AEM
Journals.ASM.org

Identification and Mechanism of Evolution of New Alleles Coding for the AIDA-I Autotransporter of Porcine Pathogenic *Escherichia coli*

Jean-Philippe Côté,[a] Frédéric Berthiaume,[a] Sébastien Houle,[b] John M. Fairbrother,[a] Charles M. Dozois,[b] and Michael Mourez[a]

Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Québec, Canada,[a] and INRS-Institut Armand-Frappier, Laval, Québec, Canada[b]

Autotransporters are a large family of virulence factors of Gram-negative bacterial pathogens. The autotransporter adhesin involved in diffuse adherence (AIDA-I) is an outer membrane protein of *Escherichia coli*, which allows binding to epithelial cells as well as the autoaggregation of bacteria. AIDA-I is glycosylated by a specific heptosyltransferase encoded by the *aah* gene that forms an operon with the *aidA* gene. *aidA* is highly prevalent in strains that cause disease in pigs. Nevertheless, there are only two published whole-length sequences for this gene. In this study, we sequenced the *aah* and *aidA* genes of 24 *aidA*-positive porcine strains harboring distinct virulence factor profiles. We compared the obtained sequences and performed phylogenetic and pulsed-field electrophoresis analyses. Our results suggest that there are at least 3 different alleles for *aidA*, which are associated with distinct virulence factor profiles. The genes are found on high-molecular-weight plasmids and seem to evolve via shuffling mechanisms, with one of the sequences showing evidence of genetic recombination. Our work suggests that genetic plasticity allows the evolution of *aah-aidA* alleles that are selected during pathogenesis.

Escherichia coli is an important cause of severe diarrhea resulting in mortality and morbidity in humans and animals worldwide (26). Adhesion is often a critical first step in the infection process, and pathogenic *E. coli* strains employ a vast array of fimbrial or afimbrial adhesins (27). The adhesin involved in diffuse adherence (AIDA-I) is one such adhesin and was first identified in a strain isolated from a case of infantile diarrhea (1). However, AIDA-I is rarely found in *E. coli* strains that cause disease in humans (18, 29). In contrast, it is frequently found in *E. coli* strains that cause postweaning diarrhea or edema disease in pigs (24, 28, 29, 36, 37). Indeed, experimental infections have shown that AIDA-I is important for the development of disease in pigs (31).

AIDA-I belongs to the family of autotransporters, outer membrane proteins of Gram-negative bacteria characterized by their modular organization and secretion mechanism (13, 17). Autotransporters comprise (i) an N-terminal signal sequence for Sec-dependent translocation across the inner membrane, (ii) an extracellular central "passenger" domain, and (iii) a C-terminal domain embedded in the outer membrane. In several autotransporters, a conserved region between the passenger domain and the membrane-embedded domain forms a junction, or "stable core," which is important for secretion (30, 32).

The N-terminal part of the passenger domain of AIDA-I is unusual: it is formed by multiple repetitions of a 19-amino-acid consensus sequence. This characteristic is shared by two other *E. coli* autotransporters, the adhesin/invasin TibA (23) and the autoaggregation factor Ag43 (16). With AIDA-I, these proteins have been called self-associating autotransporters (SAATs), as they can mediate bacterial autoaggregation through self-associations (15, 22).

Two main alleles of *agn43* code for proteins with differences in the number of consensus sequence repetitions (33, 35). An allele of *tibA* with a different number of consensus sequence repetitions was also recently identified (19). Allelic variations in *agn43* have been associated with functional differences and distinct phylogenetic groups (21, 33). This suggests that there is a mechanism of evolution that allows the appearance of variations in the number of repetitions of the consensus sequence, but such a mechanism remains to be understood. Similar studies examining putative alleles of *aidA* are lacking. The *aidA* sequences from two strains, human pathogenic strain 2787 (1) and porcine pathogenic strain E393 (37), are known. Thus, the question of the existence of *aidA* alleles is still unresolved.

In porcine pathogenic strains of *E. coli*, *aidA* has been found in isolates associated with different pathogenicity mechanisms, corresponding to its association with distinct virulence factors. AIDA-I has been associated with heat-stable enterotoxin b (STb) of enterotoxigenic *E. coli* (ETEC) (14, 34). It has also been associated with the Stx2 toxin and F18 fimbriae of Shiga toxin-producing *E. coli* (STEC) or with F18 alone (nonpathogenic) (29). Finally, AIDA-I is present in some strains of enteropathogenic *E. coli* (EPEC), which are characterized by attaching/effacing lesions caused by intimate adhesion mediated in part by the adhesin intimin (6). These strains cause various diseases in pigs, including edema disease and postweaning diarrhea. Thus, it is possible that *aidA*, like *agn43*, has functionally distinct alleles.

In this study, we have sequenced the *aah-aidA* operon from 24 porcine pathogenic *E. coli* strains in order to investigate the genetic organization and the allelic distribution of *aah-aidA*. By comparing these sequences, along with ones previously known, we found that there are three different alleles for *aidA* and that the sequences are distributed, in part, according to virulence factor profiles. Our data also reveal insights into the functionality of AIDA-I and the mechanism of evolution of the *aidA* alleles.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *E. coli* strains 2787 and PD20 and 23 other porcine pathogenic *E. coli* strains isolated from pigs presenting dis-

Received 21 March 2012 Accepted 11 April 2012
Published ahead of print 20 April 2012
Address correspondence to Michael Mourez, m.mourez@umontreal.ca.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.00906-12

TABLE 1 Primers used in this study

Primer	Sequence (5′→3′)	Function
AIDA_F	TAATAACCATGGGTAACGCCTACAGTATCATATGG	Amplification and cloning of *aidA*
AIDA_R	TTATCTAGATTATCAGAAGCTGTATTTTATCC	
Aah_F	ATGACTTTCTTATCACCACC	Amplification of *aah*
Aah_R	AATGATCCTGTCATGATTGG	
Promo_F	CTCCGTATGTAACGCTGTG	Amplification of the region upstream of *aah*
Promo_R	GGTATCTCTGGTGGTGATAAG	
Seq_aidA1	ATGGGTAAGGCCTACAGTATCATATGG	Sequencing of *aidA*
Seq_aidA2	GGCGCTAATGCCACTGAGAC	
Seq_aidA3	CTGTGGCGAGGTATCTGG	
Seq_aidA4	CAGAGTTGTCAGGTAGTGG	
Seq_aidA5	TATGTCAATGAAGATGGCAG	
Seq_aidA6	GCTTCTGCAGGTGGGGG	
Seq_aidA7	GGTTTTCCCACCATTATTGACG	
Seq_aah1	CGGATAACCAGCCTGTGG	Sequencing of *aah*
Seq_aah2	CCACTTCAGCCCACCCC	
Seq_prom	CTCCGTATGTAACGCTGTG	Sequencing of the region upstream of *aah*
AIDA_1	ATGGGTAAGGCCTACAGTATCATATGG	Generation of DIG-coupled probes
DIG-AIDA-R	GTGGCATTAGCGCCAGCAG	
F18_Probe_F	GTGAAAAGACTAGTGTTTATTTC	
F18_Probe_R	CTTGTAAGTAACCGCGTAAGC	
STb_Probe_F	ATGAAAAAGAATATCGCATTTCTTC	
STb_Probe_R	TTAGCATCCTTTTGCTGC	

ease in Québec, Canada, in 2007 were provided by the Reference Laboratory for *E. coli* (Faculty of Veterinary Medicine, University of Montreal). The presence of genes encoding virulence factors was tested by conventional PCR and by colony hybridization. Bacteria were grown on Luria-Bertani (LB) agar plates or in liquid LB broth at 37°C overnight. A PD20 Δ*aidA* strain was previously generated by allelic exchange (31). Antimicrobial susceptibility was tested by using a disk diffusion assay (Kirby-Bauer). The following 8 antimicrobials were used: ampicillin (AMP) (10 μg), apramycin (AM) (15 μg), ceftiofur (TIO) (30 μg): florfenicol (FLR) (30 μg), neomycin (NEO) (30 μg), spectinomycin (SPCM) (100 μg), tetracycline (TET) (30 μg), and trimethoprim-sulfamethoxazole (SXT) (25 μg). The breakpoints used were those recommended by the CLSI (12).

E. coli strain C600 (F^- *thr-1 leuB6 thi-1 lacY1 supE44 rfbD1 fhuA21*) (New England Biolabs) was used for phenotypic assays. Bacteria containing the different plasmids were grown at 30°C on LB agar plates or in liquid LB medium containing 100 μg/ml ampicillin and 50 μg/ml chloramphenicol. Growth was monitored by measuring the turbidity as the optical density at 600 nm (OD_{600}). At an OD_{600} of 0.8, cultures were induced with 10 μM isopropyl-β-D-thiogalactopyranoside (IPTG). This low concentration of IPTG was used to limit toxicity associated with the overexpression of AIDA-I and/or Aah.

Sequencing of the *aah* and *aidA* genes. DNA was extracted by using the DNeasy blood and tissue kit (Qiagen) according to the instructions provided by the manufacturer. Following extraction, *aah*, *aidA*, and the region upstream of *aah* were amplified by PCR and sequenced by using the primers listed in Table 1.

Phylogenetic analysis. Sequence alignments were made with Geneious Pro 5.0.4 software (Biomatters Ltd.), using BLOSUM matrix 62. Phylogenetic trees were generated by using the neighbor-joining method and the Jukes-Cantor model. Silent variations, conservative variations, and nonconservative variations were designated by comparisons with the sequence of strain 2787.

SDS-PAGE and immunoblotting. Cultures (5 ml) grown overnight were adjusted to an identical optical density measured at 600 nm and centrifuged for 10 min at 12,000 × *g* in microcentrifuge tubes, and the pellets were resuspended in 50 μl of Tris-buffered saline (TBS). Suspensions were then diluted in 2× SDS-PAGE loading buffer containing β-mercaptoethanol, denatured by heating at 100°C for 10 min, and separated by SDS-PAGE on 10% acrylamide gels. The gels were transferred onto polyvinylidene fluoride membranes (Millipore). Immunodetection was performed with a serum raised against glycosylated AIDA-I (7), diluted 1:60,000 in blocking buffer (5% skim milk, 50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.05% Triton X-100). A goat anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (Sigma) was used as a secondary antibody according to the instructions provided by the manufacturer. Immune complexes were revealed by using a 3,3′,5,5′-tetramethylbenzidine solution for membranes (Sigma).

Plasmid extraction and pulsed-field gel electrophoresis. High-molecular-weight plasmids were extracted according to a procedure described previously by Kado and Liu (20). LB cultures of the tested strains grown overnight were collected, and the pellets were suspended in E buffer (40 mM Tris, 2 mM EDTA [pH 7.8]). Bacteria were incubated in lysis buffer (50 mM Tris, 3% SDS [pH 12.6]) for 1 h at 56°C. A crude phenol-chloroform mix (1:1) was added, and the solution was emulsified by inversion. The emulsion was broken by centrifugation, and the upper aqueous phase was transferred into a microcentrifuge tube. Electrophoresis was performed on 1% agarose gels in 0.5× Tris-borate-EDTA (TBE) buffer on a CHEF-Mapper XA pulsed-field gel electrophoresis (PFGE) system (Bio-Rad) for 20 h at 14°C, with run conditions of 6 V/cm, a pulse angle of 120°, and pulse times from 0.47 to 10.29 s.

DNA hybridization. Southern hybridization was performed according to standard methods. Briefly, high-molecular-weight plasmids were transferred from the gel by a modified downward alkaline capillary transfer method (11). The gel was stained with a 1-μg/ml ethidium bromide solution for 30 min, and the DNA was nicked with the use of a GS Gene Linker UV chamber (Bio-Rad) at 60 mJ. The gel was then destained for 20 min and equilibrated for 15 min in 0.4 N NaOH–1.5 M NaCl. The gel was then put onto a positively charged nylon membrane (Boehringer Mannheim), and the transfer proceeded for 40 h. The membrane was finally neutralized in 0.5 M

TABLE 2 Strains used in this study

Strain	Virulence factor profile[a]	Host	Location	Antimicrobial resistance(s)[b]	Reference
14430	AIDA, F18	Pig	Canada	AMP, AM, TET	This study
15139	AIDA, F18	Pig	Canada	AMP, TET, SXT	This study
15194	AIDA, F18	Pig	Canada	AMP, AM, FLR, TET	This study
17530	AIDA, F18	Pig	Canada	AM, NEO	This study
15334	AIDA, F18	Pig	Canada	AMP, FLR, TET	This study
15372	AIDA, F18	Pig	Canada	AMP, FLR, NEO, TET	This study
14596	Stx2, East1, AIDA, F18	Pig	Canada	AMP, NEO, TET, SXT	This study
14603	Stx2, East1, AIDA, F18	Pig	Canada	AMP, TET, SXT	This study
15478	Stx2, East1, AIDA, F18	Pig	Canada	AMP, FLR, TET	This study
14377	Stx2, East1, AIDA, F18	Pig	Canada	AMP, FLR, TET, SXT	This study
15742	Stx2, East1, AIDA, F18	Pig	Canada	AMP, NEO, TET, SXT	This study
15941	Stx2, East1, AIDA, F18	Pig	Canada	AMP, FLR, TET	This study
14748	Eae, AIDA	Pig	Canada	AMP, TET	This study
14503	Eae, AIDA	Pig	Canada	TET	This study
2010	Eae, AIDA	Pig	Canada	AMP, TIO, FLR, TET	This study
2020	Eae, AIDA	Pig	Canada	FLR, TET	This study
2034	Eae, AIDA	Pig	Canada	AMP, FLR, TET	This study
14466	STb, AIDA, East1	Pig	Canada	AMP, TET	This study
14326	STb, AIDA, East1	Pig	Canada	AMP, AM, SPCM, TET, SXT	This study
14707	STb, AIDA, East1	Pig	Canada	FLR, SPCM, TET	This study
14675	STb, AIDA, East1	Pig	Canada	AMP, SPCM, TET, SXT	This study
15146	STb, AIDA, East1	Pig	Canada	AMP, AM, NEO, TET, SXT	This study
14187	STb, AIDA, East1	Pig	Canada	AMP, TET, SXT	This study
2787	East1, AIDA	Human	Germany	—[c]	1
PD20	STb, AIDA	Pig	Canada	—[c]	28
E393	STb, East1, AIDA	Pig	China	ND[d]	37

[a] Strains tested by colony hybridization using radioactively labeled (^{32}P) DNA probes for the genes encoding the virulence factors STa, STb, heat-labile enterotoxin (LT), East1, Stx1, Stx2, cytotoxic necrotizing factor (CNF), F4, F17, F18, F5, F6, F41, P, Eae, AIDA, Paa, AFA, Tsh, and aerobactin.
[b] The antimicrobials tested were as follows: ampicillin (AMP), apramycin (AM), ceftiofur (TIO), florfenicol (FLR), neomycin (NEO), spectinomycin (SPCM), tetracycline (TET), and trimethoprim-sulfamethoxazole (SXT).
[c] —, no antimicrobial resistance was found.
[d] ND, no data on the antimicrobial resistance of this strain.

Tris-HCl (pH 7) for 5 min, rinsed briefly in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and dried on blotting paper.

The membranes were rehydrated in DIG Easy Hyb buffer (Roche) at 42°C for 1 h. Hybridization was performed with digoxigenin (DIG)-coupled probes, generated by PCR using DIG DNA labeling mix (Roche) and corresponding to the 683 nucleotides at the N terminus of *aidA*, to the *estB* gene, or to the 230 nucleotides at the N terminus of *fedA* (29), the gene coding for the main subunit of F18 fimbriae (Table 1). Hybridization was performed with the denatured probes (10 min at 100°C), diluted 1:1,000 in DIG Easy Hyb buffer at 42°C for 18 h with gentle agitation. The probe was then washed with 2× SSC–0.1% SDS and revealed with an anti-DIG antibody conjugated to horseradish peroxidase (Anti-DIG-POD; Roche) diluted 1:1,000 in blocking buffer. Immune complexes were revealed by using a 3,3′,5,5′-tetramethylbenzidine solution for membranes (Sigma).

Cloning of AIDA-I alleles. The *aidA* genes from strains 2787, 14326 (STb), and 17530 (F18) were amplified by PCR using primers AIDA_F and AIDA_R (Table 1), using the Expand Long Template PCR system (Roche) according to the instructions provided by the manufacturer. The PCR products were then cloned into the pTRC99A vector (Pharmacia Biotech), using NcoI and XbaI. Constructions were verified by restriction mapping and sequencing.

Phenotypic assays. Autoaggregation, biofilm formation, and adhesion assays were performed as previously described (10). *E. coli* C600 cells harboring a plasmid allowing the expression of Aah (8) and the pTRC99A vector or plasmids allowing the expression of the different alleles of AIDA-I were induced with 10 μM IPTG at an OD_{600} of 0.8 and grown overnight in LB (autoaggregation and adhesion assays) or M9 (biofilm assay) medium. For the autoaggregation assay, the cultures grown overnight were normalized in 5 ml of LB to an OD_{600} of approximately 1.5 in culture tubes and left standing at 4°C. Samples (100 μl) were taken 1 cm below the surface at the beginning of the assay and after 120 min, and the OD_{600} values of the samples were measured. OD_{600} values at 120 min were compared to OD_{600} values at the beginning of the assay, and results are shown as percentages of the initial OD_{600}. For the biofilm formation assay, the cultures grown overnight were normalized in M9 minimal medium and grown for 24 h at 30°C in plastic microtiter plates. Biofilms were stained with crystal violet. After washes, the dye was then solubilized with a mixture of ethanol and acetone (4:1), and the absorption of the solution at 595 nm was measured. For the adhesion assay, the cultures grown overnight were inoculated onto monolayers of confluent HEp-2 cells in a 24-well plate (approximately 2.5×10^5 cells), using 10^6 CFU per well. After 3 h, the cells were washed with phosphate-buffered saline (PBS), and the adhering bacteria were recovered with 100 μl of Triton X-100 (1%), plated, and counted.

All functional assays were performed at least three times in duplicate or triplicate. For each assay, the results obtained were compared to the 2787 allele by performing an analysis of variance (ANOVA) and Dunnett posttests using Prism 4.0 software (Graphpad Software).

Nucleotide sequence accession numbers. The sequences of *aah*, *aidA*, and the region upstream of *aah* were deposited into the GenBank database under accession numbers HQ641298 to HQ641310 and JQ044400 to JQ044410.

RESULTS AND DISCUSSION

Assembly of a collection of *aidA*$^+$ strains with differing virulence gene profiles. We selected 24 strains of *E. coli* isolated from pigs and harboring the *aidA* gene, as demonstrated by PCR. Our selection included the previously described porcine pathogenic strain PD20 (28). The presence of the various virulence genes was determined at the Reference Laboratory for *E. coli* of the Faculty of

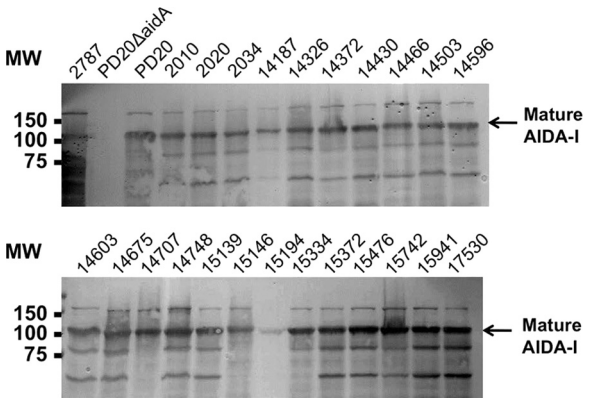

FIG 1 Expression of AIDA-I. Whole-cell lysates from normalized cultures of 24 porcine isolates grown overnight were separated by SDS-PAGE and probed by immunoblotting with an antibody raised against glycosylated AIDA-I. Whole-cell lysates from strain 2787 and the PD20 Δ*aidA* strain were added as a positive control and as a negative control, respectively. MW, molecular weight (in thousands).

Veterinary Medicine, Université de Montréal (Table 2). These strains represent the 4 distinct virulence factor profiles and pathogenicity mechanisms described above: (i) F18 alone (nonpathogenic); (ii) F18, Stx2, and the thermostable toxin East1 (STEC); (iii) the intimin Eae (EPEC); or (iv) STb with or without East1 toxins (ETEC). The two strains for which *aidA* sequences are known are ETEC strains, with the profiles STb-East1 (E393) or East1 (2787). Other similarities and differences exist between the 24 porcine strains and the previously studied ones: all newly selected strains were isolated in Canada, whereas 2787 was isolated in Europe and E393 was from China. In addition, all strains were recovered from pigs, except for strain 2787, which was recovered from a child. To assess the diversity of our strains, we also determined their antimicrobial resistance profiles (Table 2). Different resistance profiles were observed, showing that the strains were not identical clones.

Expression of AIDA-I. We first looked at the expression of AIDA-I from the 24 porcine strains by blotting total protein extracts and probing them with an antibody raised against AIDA-I (Fig. 1). In the blot, we included total protein extracts from strain 2787 as a positive control and from PD20 where *aidA* had been deleted as a negative control (PD20 Δ*aidA*) (31). All strains expressed AIDA-I at similar levels, except for strain 15194. In this strain, AIDA-I is not well expressed. There is no difference in the sequences of the promoter regions, *aah*, or *aidA* that could explain this lowered expression level (see below). It is therefore possible that the *aah-aidA* operon is regulated differently in this strain.

AIDA-I is matured by self-cleavage (9) and is glycosylated by the heptosyltransferase Aah (2). Our antibody specifically recognizes the glycosylated AIDA-I fragment (8). Thus, our results show that both Aah and AIDA-I are expressed, consistent with the fact that genes coding for these proteins form an operon (3, 4). Furthermore, the polypeptides for each of the strains had the same apparent size, which suggests that there was no important variation in the number of consensus sequence repetitions.

Building of sequence alignments and phylogenetic trees. From the 24 porcine strains, we amplified by PCR and sequenced *aah*, *aidA*, and the promoter region upstream of *aah*. We aligned the resulting sequences of *aah* and *aidA*, along with the sequences from 2787 and E393. All *aah* and *aidA* sequences had the same length, confirming that there was no variation in

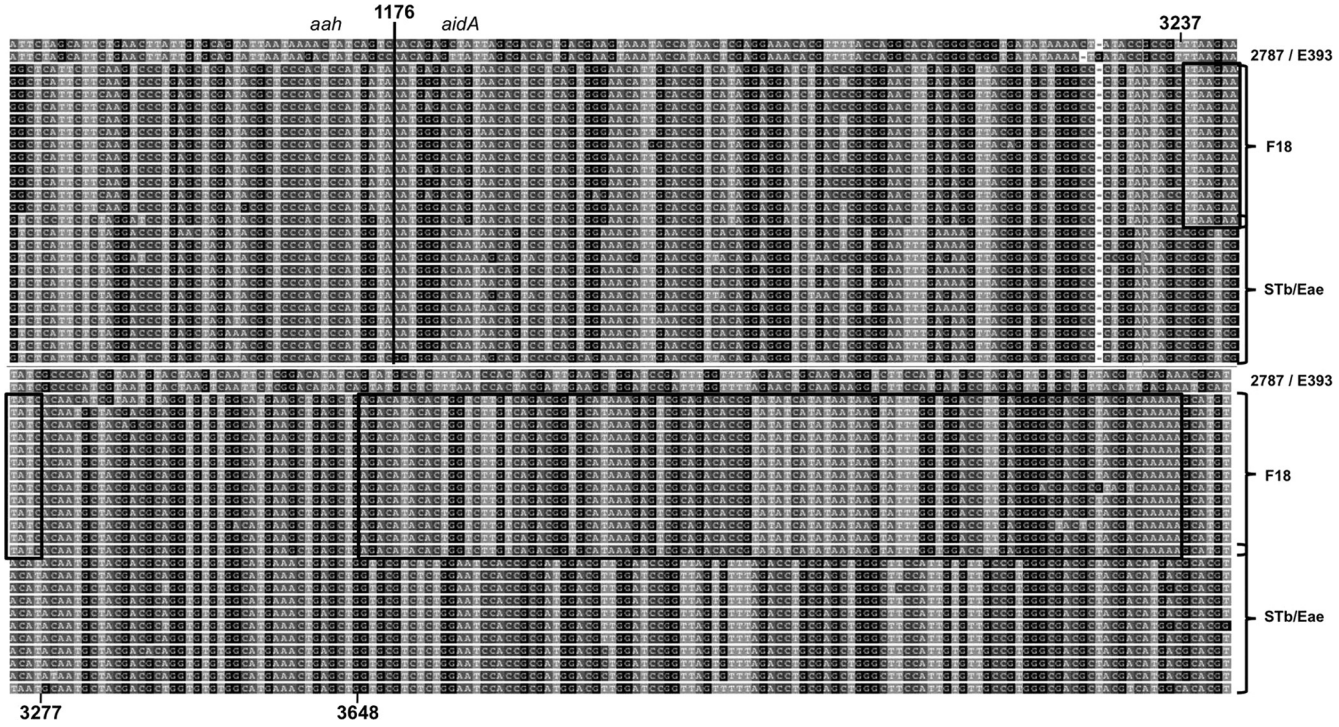

FIG 2 Polymorphism in the *aah-aidA* operon. The sequences of the operons were aligned by using the BLOSUM 62 matrix. Shown here are only the nucleotides showing polymorphisms. We omitted all the nucleotides that were identical for all the sequences in order to show the pattern of polymorphisms. Position 1176 corresponds to the end of *aah* and the start of *aidA*.

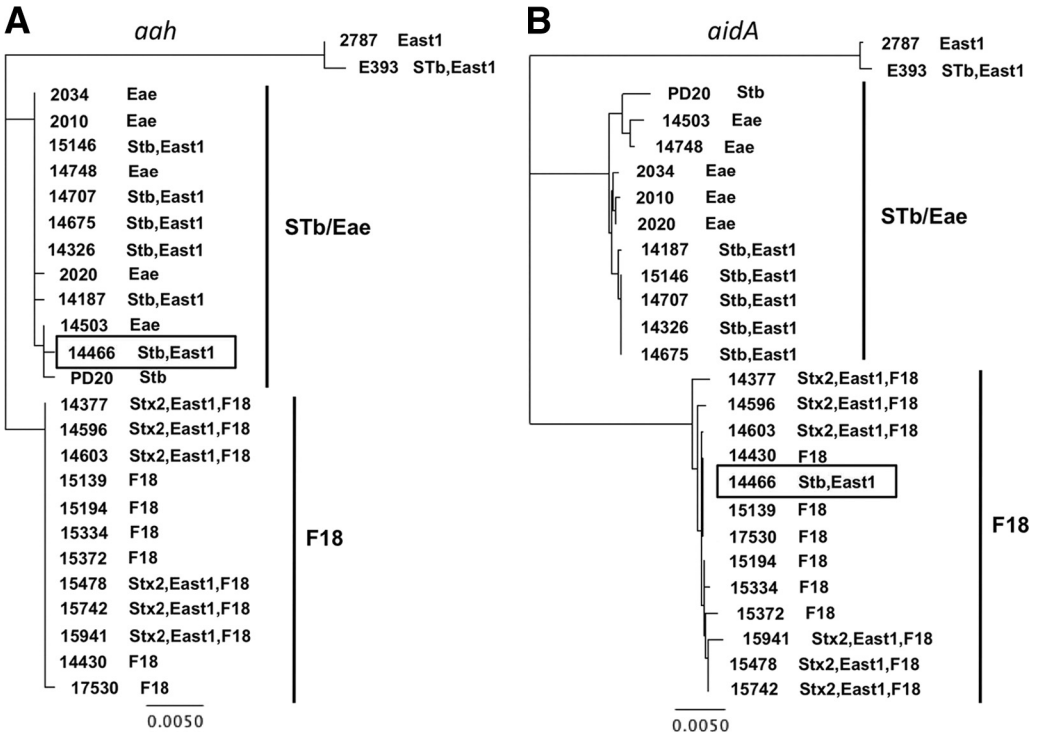

FIG 3 Phylogenetic analysis of *aah* and *aidA*. *aah* (A) and *aidA* (B) were aligned separately, and a phylogenetic tree was built from these alignments by using the neighbor-joining method.

the number of repetitions in *aidA*, and were identical over 96% and 93.2% of the gene, respectively. In Fig. 2, we report all the nucleotide positions for which there were variations among the sequences and only these positions. Variations in the part of the sequences corresponding to *aah* and the first half of *aidA* reveal two distinct groups: the sequences from 2787 and E393 are similar to each other and differ from those of all the other strains. Variations in the second half of *aidA* (corresponding to the junction region and the membrane-embedded domain) reveal three groups: the sequences of 2787 and E393 are still similar to each other and are different from those of the other strains. However, the latter strains were separated into two groups: one consisting of the strains with genes coding for the F18 fimbriae and the other consisting of the strains with genes coding for the STb toxin or intimin (Eae). The presence of three distinct groups of sequences was further evidenced by the building of phylogenetic trees with the *aah* and *aidA* sequences (Fig. 3). Again, strains 2787 and E393 clustered away from all the other strains, which were split into the F18 and STb/Eae groups. The sequence from one strain, 14466, clustered with the F18 or the STb/Eae group depending on whether it was the *aah* or *aidA* sequence that was considered in the analysis (Fig. 3). This is discussed below.

Finally, we amplified the promoter region upstream of *aah*, using primers based on the known sequence from 2787 (4). The sequencing of the PCR fragments indicated a difference between strain 2787 and all the other strains (Fig. 4). For instance, one of the two directly repeated sequences upstream of the *aah* Shine-Dalgarno sequence is absent in all strains but 2787 (Fig. 4A). The sequences of the promoter also highlight some differences between the STb- and the F18-expressing strains (Fig. 4B). Of note, there is an insertion of 3 nucleotides (TAT) near the −10 region of the promoter in the STb-expressing strains and some Eae-expressing strains. The transcription start site of P149 in these strains might therefore be slightly different. The consequence of these changes should be investigated in future experiments.

Taken together, our results suggest that one determinant for sequence variations is the presence of specific virulence genes, namely, those encoding F18 and STb/Eae. Since the receptor for F18 is expressed only in older pigs, strains possessing F18 usually infect older animals (25), whereas strains producing STb or intimin may be present in younger pigs. Thus, strains expressing these virulence factors might segregate, which could result in sequence differences. In addition, the similarity of sequences from strains expressing STb or Eae is surprising, since they correspond to different pathogenicity mechanisms, those of ETEC and EPEC, respectively.

Additional factors play an important role, however, as the sequence from E393 was very different from that of strains with the same STb-East1 profile. Such a difference could be due to the geographical origin of the strains: since the sequences from 2787 and E393 are highly similar, it is possible that strains from Eurasia differ from those isolated in North America. More sequences from Eurasian strains are needed to validate this hypothesis. Finally, the host species origin does not seem to be a factor for sequence variability, since E393 and 2787 have similar sequences but were isolated from human and pig, respectively.

Mapping of the positions for which there are genetic variations. We observed a bias when we averaged the number of positions for which there are variations among the *aah-aidA* sequences over a window of 20 nucleotides. Far more variations clustered in the region coding for the junction region of AIDA-I

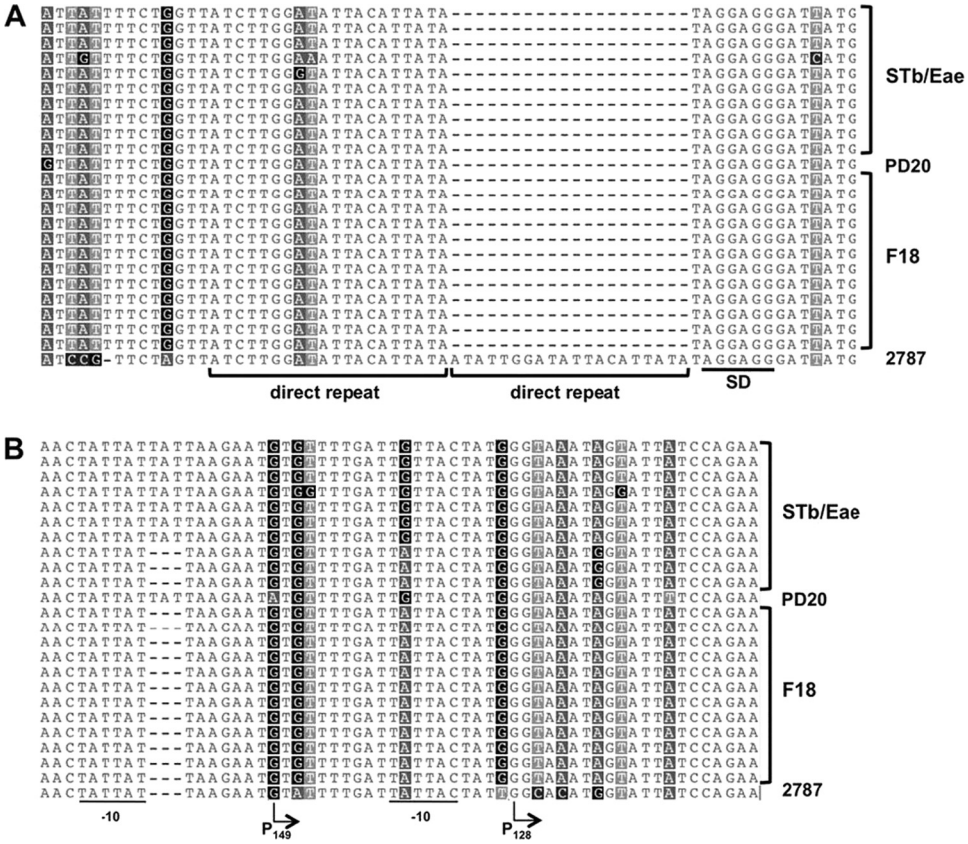

FIG 4 Promoter region upstream of *aah*. The sequences of the region upstream of *aah* were aligned to show the direct repeat right in front of *aah* (A) and the two promoters at positions −128 and −149 (B). SD, Shine-Dalgarno sequence.

and two smaller regions in the C terminus of the passenger domain (Fig. 5A). These regions correspond to blocks of sequence that are different between alleles. These blocks are readily identifiable in the alignment of the nucleotides for which there are variations. In fact, the block corresponding to the junction region extends to include the sequence coding for the whole membrane-embedded domain (Fig. 2). This finding suggests that the sequence of *aidA* has evolved by the acquisition of blocks of nucle-

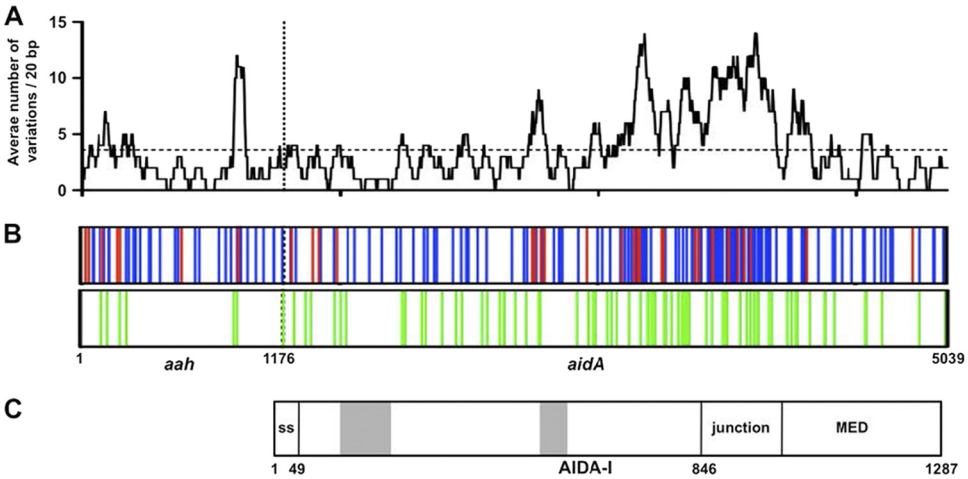

FIG 5 Regions of variability and conservation in the *aah-aidA* operon. Polymorphisms (shown in Fig. 2) were placed at their actual positions in the operon. (A) The number of variations was averaged over 20 nucleotides along the sequence, and this average number of variations was plotted for each position. Over the length of the *aah-aidA* operon, there are 3.39 variations/20 nucleotides. (B) Variations were separated into three categories: silent variations (blue), conservative variations (red), and nonconservative variations (green). To determine in which category a polymorphism should be placed, we compared it to the sequence from strain 2787. (C) Schematic representation showing the organization of AIDA-I: the signal sequence (ss); the passenger domain, with adhesion domains in gray; the junction region; and the membrane-embedded domain (MED).

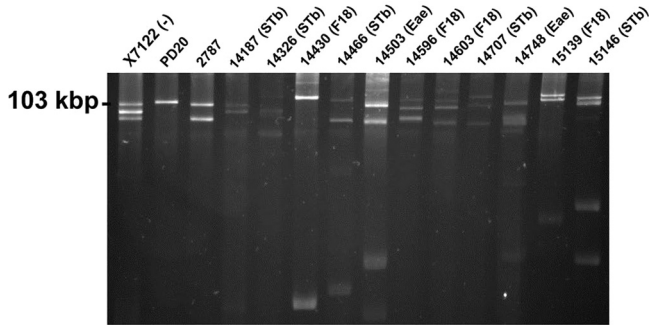

FIG 6 Localization of *aidA* on large plasmids. Plasmids from 12 of the porcine isolates as well as from strain 2787 were extracted and separated by pulsed-field gel electrophoresis and revealed by ethidium bromide staining. A plasmid extract from strain χ7122 was used as a molecular weight control.

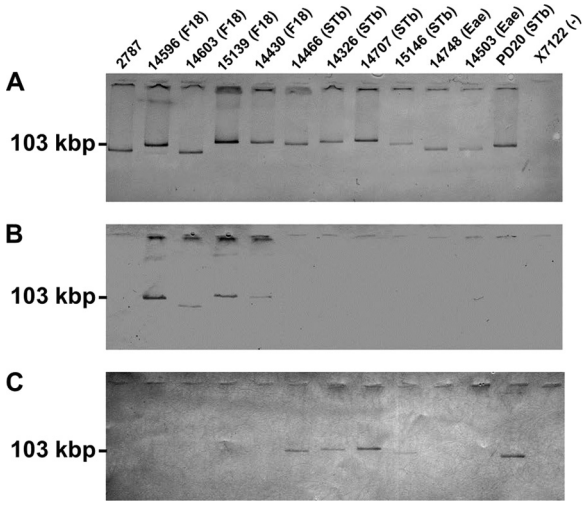

FIG 7 Colocalization of *aidA* and other virulence factors. Plasmids from 11 of the porcine isolates as well as from strain 2787 were extracted and separated by pulsed-field gel electrophoresis, transferred onto a nitrocellulose membrane, and probed with a DIG-coupled oligonucleotide specific for *aidA* (A), *fedA* (B), or *estB* (C). A plasmid extract from strain χ7122 was used as a negative control.

otides, in addition to the accumulation of point mutations. It is striking in particular that the *aidA* sequences from the STb-Eae and F18 groups have nearly identical passenger domains but distinct junction and membrane-embedded domains. A mechanism involving modular changes could also contribute to the evolution of other autotransporter genes. Indeed, domain shuffling has been implicated in the evolution of some autotransporters, such as EspP (5) and Ag43. For the latter, there are two main alleles of *agn43* that differ in length, with one allele having lost the C-terminal part of the passenger domain (35). The "full-length" allele is further subdivided into three different subfamilies. Indeed, by looking at the polymorphisms of these subfamilies, blocks of sequence that have been exchanged are also easily identifiable (see Fig. S1 in the supplemental material).

We then distinguished variations resulting in conservative or silent mutations from variations resulting in nonconservative mutations. Positions resulting in silent and conservative mutations are spread throughout the sequence, but some regions of the gene encoding AIDA-I clearly lack variations resulting in nonconservative mutations (Fig. 5B). In particular, the sequences coding for the C-terminal membrane-embedded domain and the two previously identified adhesion regions of the passenger domain (10) were highly conserved. This is not surprising, since these domains of the protein play key roles in biogenesis and function, respectively. It is surprising, however, that there are many variations resulting in nonconservative mutations in the junction region, since the junction region is known to be important for the translocation and folding of the passenger domain (30, 32), and one would expect it to be more sensitive to mutations.

Localization of the *aah-aidA* genes on large plasmids. The evolution of three *aah-aidA* alleles could be due to the localization of these genes on specific and different genetic elements. Indeed, *aidA* has been found on large plasmids also coding for the F18 fimbriae (24). It is possible that the *aidA* allele corresponding to the STb-Eae group is located on a different plasmid than that corresponding to the F18 group. Using pulsed-field gel electrophoresis, we investigated the presence of plasmids extracted from a subset of 13 strains, including strain 2787 (Fig. 6). All strains contained large plasmids, although there was no common pattern of plasmids corresponding to *aidA* alleles in the F18 or STb-Eae group. This heterogeneity confirms that none of the strains were identical clones, in agreement with the phenotypic diversity that we noted when we selected the strains (Table 2). We used an *aidA*-specific probe to determine which plasmid harbored the *aah-aidA* operon in each strain (Fig. 7A). We observed that the *aah-aidA* operon is present on large plasmids that vary in size. In a previous study, it was also reported that large plasmids extracted from strains expressing F18 and harboring *aidA* had various sizes (24). Thus, the sizes of plasmids harboring *aidA* do not seem to correlate with specific *aidA* alleles.

We further used *estB*- and *fedA*-specific probes to determine if *aidA* colocalized on plasmids encoding these other virulence factors (Fig. 7). In every strain, *aidA* was found on the same plasmid as the genes coding for either STb or the F18 fimbriae. This finding suggests that *aidA* is physically linked to these other virulence factor genes.

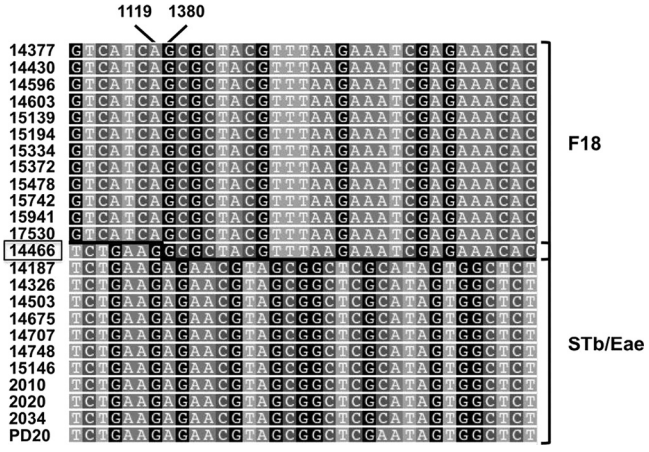

FIG 8 Recombination of the sequence from strain 14466. The first half of the polymorphisms defining the STb group and the F18 group were used to analyze the position where the recombination of strain 14466 occurred. From the alignment shown in Fig. 2, we selected only the polymorphisms specifically defining the STb group and the F18 group. We omitted strains 2787 and E393 because they form a separate group. We also omitted polymorphisms that were specific for only one sequence.

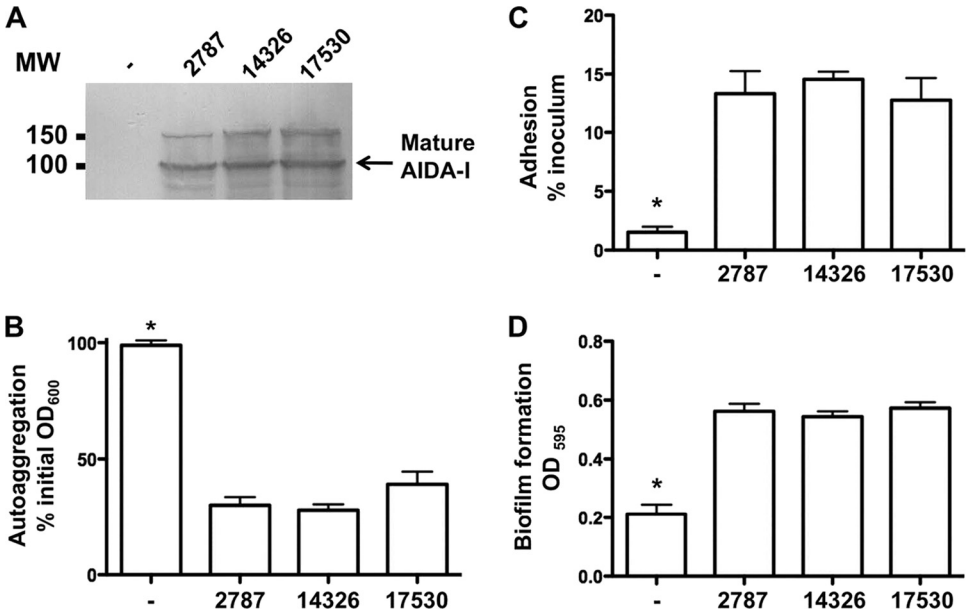

FIG 9 Functionality of the three alleles. (A) *E. coli* C600 cells bearing a plasmid allowing the expression of Aah from strain 2787 and an empty vector (−) or a plasmid allowing the expression of AIDA-I from strain 2787, 14326 (STb), or 17530 (F18) were probed with an antibody raised against glycosylated AIDA-I. (B to D) The bacteria were also assessed for their abilities to mediate autoaggregation (B), adhesion to Hep-2 cells (C), and the formation of biofilms (D). The functional assays were made in triplicate. ANOVA and Dunnett posttests were used to identify significant differences from the 2787 allele (∗, $P < 0.05$).

Overall, our results suggest that *aidA* alleles are not harbored by specific plasmids. Therefore, the association of different alleles with distinct virulence factors might stem from the fact that *aidA* genes are readily moving between plasmids and/or that the plasmids harboring *aidA* are easily gaining or losing genes.

Fortuitously, the sequences which we obtained illustrate this plasticity. Indeed, the sequence of *aah* from strain 14466 clusters with the STb-Eae group, whereas the sequence of *aidA* from this strain clusters with the F18 group (Fig. 3). This suggests that there was a recombination event between the alleles of the two groups. When only the variations specific for the F18 and STb-Eae groups are represented (Fig. 8), it becomes evident that recombination occurred between nucleotides 1119 and 1380 of the *aah-aidA* operon, which corresponds to the limit between *aah* and *aidA*.

Functional properties of the AIDA-I variants. The three alleles of *aidA* possess a similar passenger domain. In order to determine if they are also functionally similar, we cloned one representative for each allele. We previously cloned and characterized the functionality of *aidA* from strain 2787 (10). We further cloned *aidA* from strains 14326 (STb-Eae group) and 17530 (F18 group) into the same expression vector. Glycosylation is important for the stability of AIDA-I (8). Therefore, we cotransformed the plasmids allowing the expression of the different alleles of AIDA-I with a plasmid allowing the expression of Aah from strain 2787.

The three alleles were expressed and glycosylated properly, as shown by total protein extracts immunoblotted with an antibody raised against glycosylated AIDA-I (Fig. 9A). We then tested the constructs for their abilities to mediate adhesion to cultured epithelial cells, the autoaggregation of bacteria, and the formation of biofilms. The three alleles mediated the same levels of adhesion, autoaggregation, and formation of biofilms (Fig. 9B to D). This is not surprising, since the sequences indicated that the passenger domain is well conserved, especially for the parts of the passenger domain that are associated with the functionality of the protein (Fig. 5).

In conclusion, we propose the existence of at least three different alleles of the *aah-aidA* operon. One group corresponds to strains 2787 and E393, which could reflect their Eurasian geographical origin. The two remaining groups are divided according to their virulence gene profiles: the presence of the F18-encoding genes versus the presence of the STb toxin- or the Eae adhesin-encoding genes. The evolution of these alleles seems to rely on a high genetic plasticity involving mutations as well as the shuffling of whole blocks of sequence and recombination. Sequence variations could then become fixed during the segregation of strains, through self-interactions of the new alleles and/or other factors. Such a mechanism allowing a quick apparition and spread of new alleles of a self-associating surface protein could be an important mechanism to permit the recognition of "self" in bacteria. In this respect, the association of alleles with specific virulence factors may indicate the emergence of strains with a specific pathogenic process. This idea remains to be tested.

ACKNOWLEDGMENTS

This work was supported by the Canadian Institutes for Health Research (M.M.) and the Natural Sciences and Engineering Research Council (C.M.D.) as well as by funds from the Groupe de Recherche sur les Maladies Infectieuses du Porc and the Canada Research Chair program (M.M. and C.M.D.). J.-P.C. was supported by a graduate fellowship from the Natural Sciences and Engineering Research Council.

REFERENCES

1. **Benz I, Schmidt MA.** 1989. Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic Escherichia coli. Infect. Immun. **57**:1506–1511.
2. **Benz I, Schmidt MA.** 2001. Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin. Mol. Microbiol. **40**:1403–1413.

3. **Benz I, van Alen T, Bolte J, Wormann ME, Schmidt MA.** 2010. Modulation of transcription and characterization of the promoter organization of the autotransporter adhesin heptosyltransferase and the autotransporter adhesin AIDA-I. Microbiology **156**:1155–1166.
4. **Berthiaume F, Leblond MF, Harel J, Mourez M.** 2010. Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic Escherichia coli. FEMS Microbiol. Lett. **311**:176–184.
5. **Brockmeyer J, et al.** 2007. Subtypes of the plasmid-encoded serine protease EspP in Shiga toxin-producing Escherichia coli: distribution, secretion, and proteolytic activity. Appl. Environ. Microbiol. **73**:6351–6359.
6. **Bruant G, et al.** 2009. Two distinct groups of porcine enteropathogenic Escherichia coli strains of serogroup O45 are revealed by comparative genomic hybridization and virulence gene microarray. BMC Genomics **10**:402. doi:10.1186/1471-2164-10-402.
7. **Charbonneau ME, Berthiaume F, Mourez M.** 2006. Proteolytic processing is not essential for multiple functions of the Escherichia coli autotransporter adhesin involved in diffuse adherence (AIDA-I). J. Bacteriol. **188**:8504–8512.
8. **Charbonneau ME, et al.** 2007. O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. J. Bacteriol. **189**:8880–8889.
9. **Charbonneau ME, Janvore J, Mourez M.** 2009. Autoprocessing of the Escherichia coli AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. J. Biol. Chem. **284**:17340–17351.
10. **Charbonneau ME, Mourez M.** 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. J. Bacteriol. **189**:9020–9029.
11. **Chomczynski P.** 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. Anal. Biochem. **201**:134–139.
12. **Clinical and Laboratory Standards Institute.** 2008. Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals; approved standard. CLSI document M31-3, 3rd ed. Clinical and Laboratory Standards Institute, Wayne, PA.
13. **Dautin N, Bernstein HD.** 2007. Protein secretion in gram-negative bacteria via the autotransporter pathway. Annu. Rev. Microbiol. **61**:89–112.
14. **Dubreuil JD.** 2010. STb and AIDA-I: the missing link? Crit. Rev. Microbiol. **36**:212–220.
15. **Girard V, et al.** 2010. Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. J. Biol. Chem. **285**:10616–10626.
16. **Henderson IR, Meehan M, Owen P.** 1997. Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in Escherichia coli K-12. FEMS Microbiol. Lett. **149**:115–120.
17. **Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D.** 2004. Type V protein secretion pathway: the autotransporter story. Microbiol. Mol. Biol. Rev. **68**:692–744.
18. **Jallat C, Livrelli V, Darfeuille-Michaud A, Rich C, Joly B.** 1993. Escherichia coli strains involved in diarrhea in France: high prevalence and heterogeneity of diffusely adhering strains. J. Clin. Microbiol. **31**:2031–2037.
19. **Johnson TJ, et al.** 2010. Pyrosequencing of the Vir plasmid of necrotoxigenic Escherichia coli. Vet. Microbiol. **144**:100–109.
20. **Kado CI, Liu ST.** 1981. Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. **145**:1365–1373.
21. **Klemm P, Hjerrild L, Gjermansen M, Schembri MA.** 2004. Structure-function analysis of the self-recognizing antigen 43 autotransporter protein from Escherichia coli. Mol. Microbiol. **51**:283–296.
22. **Klemm P, Vejborg RM, Sherlock O.** 2006. Self-associating autotransporters, SAATs: functional and structural similarities. Int. J. Med. Microbiol. **296**:187–195.
23. **Lindenthal C, Elsinghorst EA.** 1999. Identification of a glycoprotein produced by enterotoxigenic Escherichia coli. Infect. Immun. **67**:4084–4091.
24. **Mainil JG, Jacquemin E, Pohl P, Kaeckenbeeck A, Benz I.** 2002. DNA sequences coding for the F18 fimbriae and AIDA adhesin are localised on the same plasmid in Escherichia coli isolates from piglets. Vet. Microbiol. **86**:303–311.
25. **Nagy B, Casey TA, Whipp SC, Moon HW.** 1992. Susceptibility of porcine intestine to pilus-mediated adhesion by some isolates of piliated enterotoxigenic Escherichia coli increases with age. Infect. Immun. **60**:1285–1294.
26. **Nataro JP, Kaper JB.** 1998. Diarrheagenic Escherichia coli. Clin. Microbiol. Rev. **11**:142–201.
27. **Nataro JP, et al.** 1987. Patterns of adherence of diarrheagenic Escherichia coli to HEp-2 cells. Pediatr. Infect. Dis. J. **6**:829–831.
28. **Ngeleka M, Pritchard J, Appleyard G, Middleton DM, Fairbrother JM.** 2003. Isolation and association of Escherichia coli AIDA-I/STb, rather than EAST1 pathotype, with diarrhea in piglets and antibiotic sensitivity of isolates. J. Vet. Diagn. Invest. **15**:242–252.
29. **Niewerth U, et al.** 2001. The AIDA autotransporter system is associated with F18 and stx2e in Escherichia coli isolates from pigs diagnosed with edema disease and postweaning diarrhea. Clin. Diagn. Lab. Immunol. **8**:143–149.
30. **Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC.** 2003. A conserved region within the Bordetella pertussis autotransporter BrkA is necessary for folding of its passenger domain. Mol. Microbiol. **47**:1367–1383.
31. **Ravi M, et al.** 2007. Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic Escherichia coli and to intestinal colonization through biofilm formation in pigs. Vet. Microbiol. **120**:308–319.
32. **Renn JP, Clark PL.** 2008. A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. Biopolymers **89**:420–427.
33. **Restieri C, Garriss G, Locas MC, Dozois CM.** 2007. Autotransporter-encoding sequences are phylogenetically distributed among Escherichia coli clinical isolates and reference strains. Appl. Environ. Microbiol. **73**:1553–1562.
34. **Taillon C, Nadeau E, Mourez M, Dubreuil JD.** 2008. Heterogeneity of Escherichia coli STb enterotoxin isolated from diseased pigs. J. Med. Microbiol. **57**:887–890.
35. **van der Woude MW, Henderson IR.** 2008. Regulation and function of Ag43 (flu). Annu. Rev. Microbiol. **62**:153–169.
36. **Zhang W, Zhao M, Ruesch L, Omot A, Francis D.** 2007. Prevalence of virulence genes in Escherichia coli strains recently isolated from young pigs with diarrhea in the US. Vet. Microbiol. **123**:145–152.
37. **Zhao L, Chen X, Xu X, Song G, Liu X.** 2009. Analysis of the AIDA-I gene sequence and prevalence in Escherichia coli isolates from pigs with post-weaning diarrhoea and oedema disease. Vet. J. **180**:124–129.