrapping (*N* = 1000) parameters. The similar approach was used to construct the phylogenetic tree of VgrG proteins, which only differs in using the Poisson correction instead of *P*-distance.

2.5. Sequence annotation and GC content analysis

T6SS-related genes were named according to the T6SS genes of *Vibrio cholerae* (Pukatzki et al., 2006). Maps of the genomic islands were constructed manually in the VECTOR NTI program and Microsoft PowerPoint. Visual representation of the alignments using nucleotide similarities (tblastx) of the T6SS loci and VgrG genes were carried out with the Artemis Comparison Tool (ACT) (Carver et al., 2005). The nucleic acid or translated proteins were compared with those in GenBank database by the BLAST network service. The average GC content of the whole T6SS loci and the VgrG islands (see below) were determined using the Bioedit (v.7.0.5.3) package.

2.6. Analyses of the VgrG islands

VgrG islands were defined according to procedures outlined by De Maayer et al. (2011). In this study, only the APEC-T6SS1 locus encoded visible VgrG islands. The amino acid sequences for the proteins encoded in the VgrG islands were analyzed for sequence identity by Blast *P* analysis against the NCBI protein database. The presence of conserved domains was identified by Blast analysis against the Conserved Domain Database (CDSearch). In addition, the GC contents of the VgrG genes were determined for the conserved N-terminal region, including the conserved Vgr and Gp5 domains, as well as the C-terminal extensions which were all considered to be nucleotides located at the 3' end of the Gp5 domain. All of the proteins containing conserved domains were used to analyze possible functions.

2.7. Proteins structure prediction and modeling

The amino-terminal and central domains of VgrG proteins resemble the T4 bacteriophage gp27 protein and the C-terminal domain of gp5, respectively (Pukatzki et al., 2007). In this study, the structure comparison between gp27 domain of VgrG and gp27 protein of T4 bacteriophage (gp27-T4) was performed. All of the models were generated by employing the SWISS-MODEL server (<http://swissmodel.expasy.org>) (Arnold et al., 2006). The model of gp27-T4 protein was released by Kanamaru et al. (2002). The model of gp27-T6SS1 was drawn based on template 2p5z.X (VgrG protein from *E. coli* CFT073 c3393), sharing 89.84% sequence identity. The model of gp27-T6SS2 was drawn based on template 1wru.A (tail protein from bacteriophage Mu gp44) (Kondou et al., 2005), sharing 31.48% sequence identity and the modeled domain ranged from residues 8–358 aa. As no a highly similar template could be referenced, the gp27-T6SS3 could not be drawn a high-quality model. However, a referential model was still drawn based on 2p5z.X, sharing only 14.86% sequence identity, while the modeled domain ranged from residues start to end.

2.8. Analysis of VgrGs expression

To analyze the expression competence of VgrG proteins in different *E. coli* pathotypes, an indirect ELISA assay was carried out using antibody against VgrG protein. Briefly, wells of an ELISA plate were coated using 100 μ l bacterial ($\approx 10^9$ cfu/ml) supersonic lysate at 4 °C for 12 h. Then, VgrG-ELISA was performed as reported previously (Dreyfus et al., 2004), and each sample was measured for three times. A sample A450 value/negative control A450 value (S/N) >2 was used as the positive standard. 15% mice macrophage lysate was added in Trypticase Soy Broth (Difco Laboratories, Detroit, MI) for bacterial strains culture, which could help to activate the T6SS vitality (Ma et al., 2009). Because the T6SS3 locus encoded an incomplete T6SS cluster and was not encoded in most of *E. coli* strains, only VgrG-T6SS1 and VgrG-T6SS2 were analyzed by ELISA assays. The VgrG-T6SS1 and VgrG-T6SS2 monoclonal antibodies used in this study were prepared by GenScript USA Inc. as the service SC1040 (Kohler and Milstein, 1975).

2.9. Statistical analysis

The data were analyzed by using SPSS version 17.0 (SPSS Inc., Chicago, IL). The difference between mean values among groups was evaluated, first by one-way analysis of variance (ANOVA) and then by pairwise comparison of the mean values between the two groups, followed by Tukey's student rank test. Differences in a *P* value of <0.05 were considered significant, while a *P* value of <0.01 was considered greatly significant.

3. Results and discussion

3.1. Identification of orthologous T6SS loci occur in APEC

By baiting with conserved proteins from the characterized T6SS loci in human pathogenic *E. coli*, three distinguishable T6SS loci were identified in the draft genome sequences of APEC ED205 strains (Fig. 1). The embedded orders of conserved T6SS core genes in T6SS1, T6SS2 and T6SS3 loci were fully consistent with the Scil locus of EAEC 042 strain, the T6SS locus of NMEC RS218 strain and the T6SS locus of EAEC 55989 (Fig. 1) respectively. The prevalence of T6SS1, T6SS2 and T6SS3 loci were 14.62% (69/472), 2.33% (11/472) and 0.85% (4/472) positive in the APEC collections, respectively, and revealed that >85% of the strains contained T6SS loci which consisted of the virulent phylogenetic groups D and B2 (Fig. 2). 11 APEC isolates contained within the T6SS2 locus also encoded the T6SS1 locus, and only ED205 strains contained all of three T6SSs. The phylogenetic tree of whole T6SS clusters displayed the sequence identity of three branches less than 15% (Fig. 3), which indicated that these clusters were three distinguishable genetic T6SS loci again (see Table 1).

The T6SS1 loci of APEC isolates were between 27.1 and 33.6 kb in size and encompass 18–25 protein coding sequences. Fifteen of these belonged to conserved T6SS proteins outlined by Boyer et al. (2009) (Figs. 1 and 3B). Almost all of the T6SS2 loci, consisted of only seventeen T6SS core genes, ranged a size of around 23 kb, of which the cluster from ED069 strain was the only exception (Fig. 3). The T6SS3 loci ranged a size of around 18.5 kb and encompassed 16 ORFs (Figs. 1 and 3B).

3.2. VgrG islands analysis in T6SS loci

The concept of VgrG islands was first proposed by De Maayer et al. (2011) and has been applied to most of the T6SS loci from a variety of Gram negative bacteria. In this study, we also made a VgrG island analysis in APEC isolates, indicating that only T6SS1

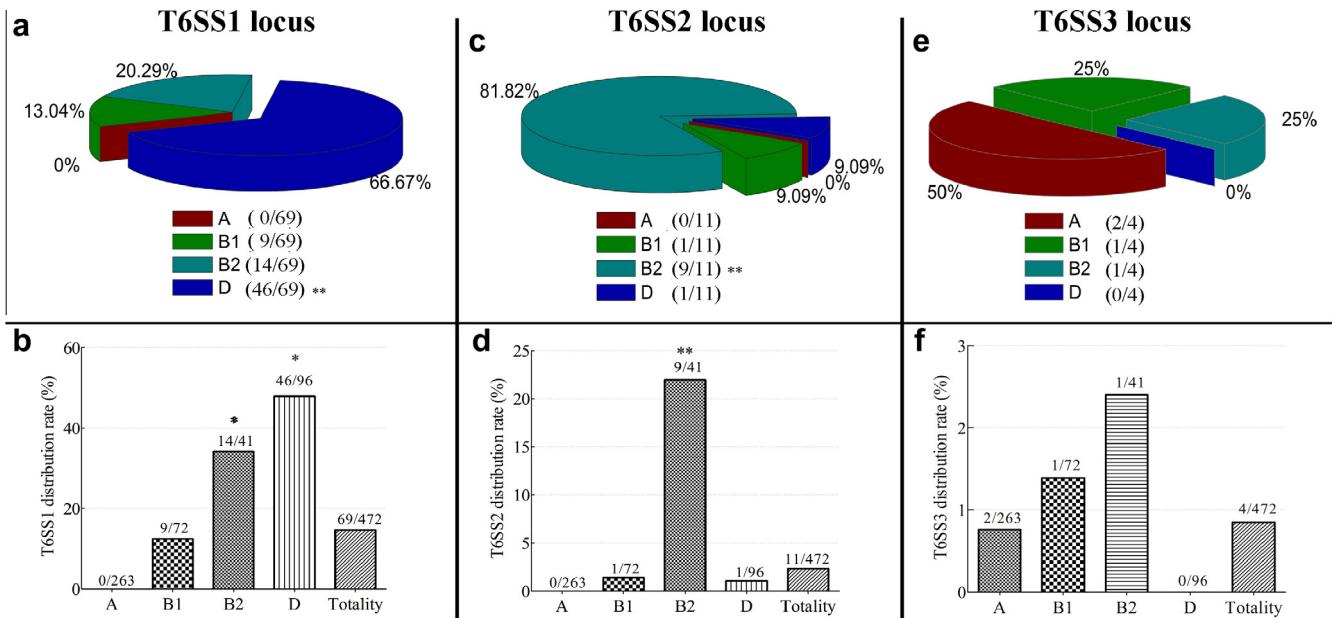


Fig. 2. The prevalence and phylogenetic groups analysis of T6SS loci among APEC isolates. (a), (c) and (e): The proportion of T6SS-positive isolates belongs to different phylogenetic groups. (b), (d) and (e): The distribution rate of three T6SS loci in different phylogenetic groups of APEC respectively. The T6SS1 is described by figures (a) and (b), the T6SS2 is described by figures (c) and (d), the T6SS3 is described by figures (e) and (f). Statistical significance was determined by a Student's t test based on comparisons with other groups (** $P < 0.01$; * $P < 0.05$).

loci contained both “COG4253” VgrGs and a visible island structure (Figs. 4 and 5). Although the T6SS1 locus shared conserved and syntenous cores in this study, considerable variability in the VgrG downstream regions was observed. Interestingly, the conserved core proteins of T6SS1 locus were arranged in two syntenous clusters (blocks I and III, see Fig. 5C), between which sandwiched a variable region (block II, see Fig. 5C), which was speculated to link with the VgrG gene encoded in the middle region of these clusters (De Maayer et al., 2011). The COG4235 is a part of the DUF 2345 superfamily, which contains VgrG-Rhs (rearrangement hot spot) proteins. The Rhs domain of the VgrGs was verified to potentially drive effector acquisition and diversity in a previous study (Wang et al., 1998).

The GC content analysis showed that the variable and unknown regions (43.97%) had a lower average value than the conserved core regions (56.9%) (Table S2). A phenomenon was noticed that the C-terminal extensions (putative effector-activity domain) of VgrGs had a similar value (48.24%) with the variable regions, while the N-terminal regions containing gp27 and gp5 domains (Cascales, 2008) (Fig. 5A) shared a approximate value (58.95%) with the conserved regions (Table S2). The presence of multiple non-homologous or highly divergent forms of unexploited genes and the C-terminal extensions of VgrGs, together with the lower GC content of these regions, supporting the view that these genes may have been imported into APEC (or their ancestors) on multiple occasions by an unknown mechanism of VgrGs.

To further prove this hypothesis, the proteins encoded in the variable VgrG islands were analyzed for its sequence similarity, structural homology and the presence of conserved domains to known proteins. The majority of proteins encoded on the islands showed homology to proteins of unknown function. However, a number of island proteins shared high sequence identity and contained conserved domains which suggested that they might represent T6SS effectors with putative functions in host-microbe and inter-bacterial interactions (Table 2). We found 12 proteins that contained conserved domains (named Pro1 to Pro12, Table 2, Fig. 5C). Each of them could be encoded in multiple APEC isolates.

3.3. The T6SS1 locus encode a multipurpose T6SS

The prevalence analysis displayed that the T6SS1 locus was more widely distributed in 472 APEC isolates (14.62%) compared with the T6SS2 and 3 (Fig. 2). Of 69 T6SS1 strains, 42 (66.67%) belonged to phylogenetic group D, while 14 (20.29%) belonged to group B2 and 9 (13.04%) belonged to group B1, but no strain belonged to group A (Fig. 2a). These data indicated that the T6SS1 locus had extensive adaptability in different background APEC strains, and also suggested the function diversity of T6SS1. Moreover, the distribution rate of T6SS1 locus was significantly higher in group D (47.92%) and B2 (34.15%), suggesting that T6SS1 was mainly encoded in virulent isolates of groups B2 and D (Johnson et al., 2006) and thus made contribution to APEC pathogenicity.

The T6SS1 locus of APEC isolates could be clearly divided into seven branches by the phylogenetic analysis (Fig. 3A), while the similar situation did not appear in the T6SS2 and 3 loci. Likewise, DNA sequence comparisons and linear interactive plots displayed that the T6SS1 locus presented great variability in both gene number and sequence polymorphism (Fig. 3B). This likely reflected the complex evolutionary mechanisms in APEC which facilitated the adaptation of T6SS1 to different hosts. A longer sequence and more unknown genes of T6SS1 locus provided the possibility to perform adjustable functions aiming different host cells.

The functional diversity can be observed more directly by analyzing of effectors encoded in VgrG islands. Based on the analysis of a substantial number of hypothetical effector proteins, we supposed that the VgrG islands of T6SS1 locus function in mediating virulence, competition proliferation and bacteria evolution. The importance of T6SS in pathogenesis is becoming increasingly obvious. In this study, a zinc-dependent metallopeptidase (Pro6), which is associated numerous times with pathogenicity in previous researches (Wassif et al., 1995; Woods et al., 2001), was found to be encoded in VgrG islands of APEC. Furthermore, the PGAP1-like protein (Pro1) found in this study contained a conserved domain of lipases, which had been shown to represent a major virulence factor in both animal and plant pathogens (Nardini et al., 2000; Ham et al., 2011). However, *E. coli* is an opportunistic path-

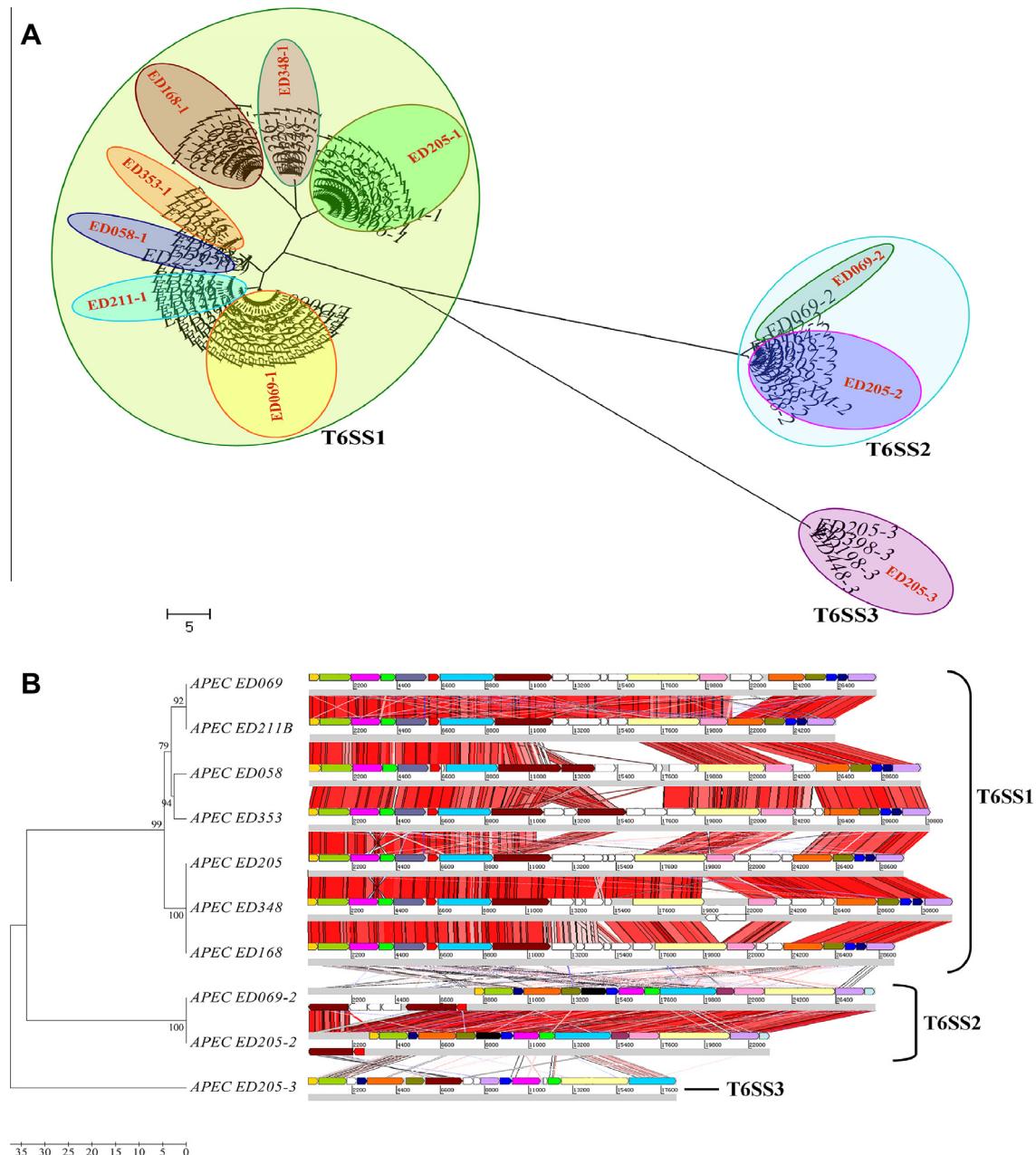


Fig. 3. The phylogenetic trees construction and comparative genomics analysis. (A) Sequence relationship of all T6SS clusters. The whole T6SS nucleotide sequences were aligned with ClustalW, and the MEGA software version 5.0 was used to perform a 2000 bootstrap phylogenetic analysis using the neighbor joining method. Large clades are indicated by colored and named by representative isolates. (B) Comparative genome alignments of 10 T6SS loci selected according to the result of Fig. 3A. The color key for the functional classes of genes and relevant information are the same as shown in Fig. 1. Phylogenetic relationships of the T6SS loci were obtained using a neighbor-joining tree (bootstrap $N = 1000$) based on a ClustalW alignment of the complete T6SS nucleotide sequences. Visual representation of the alignments using nucleotide similarities (tblastx) of the T6SS loci and were carried out with the Artemis Comparison Tool (ACT) (Carver et al., 2005).

ogen, and many *E. coli* strains with genomes encoding T6SS locus are not known to be pathogenic or even symbionts. In this study, the analysis of hypothetical effectors encoded in VgrG islands revealed that many proteins were not associated with pathogenicity, which suggested T6SS may also function in nonpathogenic bacteria-host interactions. Based on our analysis of the VgrG islands proteins contained conserved domains, the nonpathogenic function of T6SS1 locus could be summarized as competition proliferation and bacteria evolution. The PAAR-repeat domain of Pro5, 11 and 12 in VgrG islands could sharpen and diversify the T6SS spike and were required for full functionality of the T6SS in *V. cholerae* and *Acinetobacter baylyi* (Shneider et al., 2013). The presence of

bacteriocin-like proteins (Pro11) in the T6SS1 locus of these APEC isolates supported the finding of a potential function for the T6SS in antibiosis and competition (Hood et al., 2010; Russell et al., 2011). The bacteriocins could take part in killing closely related bacterial species (Lesic et al., 2009). Furthermore, the VgrG trimer showed structural resemblance with the (gp27)₃-(gp5)₃ spike complex of T4 bacteriophage, which was used as the cell-puncturing apparatus to deliver viral DNA into bacterial target cells (Cascales, 2008). Coincidentally, many VgrG island proteins of APEC in this study contained integrase (Pro10) and helicase (Pro8) core domains and performed a transposase function, and these transposase are considered to be correlated with bacteriophage. These

Table 1

The large clades of phylogenetic tree of T6SS nucleotide sequences and their representative isolate.

T6SS locus	Group	Number of orthologous locus	Representative isolates			Length of the T6SS locus ^a (bp)	Number of orfs ^b	GenBank accession No.
			Name	Serogroups	Phylogenetic groups			
T6SS1	I	3	ED058	07	B1	30,000–31,000	22	KF678356
	II	21	ED069	06	D	28,000–29,000	20–22	KF678354
	III	11	ED168	Unknown	B2	29,000–30,000	20–22	KF678357
	IV	17	ED205	02	B2	29,000–30,000	22–24	KF678351
	V	7	ED211B	075	D	26,000–27,000	17–19	KF678358
	VI	6	ED348	01	B2	32,000–33,000	23–25	KF678359
	VII	4	ED353	Unknown	B2	30,500–31,500	24	KF678360
	Total 69							
T6SS2	VIII	10	ED205	02	B2	22,000–24,000	17	KF678352
	XI	1	ED069	06	D	28,461	21	KF678355
T6SS3	Total 11							
	X	4	ED205	02	B2	18,000–19,000	16	KF678353
Total 4								

^a From the first to last conserved core gene of T6SS.^b Ignoring flanking genes.

observations suggested that T6SS1 locus might play an important role in the evolutionary mechanism of APEC, which relies on the bacteriophage to evolve.

A phylogeny based on the VgrG-T6SS1 proteins showed that all of them were from different *E. coli* pathotypes, but mixed together. It was difficult to pinpoint which one was from ExPEC or which one was from intestinal pathogenic *E. coli* (Fig. 4). Interestingly, the expression competence of each VgrG proteins was very close in levels, thus no significant difference was observed among seven pathotypes (Fig. 5B). It seemed that the T6SS1 locus could function in every *E. coli* pathotype and would play versatile roles in a variety of survival environments.

The original study of T6SS was in the context of pathogenic bacterial–host interactions, though it should not be limited in such narrow space. Several papers had highlighted the diverse potentials of T6SS functions (Jani and Cotter, 2010). The versatile T6SS may also function to promote commensal or mutualistic relationships between bacteria and eukaryotes (Hood et al., 2010), or to mediate both cooperative and competitive interactions between bacteria (Robinson et al., 2009). These latest progresses seemed to explain why the APEC T6SS1 locus had the sequence and effectors diversity, wide distribution and expression. In conclusion, we speculated that the APEC T6SS1 locus was not just for pathogenesis anymore.

3.4. The T6SS2 loci encode a single-minded T6SS

The prevalence of T6SS2 locus was very single-minded, and had an addiction to encode in phylogenetic group B2. Of 472 APEC isolates, 11 strains (2.33%) were confirmed to encode the T6SS2 locus. Among them, 9 (81.82%) belonged to group B2, and the remaining two belonged to group B1 and D respectively (Fig. 2c). Group B2 (21.95%) was the most prevalent phylogenetic group within the T6SS2 locus (Fig. 2d). These results indicated that the function of T6SS2 locus seemed to focus on an unknown aspect. In the previous study, the group B2 was proved to have a higher chance to produce virulent strains, which seemed to suggest that the T6SS2 locus make contribution to APEC virulence.

Although 11 T6SS2 loci of APEC isolates could be divided into two branches by the phylogenetic analysis (Fig. 3A), the small branch contained only one strain (ED069). T6SS2 loci of the ten APEC isolates from the big branch shared a sequence of more than 95% identity, and retained a conserved and compact gene cluster structure by observing the representative T6SS2 cluster from ED205 (Fig. 3). There were not any variable and unknown regions inserted into sequences of most T6SS2 locus. The above description

suggested that the T6SS2 might have particular function aiming specific target cells, and was critical in some special mechanisms.

The phylogenetic tree of VgrG-T6SS2 proteins showed that they branched into two distinct clades (Fig. 4), which suggested distinct evolutionary backgrounds of these paralogous proteins. All of the VgrGs from APEC T6SS2 locus were assigned to ExPEC branch, which indicated that the same function was performed in extra-intestinal survival proliferation or pathogenic mechanism. Surprisingly, their expression competence of VgrG-T6SS2 also had huge difference between ExPEC and intestinal pathogenic *E. coli*. Only the NMEC and APEC had a high expression level, while the UPEC barely reached a normal level (Fig. 5B). The same result was shown by the research of de Pace et al. (2010). According to him, the APEC strain expressed T6SS genes at significantly higher levels than EHEC strain did through a transcriptomics analysis. Coincidentally, the T6SS2 locus of NMEC had been reported to play an important role in the invasion of human brain microvascular endothelial cells (Zhou et al., 2012). These results suggested that the T6SS2 might also perform an important function in avian *E. coli* causing meningitis.

The prevalence which only appeared in group B2, conserved and compact cluster structure, “ExPEC branch” VgrG proteins and expression only in meningitis *E. coli*, all of these evidences suggested that the APEC T6SS2 locus might perform a role in the meningitis mechanisms.

3.5. The T6SS3 loci encode a unserviceable T6SS

A prevalence analysis was performed to discover the phylogenetic distribution of the T6SS3 in 472 APEC isolates. However, unlike previous surveys of T6SS1 and 2, only 4 strains (0.85%) were revealed to encode the T6SS3 locus, and they belonged to three different phylogenetic groups (A, B1, and B2) (Fig. 2). These results were neither statistically significant, nor inspirational. The amount of APEC isolates encoded T6SS3 locus was negligible, apparently suggesting that it did not contribute to the pathogenicity of APEC or perform other special missions.

The phylogeny based on the whole T6SS loci showed a greater evolutionary distance between the T6SS3 and the other two T6SS loci (Fig. 3A), indicating that it might have been acquired through horizontal gene transfer. This could also be correlated with greater diversity in GC and gene content between the T6SS3 and other two T6SS loci (Fig. 3B). The GC content of conserved core regions was significantly lower in the T6SS3 loci (38.12%) than in the T6SS1 and 2 loci (56.93%), supporting the view that the T6SS3 locus might have been imported into APEC (or their ancestors) from distant

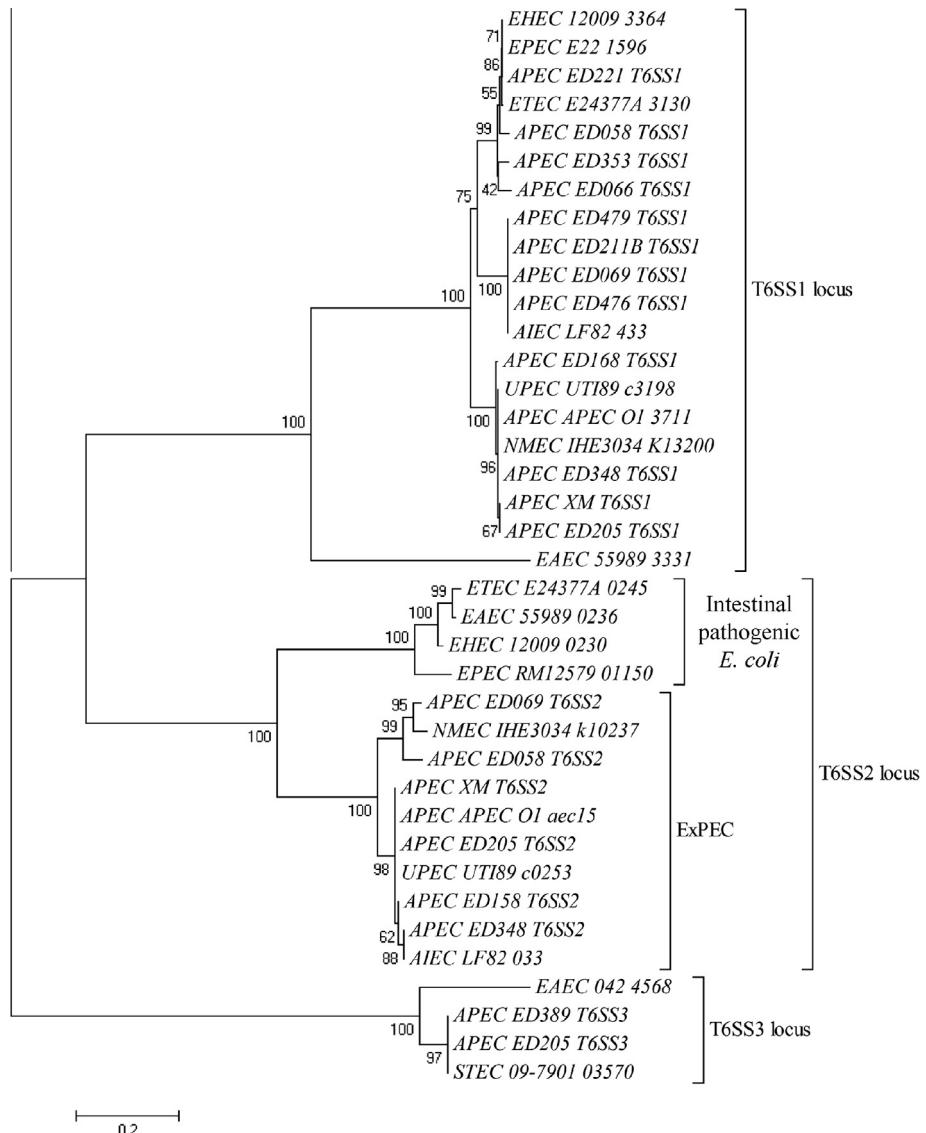


Fig. 4. Evolutionary relationships of VgrG proteins. A neighbor-joining tree (bootstrap $n = 1000$; Poisson correction) was constructed based on a ClustalW alignment of the VgrG amino acid sequences from the APEC T6SS loci and orthologous loci in other *E. coli* pathotypes.

species without correct transcription and translation guaranteed. In addition, only twelve of the T6SS conserved core proteins identified by Boyer et al. (2009) were encoded within the T6SS3 locus. Genes encoding orthologs of the Hcp-related, VC_A0112/FHA, VC_A0113, VC_A0119 and VC_A0118 family proteins are absent (Figs. 1 and 3B). Hcp-related protein is an essential component of the T6SS injection machine, and the Hcp hexamer assemble is likely the shaft of a “bolt gun” that is loaded with effectors before being fired into a target cell (Silverman et al., 2013). The phosphorylation of FHA protein can promote the clustering of ClpV, and the latter provided the force required for the assembly of the secretion machine as well as for the translocation of exported proteins (Mouggous et al., 2006). The absence of the two key proteins above indicated that the T6SS3 locus could not be assembled as a fully functional T6SS complex.

3.6. The properties difference of three T6SS loci may be supervised by the VgrG proteins

As previously described De Maayer et al. (2011) all VgrG orthologues could be divided into three different groups (“COG4253”

VgrG (Blondel et al., 2009), evolved VgrG and ordinary VgrG; Figs. 4 and 5Ab). The transportation capability of effectors genes presented a huge difference in the three VgrGs groups, which determined the high or low sequence identity of T6SS. In this study, the “COG4253” group contained all of the T6SS1 VgrG proteins, which encoded a conserved domain of unknown function (COG4253) which was absent in the T6SS2 and 3 loci VgrG proteins (Fig. 4). Interestingly, all T6SS1 clusters that encoded a “COG4253” VgrG carried more variable and unknown genes than other T6SS loci, suggesting that the VgrG island structure might be associated with the COG4253 domain. Therefore, the COG4253 conserved domain might be involved in the anchorage of the VgrG transporter to the effector proteins. These suggestions also implied that COG4253-positive and COG4253-negative T6SS VgrG proteins had different targets and functions (Blondel et al., 2009). Furthermore, some of the proteins encoded in the hypothetical VgrG islands showed sequence homology and contained conserved domains found in the C-terminal extensions of VgrG proteins. For example, the Pro-4 of some APEC isolates shared 83% amino acid identity with the C-terminal region of their VgrG-T6SS1 and encoded a COG4253 domain (Table 2). Moreover, the ED058_Vn

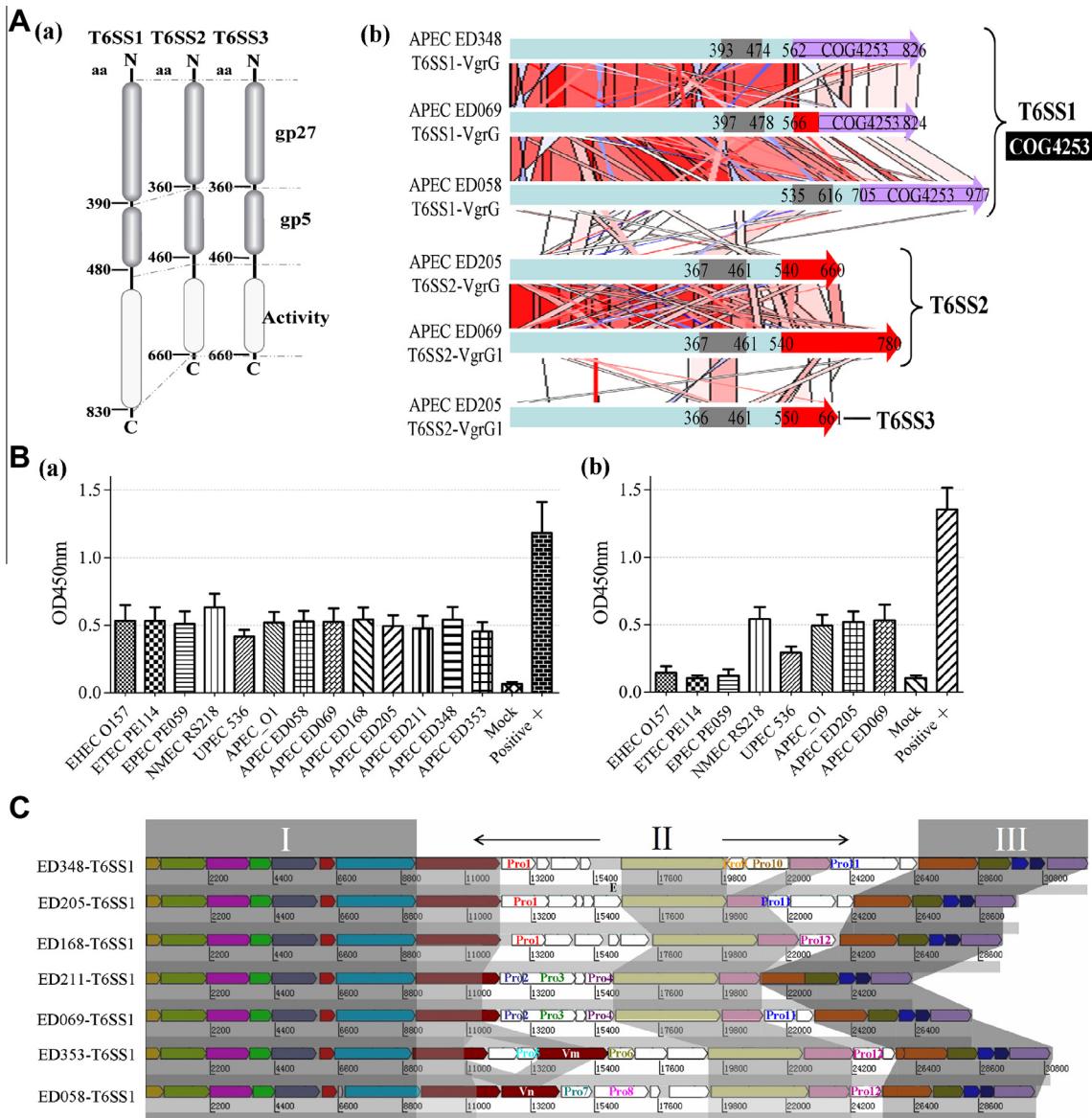


Fig. 5. The conserved domains and expression competence analysis of VgrG proteins and the hidden VgrG islands. (A) Domain organization of VgrG. (a): The conserved domains analysis of ED205 VgrG proteins, according to the domains and structures bank of NCBI. (b): Comparisons of 6 different VgrG proteins selected according to the result of Fig. 4. The conserved VgrG domains are shown in blue, Gp5 domains are shown in gray, the non-conserved C-terminal extensions are shown in red, COG4253 domains in purple. (B) ELISA analysis of the VgrGs expression. A sample A450 value/negative control A450 value (S/N)>2 was used as the positive standard. (a): The expression competence of VgrG-T6SS1 proteins in different *E. coli* pathotypes. (b): The expression competence of VgrG-T6SS2 proteins in different *E. coli* pathotypes. (C) Hypothetical VgrG islands hidden in block II of T6SS1 locus of the APEC strains. The conserved regions are shaded in grey, while non-conserved regions are not shaded. The blocks I and III sequences only encode conserved core genes and have a high identity, while the block II sequences contained several variable and unknown genes. Several color-coded genes were analyzed as the putative effectors and described in detail at the Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and ED353_Vm amino acid sequences shared 68% and 92% amino acid identity with the C-terminal region of their T6SS1 VgrG (Table 2). These above cases have been reported to present in other bacterial species, indicating that "COG4253"VgrG proteins had a stronger capability and played an important role in the formation of specific function of T6SS1. The VgrG-T6SS2 protein of APEC ED069 represented the "evolved" group, which encoded a longer variable C-terminal extensions (Fig. 5Ab) that might led to the insertion of three unknown genes and the second VgrG protein (see Fig. 3B). Almost all of the T6SS2 and T6SS3 VgrG proteins were considered to be ordinary, and kept a stable length (660 aa; Figs. 4 and 5Ab). The size of their C-terminal regions was less than 150 aa, thus was not enough to be capable of the transporting effectors genes. Therefore, both T6SS2 locus and T6SS3 locus kept high se-

quence identity and compact structure. In summary, the VgrG proteins partly determined the sequence diversity of T6SS loci and had a significant impact to the functional characteristics of T6SS among APEC isolates.

Each VgrG protein contained three parts, a gp27 domain, a gp5 domain and the variable activity domain (Cascales, 2008) (Fig. 5Aa). The similarity between gp27 domain and tail fiber protein of bacteriophage has been proved many times (Leiman et al., 2009). A recent research showed that tail fiber protein of *V. cholerae* phage VP3 could function as a receptor-binding protein and responsible for host cells recognition (Zhang et al., 2009). Therefore, we speculated that host specificity, which shared a similar mechanism with bacteriophage, was also presented in different T6SSs. In this study, the comparative analysis of gp27 domains of

Table 2

Properties of proteins in the hypothesized VgrG islands.

Number of proteins ^a	Conserved domains	Homologous function proteins (amino acid identity)	Function
Pro1	PGAP1[pfam07819]	PMI17_02526 of <i>Pantoea</i> sp. GM01 (41%)	PGAP1-like protein
Pro2	PRK01096[PRK01096]	FF52_01210 of <i>Flavobacterium</i> sp. F52 (48%)	Glycoside hydrolase
Pro3	Abhydrolase_1[pfam00561]	A225_2636 of <i>Klebsiella oxytoca</i> E718 (46%)	Transmembrane protein
Pro4	COG4253/ DUF2345[pfam10106]	VgrG C-terminal extensions sequence of every isolate, respectively (>80%)	Unknown
Pro5	COG5433/ PAAR_motif[pfam05488] ^b	ECSTEC94C_0675 of <i>Escherichia coli</i> STEC_94C (93%)	PAAR-repeat protein, transposase, IS4 family
Pro6	M35_like_1[cd11007]	ECUMN_1706 of <i>E. coli</i> UMN026 (30%)	Putative zinc-dependent metallopeptidase
Pro7	DUF4123[pfam13503]	–	Domain of unknown function
Pro8	UvrD_C_2[pfam13538]	EC55989_3316 of <i>E. coli</i> 55989 (99%)	UvrD/REP helicase
Pro9	IstB_IS21[pfam01695]	ECDEC10F_1826 of <i>E. coli</i> DEC10F (99%)	IstB-like ATP binding protein
Pro10	Sigma70_r4/RVE[pfam00665]	TnpA of <i>E. coli</i> CB9615 (99%)	Transposase for ISEC12
Pro11	PAAR_motif[pfam05448] ^b	EC30301_3278 of <i>E. coli</i> 3030-1 (85%)	PAAR/S-type pycocin domain
Pro12	PAAR_motif[pfam05448] ^b	SSJG_02256 of <i>Shigella</i> sp. D9 (67%)	PAAR/S-type pycocin domain
Vm	VgrG[COG3501]/COG4253	Evolved VgrG proteins of ED353 (92%)	The second VgrG protein in T6SS1 locus of APEC ED353
Vn	VgrG[COG3501]/COG4253	VgrG C-terminal extensions sequence of ED058 (68%)	Shortened VgrG protein in T6SS1 locus of APEC ED058

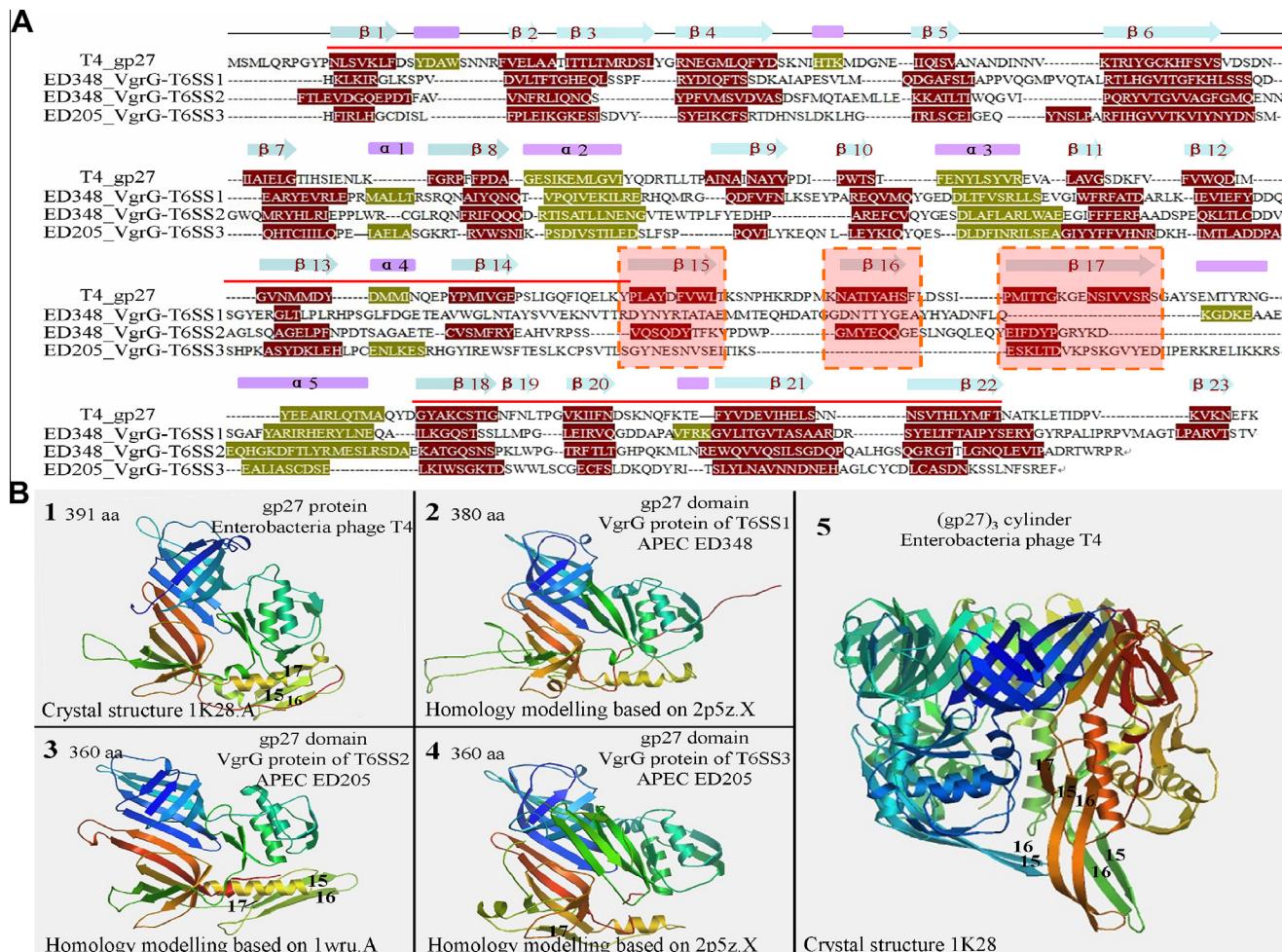
^a The number of unknown proteins in the VgrG islands is the same as that shown in Fig. 5C.^b The PAAR motif has been verified for the full functionality of T6SS is a must in a recent study (Shneider et al., 2013).

Fig. 6. The structure prediction and modeling of VgrGs proteins. (A) Sequence alignment and secondary structure of VgrGs proteins. The red line at the top of amino acid sequences indicates the first domain. The secondary structural elements are shown in the sequences as dark green blanket (α -helices) and dark red blanket (β -sheets). The α -helices and β -sheets of phage gp27 proteins are indicated as purple cylinders and blue arrows above the sequences again, respectively. The secondary structures of VgrG proteins are assigned using define secondary structure of proteins (DSSP) criteria of SWISS-MODEL (Arnold et al., 2006). (B) Homology modeling of gp27 homologues based on the crystal structure of c393 of *E. coli* CFT073 (PDB code: 2p5z.X) or tail protein of bacteriophage Mu (PDB code: 1wru.A). Stereoview of gp27 domains in ribbon representation and rainbow coloring, from blue (N-term) to red (C-term). These models were generated by employing the SWISS-MODEL server (<http://swissmodel.expasy.org>) (Arnold et al., 2006). Model 1 is the ribbon representation of the gp27 protein crystal structure (PDB code: 1K28.A) from phage T4. Model 5 is the gp27 trimer of phage T4 (PDB code: 1K28). The different parts of gp27's domain are labeled 15 β , 16 β and 17 β sheets among the four proteins according to the result of Fig. 6A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

three T6SS loci and gp27 protein of bacteriophage T4 was performed to provide a reference to predict the target cells diversity of T6SS.

The analysis showed that each monomer of gp27 domains contained at least 3 α -helices and 18 β -sheets (Fig. 6). The gp27 domains were divided into two smaller domains (Kondou et al., 2005). The first domain (approximately, residues 10–90, residues 160–185, and residues 275–340) located on the left side of the models and consisted primarily of β -sheets. And the second domain (approximately, residues 90–160, residues 185–275, and 340–370) folded into a $\alpha\beta$ -structure. APEC VgrG gp27 domains had the same folding topology as T4 phage gp27 (Fig. 6B). Despite the high degree of structural similarity, these four gp27 domains/proteins only shared 4% amino acid identity over the relevant 250 residue sequence (Fig. 6A). In particular, the first domain was more closely matched with the four proteins mentioned previously than the second domain (Fig. 6B). The DALI algorithm assigned 101 equivalent amino acid residues (59.41%) between the first domains of the four proteins, whereas the corresponding value for the second domain was only 64 equivalent positions (27.83%). Previous research has proved that the first domain mediated associations with other baseplate proteins in bacteriophage (Kanamaru et al., 2002; Kondou et al., 2005), while the second domain carried out frequent interaction with cell surface components. These results suggested that the second domain of gp27 might be a key unit of selecting host cells. The further analysis showed that the β 15, 16 and 17 sheets were absent or shortened in the second domain of T6SS1 and 3 gp27s (Fig. 6B). By observing the gp27 trimer, these three β -sheets were found to have more opportunities to perform extensive and in-depth exchanges interacting with cell surface (Fig. 6B). It seemed to suggest that had great potential to function as a receptor-binding protein. A previous research indicated that the shortening of these three β -sheets in gp44 (a tail fiber protein of Bacteriophage Mu) caused a more positive electrostatic potential (Kondou et al., 2005). Mu phage relies upon the electrostatic interactions of gp44 with the host cell membrane to initiate infection. This infection method would confer a low degree of host specificity and allow Mu phage to infect multiple different unrelated bacteria. This phenomenon seemed to tell us that the T6SS1 and 3 of APEC may have a wide host cells spectrum. The multipurpose T6SS1 clarified by previous description also proved this view. In addition, the second domain of T6SS2 gp27 kept a high similarity with one of phage gp27 protein, even longer β 15-sheet, β 16-sheet and α 5-helice were observed in T6SS2 gp27 domain (Fig. 6). These indicated that the T6SS2 and phage T4 might share a high degree of host specificity and explained why only APEC and NMEC had high expression level of VgrG-T6SS2 protein. The above description also suggested that there was an indistinct correlation between single-minded T6SS2 and its high gp27-like VgrG protein.

In summary, the VgrG proteins were not only related to the cluster characteristics (sequence identity, effectors coding and cluster structure) of T6SS loci, but also played a key role in specific functions formation of T6SS among APEC isolates by participating in recognition of host cells. Therefore, we developed a divergent thought that the functional characteristics of T6SS loci may be supervised by the VgrG proteins.

4. Conclusions

This study provides a basis for further focused investigations on this newly discovered and poorly understood secretion system in APEC. The three T6SS loci comparative genomics showed extensive diversity in prevalence analysis, number of core genes and sequence identity, which suggested that they might play respective roles in different mechanisms. The COG4253 and gp27 domains

of VgrGs were found to be associated with the transport of putative effector islands and conferring target cells specificity respectively, which clearly indicate the properties difference of T6SS might be supervised by the VgrG proteins.

Acknowledgments

This research was supported by the Special Fund for Advantage Discipline Construction of Colleges and Universities in Jiangsu, and National Nature Science Foundation (31372455), and Fundamental Research Funds for the Central Universities (KYZ201274).

We thank Jianjun Dai (Nanjing Agricultural University) and Yueming Wang (Shantung Academy of Agriculture) for kindly providing part of avian pathogenic *E. coli* isolates for this study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.meegid.2013.09.031>.

References

- Arisaka, F., Kanamaru, S., Leiman, P., Rossmann, M.G., 2003. The tail lysozyme complex of bacteriophage T4. *Int. J. Biochem. Cell Biol.* 35, 16–21.
- Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace. A web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.
- Bingle, L.E., Bailey, C.M., Pallen, M.J., 2008. Type VI secretion: a beginner's guide. *Curr. Opin. Microbiol.* 11, 3–8.
- Blondel, C.J., Jimenez, J.C., Contreras, I., Santiviago, C.A., 2009. Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in *Salmonella* serotypes. *BMC Genomics* 10, 354.
- Bonemann, G., Pietrosuk, A., Diemand, A., Zentgraf, H., Mogk, A., 2009. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J.* 28, 315–325.
- Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y., Attree, I., 2009. Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources? *BMC Genomics* 10, 104.
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the Artemis Comparison Tool. *Bioinformatics* 21, 3422–3423.
- Castascales, E., 2008. The type VI secretion toolkit. *EMBO Rep.* 9, 735–741.
- Clermont, O., Bonacorsi, S., Bingen, E., 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl. Environ. Microbiol.* 66, 4555–4558.
- De Maayer, P., Venter, S.N., Kamber, T., Duffy, B., Coutinho, T.A., Smits, T.H., 2011. Comparative genomics of the Type VI secretion systems of *Pantoea* and *Erwinia* species reveals the presence of putative effector islands that may be translocated by the VgrG and Hcp proteins. *BMC Genomics* 12, 576.
- de Pace, F., Nakazato, G., Pacheco, A., de Paiva, J.B., Sperandio, V., Da, S.W., 2010. The type VI secretion system plays a role in type 1 fimbria expression and pathogenesis of an avian pathogenic *Escherichia coli* strain. *Infect. Immun.* 78, 4990–4998.
- Dho-Moulin, M., Fairbrother, J.M., 1999. Avian pathogenic *Escherichia coli* (APEC). *Vet. Res.* 30, 299–316.
- Dreyfus, A., Schaller, A., Nivollet, S., Segers, R.P., Kobisch, M., Mieli, L., Soerensen, V., Hussy, D., Miserez, R., Zimmermann, W., Inderbitzin, F., Frey, J., 2004. Use of recombinant ApxIV in serodiagnosis of *Actinobacillus pleuropneumoniae* infections, development and prevalidation of the ApxIV ELISA. *Vet. Microbiol.* 99, 227–238.
- Ewers, C., Janssen, T., Wieler, L.H., 2003. Avian pathogenic *Escherichia coli* (APEC). *Berl. Munch. Tierarztl. Wochenschr.* 116, 381–395.
- Filloux, A., Hachani, A., Bleves, S., 2008. The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology* 154, 1570–1583.
- Ham, J.H., Melanson, R.A., Rush, M.C., 2011. *Burkholderia glumae*: next major pathogen of rice? *Mol. Plant Pathol.* 12, 329–339.
- Hood, R.D., Singh, P., Hsu, F., Guvenier, T., Carl, M.A., Trinidad, R.R., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L., Li, M., Schwarz, S., Wang, W.Y., Merz, A.J., Goodlett, D.R., Mougous, J.D., 2010. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* 7, 25–37.
- Jani, A.J., Cotter, P.A., 2010. Type VI secretion: not just for pathogenesis anymore. *Cell Host Microbe* 8, 2–6.
- Johnson, J.R., Owens, K.L., Clabots, C.R., Weissman, S.J., Cannon, S.B., 2006. Phylogenetic relationships among clonal groups of extraintestinal pathogenic *Escherichia coli* as assessed by multi-locus sequence analysis. *Microbes Infect.* 8, 1702–1713.
- Johnson, T.J., Kariyawasam, S., Wannemuehler, Y., Mangamele, P., Johnson, S.J., Doekkott, C., Skyberg, J.A., Lynne, A.M., Johnson, J.R., Nolan, L.K., 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares

- strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J. Bacteriol.* 189, 3228–3236.
- Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., Rossmann, M.G., 2002. Structure of the cell-puncturing device of bacteriophage T4. *Nature* 415, 553–557.
- Kohler, G., Milstein, C., 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *J. Immunol.* 174 (2005), 2453–2455.
- Kondou, Y., Kitazawa, D., Takeda, S., Tsuchiya, Y., Yamashita, E., Mizuguchi, M., Kawano, K., Tsukihara, T., 2005. Structure of the central hub of bacteriophage Mu baseplate determined by X-ray crystallography of gp44. *J. Mol. Biol.* 352, 976–985.
- Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C., Mekalanos, J.J., 2009. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci. USA* 106, 4154–4159.
- Lesic, B., Starkey, M., He, J., Hazan, R., Rahme, L.G., 2009. Quorum sensing differentially regulates *Pseudomonas aeruginosa* type VI secretion locus I and homologous loci II and III, which are required for pathogenesis. *Microbiology* 155, 2845–2855.
- Ma, A.T., McAuley, S., Pukatzki, S., Mekalanos, J.J., 2009. Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell Host Microbe* 5, 234–243.
- Mora, A., Lopez, C., Dabhi, G., Blanco, M., Blanco, J.E., Alonso, M.P., Herrera, A., Mamani, R., Bonacorsi, S., Moulin-Schouleur, M., Blanco, J., 2009. Extraintestinal pathogenic *Escherichia coli* O1:K1:H7/NM from human and avian origin: detection of clonal groups B2 ST95 and D ST59 with different host distribution. *BMC Microbiol.* 9, 132.
- Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordonez, C.L., Lory, S., Walz, T., Joachimiak, A., Mekalanos, J.J., 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* 312, 1526–1530.
- Moulin-Schouleur, M., Reperant, M., Laurent, S., Bree, A., Mignon-Grasteau, S., Germon, P., Rasschaert, D., Schouler, C., 2007. Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *J. Clin. Microbiol.* 45, 3366–3376.
- Moulin-Schouleur, M., Schouler, C., Tailliez, P., Kao, M.R., Bree, A., Germon, P., Oswald, E., Mainil, J., Blanco, M., Blanco, J., 2006. Common virulence factors and genetic relationships between O18:K1:H7 *Escherichia coli* isolates of human and avian origin. *J. Clin. Microbiol.* 44, 3484–3492.
- Nardini, M., Lang, D.A., Liebeton, K., Jaeger, K.E., Dijkstra, B.W., 2000. Crystal structure of *pseudomonas aeruginosa* lipase in the open conformation. The prototype for family I.1 of bacterial lipases. *J. Biol. Chem.* 275, 31219–31225.
- Parsons, D.A., Heffron, F., 2005. *SciS*, an *icmF* homolog in *Salmonella enterica* serovar Typhimurium, limits intracellular replication and decreases virulence. *Infect. Immun.* 73, 4338–4345.
- Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., Mekalanos, J.J., 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. USA* 104, 15508–15513.
- Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., Heidelberg, J.F., Mekalanos, J.J., 2006. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. USA* 103, 1528–1533.
- Robinson, J.B., Telepnev, M.V., Zudina, I.V., Bouyer, D., Montenieri, J.A., Bearden, S.W., Gage, K.L., Agar, S.L., Foltz, S.M., Chauhan, S., Chopra, A.K., Motin, V.L., 2009. Evaluation of a *Yersinia pestis* mutant impaired in a thermoregulated type VI-like secretion system in flea, macrophage and murine models. *Microb. Pathog.* 47, 243–251.
- Rodriguez-Siek, K.E., Giddings, C.W., Doekott, C., Johnson, T.J., Nolan, L.K., 2005a. Characterizing the APEC pathotype. *Vet. Res.* 36, 241–256.
- Rodriguez-Siek, K.E., Giddings, C.W., Doekott, C., Johnson, T.J., Fakhr, M.K., Nolan, L.K., 2005b. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. *Microbiology* 151, 2097–2110.
- Russell, A.B., Hood, R.D., Bui, N.K., LeRoux, M., Vollmer, W., Mougon, J.D., 2011. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* 475, 343–347.
- Sarris, P.F., Zoumatakis, C., Panopoulos, N.J., Scoulica, E.V., 2011. Distribution of the putative type VI secretion system core genes in *Klebsiella* spp. *Infect. Genet. Evol.* 11, 157–166.
- Shneider, M.M., Butch, S.A., Ho, B.T., Basler, M., Mekalanos, J.J., Leiman, P.G., 2013. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* 500, 350–353.
- Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., Mougon, J.D., 2013. Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. *Mol. Cell.* 51, 584–593.
- Wang, S., Niu, C., Shi, Z., Xia, Y., Yaqoob, M., Dai, J., Lu, C., 2011. Effects of *ibeA* deletion on virulence and biofilm formation of avian pathogenic *Escherichia coli*. *Infect. Immun.* 79, 279–287.
- Wang, Y.D., Zhao, S., Hill, C.W., 1998. Rhs elements comprise three subfamilies which diverged prior to acquisition by *Escherichia coli*. *J. Bacteriol.* 180, 4102–4110.
- Wassif, C., Cheek, D., Belas, R., 1995. Molecular analysis of a metalloprotease from *Proteus mirabilis*. *J. Bacteriol.* 177, 5790–5798.
- Woods, R.G., Burger, M., Beven, C.A., Beacham, I.R., 2001. The *aprX-lipA* operon of *Pseudomonas fluorescens* B52: a molecular analysis of metalloprotease and lipase production. *Microbiology* 147, 345–354.
- Xu, J., Zhang, J., Lu, X., Liang, W., Zhang, L., Kan, B., 2013. O antigen is the receptor of *Vibrio cholerae* serogroup O1 El Tor typing phage VP4. *J. Bacteriol.* 195, 798–806.
- Zhang, J., Li, W., Zhang, Q., Wang, H., Xu, X., Diao, B., Zhang, L., Kan, B., 2009. The core oligosaccharide and thioredoxin of *Vibrio cholerae* are necessary for binding and propagation of its typing phage VP3. *J. Bacteriol.* 191, 2622–2629.
- Zhou, Y., Tao, J., Yu, H., Ni, J., Zeng, L., Teng, Q., Kim, K.S., Zhao, G.P., Guo, X., Yao, Y., 2012. Hcp family proteins secreted via the type VI secretion system coordinately regulate *Escherichia coli* K1 interaction with human brain microvascular endothelial cells. *Infect. Immun.* 80, 1243–1251.