Applied and Environmental Microbiology, Mar. 2007, p. 1553–1562 Vol. 73, No. 5
0099-2240/07/$08.00+0 doi:10.1128/AEM.01542-06

Autotransporter-Encoding Sequences Are Phylogenetically Distributed among *Escherichia coli* Clinical Isolates and Reference Strains▽†

Concetta Restieri, Geneviève Garriss, Marie-Claude Locas, and Charles M. Dozois*

INRS–Institut Armand-Frappier, Laval, Québec, Canada

Received 4 July 2006/Accepted 28 December 2006

Autotransporters are secreted bacterial proteins exhibiting diverse virulence functions. Various autotransporters have been identified among *Escherichia coli* associated with intestinal or extraintestinal infections; however, the specific distribution of autotransporter sequences among a diversity of *E. coli* strains has not been investigated. We have validated the use of a multiplex PCR assay to screen for the presence of autotransporter sequences. Herein, we determined the presence of 13 autotransporter sequences and five allelic variants of antigen 43 (Ag43) among 491 *E. coli* isolates from human urinary tract infections, diarrheagenic *E. coli*, and avian pathogenic *E. coli* (APEC) and *E. coli* reference strains belonging to the ECOR collection. Clinical isolates were also classified into established phylogenetic groups. The results indicated that Ag43 alleles were significantly associated with clinical isolates (93%) compared to commensal isolates (56%) and that *agn43K12* was the most common and widely distributed allele. *agn43* allelic variants were also phylogenetically distributed. Sequences encoding *espC*, *espP*, and *sepA* and *agn43* alleles *EDL933* and *RS218* were significantly associated with diarrheagenic *E. coli* strains compared to other groups. *tsh* was highly associated with APEC strains, whereas *sat* was absent from APEC. *vat*, *sat*, and *pic* were associated with urinary tract isolates and were identified predominantly in isolates belonging to either group B2 or D of the phylogenetic groups based on the ECOR strain collection. Overall, the results indicate that specific autotransporter sequences are associated with the source and/or phylogenetic background of strains and suggest that, in some cases, autotransporter gene profiles may be useful for comparative analysis of *E. coli* strains from clinical, food, and environmental sources.

Escherichia coli is a commensal resident of the intestine as well as a pathogen causing various diseases of humans and animals (7, 28). *E. coli* strains can be grouped into three major categories: diarrheagenic *E. coli*, commensal *E. coli*, and extraintestinal pathogenic *E. coli* (ExPEC) (28, 54). Diarrheagenic *E. coli* strains cause different types of diarrheal diseases and are among the most important enteric bacterial pathogens (34). Commensal *E. coli* strains are members of the normal intestinal flora of humans, other mammals, and birds (7, 54). Certain commensal intestinal isolates are capable of causing infections in extraintestinal tissues, and these pathogenic strains have been collectively termed ExPEC (55). Among ExPEC, uropathogenic *E. coli* (UPEC) is the most common cause of human urinary tract infections (UTIs) (32, 54). Avian pathogenic *E. coli* (APEC) strains share some virulence traits with ExPEC strains from human infections and are associated with extraintestinal infections of poultry, such as airsacculitis, pericarditis, cellulitis, and septicemia (7, 51, 52). The ECOR collection comprises 72 *E. coli* strains isolated from humans and other mammals from diverse geographic areas, and represents some of the clonal diversity of *E. coli* (39). Phylogenetic analyses based on the ECOR collection have classified *E. coli* into four major clonal groups: A, B1, B2, and D (22, 39). ExPEC strains commonly belong to phylogenetic group B2 and contain numerous virulence determinants implicated in extraintestinal infections (12, 47). In addition, some ExPEC strains belong to phylogenetic group D and generally contain less of the characteristic virulence factors associated with ExPEC strains belonging to group B2 (12, 47). Strains from groups B1 and A are rarely associated with extraintestinal infections, and these groups commonly contain commensal strains (12, 47). Although certain diarrheagenic *E. coli* pathotypes belong to restricted clonal groups, globally, in contrast to ExPEC, diarrheagenic *E. coli* strains are more phylogenetically dispersed and can belong to any of the *E. coli* phylogenetic groups (12).

Autotransporters are a family of secreted proteins from gram-negative bacteria that direct their own secretion across the outer membrane (this secretion mechanism has also been described as type V secretion) (11, 20). Autotransporters all possess the same general structure, comprised of three domains: an amino-terminal leader peptide (which initiates transport of the precursor across the inner membrane), an α or passenger domain (which confers the function of the secreted protein) and a carboxy-terminal domain (the β domain) that mediates secretion through the outer membrane (20). The passenger domain can either be released at the surface of the outer membrane or remain associated with the bacterial cell surface (20). The autotransporter family comprises a number of putative virulence factors that function as cytotoxins, enterotoxins, immunoglobin proteases, mucinases, heme binding proteins, and adhesins in *E. coli* and other gram-negative bacteria (11, 20). The *s*erine *p*rotease *a*uto*t*ransporters of *Enterobacte-*

* Corresponding author. Mailing address: INRS–Institut Armand-Frappier, 531 Boul. des Prairies, Laval, Québec, Canada H7V 1B7. Phone: (450) 687-5010, ext. 4221. Fax: (450) 686-5501. E-mail: charles.dozois@iaf.inrs.ca.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 12 January 2007.

riaceae (SPATEs) include adhesins, toxins, and proteases that can contribute to the virulence of different *E. coli* pathotypes (42). No SPATEs are present in *E. coli* K-12 strains, and with the exception of the nearly identical *eaaA* and *eaaC* genes (hereafter described as *eaa*) (56), SPATE-encoding genes have not been specifically identified in nonpathogenic *E. coli* (20). Currently, with the exception of *vat* and *pic* (41, 42), the specific distribution of SPATE-encoding genes among a diversity of *E. coli* isolates has not been investigated.

In addition to SPATEs, other autotransporters identified in *E. coli* include AIDA-I and antigen 43 (Ag43). AIDA-I and Ag43 are autotransporters that undergo glycosylation by heptosyl transferases, and this glycosylation affects the capacity of these autotransporters to mediate adherence to eukaryotic cells (59, 60). In addition, both AIDA-I and Ag43 contribute to bacterial intercellular aggregation and biofilm formation (6, 57, 60). AIDA-I was originally identified in diffusely adhering diarrheagenic *E. coli* strains and is also produced by some *E. coli* strains associated with edema disease and diarrhea in pigs (2, 36). Ag43 is a common phase-variable surface protein on many *E. coli* strains, and it is encoded by the *flu* (*agn43*) gene in *E. coli* K-12 (6, 8, 21). Ag43 is a highly expressed protein and can be present in numerous chromosomal copies in certain strains, and at least five allelic variants have been identified among *E. coli* strains (29, 50).

Most of the characterized autotransporters were initially identified in pathogenic *E. coli* strains, and with the exception of antigen 43, autotransporters are absent from *E. coli* K-12 strains. Despite being associated with certain pathotypes of *E. coli*, the presence of specific autotransporter sequences has not been extensively investigated among *E. coli* strains belonging to different phylogenetic groups or pathotypes. The major objectives of this study were to (i) establish a multiplex PCR assay specific for different autotransporter sequences, including 11 different SPATEs and 2 other autotransporters, Ag43 (including five allelic variants) and AIDA-1; (ii) evaluate the presence of these autotransporter sequences among *E. coli* strains from different sources or categories, including the ECOR collection, diarrheagenic *E. coli*, ExPEC, and APEC; and (iii) determine if there were differences in the incidence of autotransporter genes among *E. coli* strains based on the origins of the isolates and/or phylogenetic relationships.

MATERIALS AND METHODS

Bacterial strains and sample preparation. The ECOR collection comprises 72 strains, including 61 fecal isolates from mammals, 10 UTI isolates, and 1 asymptomatic bacteriuria isolate (39). The *E. coli* clinical isolates examined included 73 ExPEC isolates from urinary tract infections or urosepsis cases (kindly provided by James R. Johnson, Veterans Administration Medical Center, Minneapolis, MN) (24, 25). Twenty-three diarrheagenic *E. coli* strains belonging to various pathotypes (12 enterotoxigenic *E. coli* [ETEC], 6 enteropathogenic *E. coli* [EPEC], and 5 enterohemorrhagic *E. coli* [EHEC] pathotypes) were kindly provided by J. M. Fairbrother (Faculty of Veterinary Medicine, University of Montreal, Canada). The 323 APEC strains included 295 isolates from France and Canada (10) and 28 isolates from Italy (15) that were provided by M. Moulin-Schouleur (INRA, Tours, France) and L. Sperati Ruffoni (Laboratorio Tre Valli, Verona, Italy), respectively. The 295 APEC strains were previously classified into three virulence groups based on a lethality assay in 1-day-old chicks (10). Strains were stored at −80°C in 20% glycerol following overnight culture in tryptic soy broth. Strains were cultured on Luria-Bertani (LB) agar plates to isolate colonies. Overnight cultures grown at 37°C in LB broth were then used to obtain DNA samples for PCR assays. Bacterial DNA samples were obtained by alkaline lysis. Briefly, 125 μl of an overnight culture was centrifuged and the medium was removed from the pellet. The cells were suspended in 25 μl of 0.5 N NaOH and incubated at room temperature for 5 min. A 25-μl volume of 1 M Tris buffer (pH 7.5) was added to neutralize the pH, and 450 μl of water was then added to each sample. All lysates were conserved at −20°C.

The phylogenetic group (A, B1, B2, and D) of the strains was either previously known by multilocus enzyme electrophoresis (39) or determined by a triplex PCR method which uses phylogenetic group-specific primers for two genes (*chuA* and *yjaA*) and an anonymous DNA fragment (*tspE4C2*) coupled with a dichotomous decision tree (5). Samples that were positive for *chuA* and *yjaA* were classed as group B2. Samples testing positive for *chuA* and negative for *yjaA* were classed as group D. Samples testing negative for *chuA* and positive for *tspE4C2* were classed as group B1, whereas samples testing negative for *chuA* and *tspE4C2* were classed as group A.

Primer and assay design. Primers for detecting the presence of sequences encoding 11 SPATEs (Tsh, Vat, Sat, Pic [PicU], SigA, SepA, EspP [PssA], EspC, Pet, Eaa, and EatA) and 2 other *E. coli* autotransporters (Ag43 and AIDA-I) were designed and validated. In addition, we also determined the presence of sequences encoding five known variants of Ag43: K12 (*flu*), RS218, EDL933 (also called Cah [63]), CFT073a, and CFT073b. The primers, the sequences they were derived from, and positive control strains used for PCRs are presented in Table 1. Primers were designed and verified for specificity by comparative sequence analysis using sequence databases, including GenBank (http://www.ncbi.nlm.nih.gov/) and *coli*BASE, an online database for comparative genomics of *E. coli* and related enterobacteria (http://colibase.bham.ac.uk/) (4). The specificity of each of the primers and their predicted amplification products was also verified by comparative genomics/bioinformatics analysis against nucleotide sequence databases to confirm that each primer set was specific only for the given autotransporter sequence and exhibited low identity to any other autotransporter sequences. The primers were designed to have similar melting temperatures (60 to 64°C) and to amplify products of various sizes to permit the use of mixtures of primer pairs and simultaneous detection of different PCR products in multiplex reactions. A number of different alleles encoding variants of Ag43 have been described among *E. coli* strains (29). For analysis of sequences encoding five different variants of *agn43*, primer pairs specific to five distinct alleles were developed. *agn43* allele-specific primers targeted distinct sequences in the variable 5′ region of the five different alleles, whereas the generic primers for *agn43* targeted a 3′ region which is conserved among all allelic variants. The positive controls were as follows: *agn43K12* (*E. coli* MG1655), *agn43RS218* (*E. coli* RS218), *agn43aCFT073* and *agn43bCFT073* (*E. coli* CFT073), and *agn43EDL933* (*E. coli* EDL933). The specificity of each of these primer pairs was confirmed by its failure to amplify products from control strains known to possess the other *agn43* allelic variants. In addition, the strains used for detection of the five specific *agn43* alleles were used as positive controls for testing a conserved *agn43* primer pair designed to amplify all known copies of *agn43* alleles. For all primer sets, with the exception of *agn43* and *agn43K12*, *E. coli* K-12 strain MG1655 was used as the negative control. For the multiplex PCR assays, six sets of multiplex reactions using three different primer pairs each were used to screen samples. Specific multiplex sets targeting the different autotransporter sequences were as follows: M1 comprised *tsh*, *vat*, and *sat*; M2 comprised *espP*, *sigA*, and *sepA*; M3 comprised *espC*, *agn43*, and *eatA*; M4 comprised *pet*, *aida-I*, and *eaa*; M5 comprised *agn43K12*, *pic*, and *agn43EDL933*; and M6 comprised *agn43RS218*, *agn43aCFT073*, and *agn43bCFT073*. Each multiplex set was then tested, and the PCR conditions were adjusted to yield optimal simultaneous amplification of each of the target products.

Multiplex PCR conditions. Amplification reactions were done in a 50-μl volume. PCR assays were done using either 1-μl samples of purified genomic DNA resuspended in 9 μl of Tris buffer (pH 7.4) or 10-μl samples from crude bacterial cell lysates as the template DNA, 0.2 mM of deoxynucleoside triphosphates (dNTPs) (Invitrogen), various concentrations of each primer set (Table 1), and 1 U of *Taq* DNA polymerase (New England Biolabs) in 1× Thermopol reaction buffer (NEB). Reactions were performed in an automated thermal cycler (MasterCycler; Eppendorf). The reaction mixture was heated to 95°C for 1 min as a hot start. This was followed by 30 cycles of denaturing (95°C for 1 min), annealing (58°C for 1 min), and extension (72°C for 1 min, 20 s) and a final extension (72°C for 10 min). The samples were then electrophoresed in 1% agarose gels, stained with ethidium bromide, and visualized with a UV transilluminator and a Chemigenius^2 bioimaging system (Syngene).

Southern blot. Probes were synthesized with the PCR digoxigenin-labeling mix (Roche, Mannheim, Germany) in a 100-μl reaction mixture containing 1 μl of template DNA, 0.12 mM of dNTP PCR digoxigenin-labeling mix, 0.08 mM of dNTPs (Invitrogen), 0.25 μM of each primer, and 2 units of *Taq* DNA polymerase (NEB) in 1× Thermopol reaction buffer (NEB). The reactions were performed as described for the multiplex PCR conditions. Total bacterial genomic

TABLE 1. Primers for multiplex PCR and DNA probe synthesis

Gene or probe	Primer concn (μM)	Primer sequence (5′–3′)	Length of PCR product (CDS) in base pairs[a]	Accession no.[b]	Control strain	Reference(s)
Autotransporter gene						
tsh	0.25	CCGTACACAAATACGACGG GGATGCCCCTGCAGCGT	304 (4,134)	AF218073	*E. coli* χ7122	10
vat	0.25	AACGGTTGGTGGCAACAATCC AGCCCTGTAGAATGGCGAGTA	420 (4,134)	AY151282	*E. coli* CFT073	44
sat	0.5	GGTATTGATATCTCCGGTGAAC ATAGCCGCCTGACATCAGTAAT	779 (3,900)	AE014075	*E. coli* CFT073	63
espP	0.125	GTCCATGCAGGGACATGCCA TCACATCAGCACCGTTCTCTAT	547 (3,903)	NC_002128	*E. coli* EDL933	31
sigA	0.125	ACAGGGGTTACGGAGAGTGC TCCGGCAACGCCCTTAACATT	674 (3,858)	NC_004337	*S. flexneri* 2a	23
sepA	0.125	GCAGTGGAAATATGATGCGGC TTGTTCAGATCGGAGAAGAACG	794 (4,101)	Z48219	*S. flexneri* 2a	3
espC	0.75	TAGTGCAGTGCAGAAAGCAGTT AGTTTTCCTGTTGCTGTATGCC	301 (3,918)	AF297061	*E. coli* E2348/69	33
eatA	0.75	CAGGAGTGGGAACATTAAGTCA CGTACGCCTTTGATTTCAGGAT	743 (4,095)	AY163491	*E. coli* H844-7056	45
pet	0.125	GGCACAGAATAAAGGGGTGTTT CCTCTTGTTTCCACGACATAC	302 (3,888)	AF056581	*E. coli* O42	13
eaaA/eaaC	0.875	ATCCGGTGTTTTATCGTATGGG TACGGACTTCACTGAATGCCC	830 (4,008)	AF151091 AF151674	*E. coli* ECOR-9	56
pic	0.25	ACTGGATCTTAAGGCTCAGGAT GACTTAATGTCACTGTTCAGCG	572 (4,119)	AF097644	*E. coli* CFT073	19
aida-I	0.5	ACAGTATCATATGGAGCCACTC[c] TGTGCGCCAGAACTATTAATGG[c]	587 (4,080)	X65022	*E. coli* 2787	2
agn43	0.06	CTGGAAACCGGTCTGCCCTT CCTGAACGCCCAGGGTGATA	433 (3,120)	U24429	*E. coli* MG1655	21, 29
agn43K12 (flu)	0.25	CCGGCGGGCAATGGGTACA CAGCTCTCACAATCTGGCGAC	386 (3,120)	U24429	*E. coli* MG1655	21
agn43 EDL933 (cah)	0.5	CGTATCGCTGTGCCCGATAAC CCGTATACGAGTTGTCAGAATCA	707 (3,018)	AE005174	*E. coli* EDL933	46
agn43 RS218	0.375	CGGATTTCACCACCGTTAACC CATCCACCAGTGTTTTCCAGG	240 (2,847)	AJ617685	*E. coli* RS218	58
agn43aCFT073	0.25	AGGCAGGAGGAACTGCCAGT TAAATGAGGGTGTCCCGTGCC	340 (3,120)	AE014075	*E. coli* CFT073	63
agn43 bCFT073	0.25	CAGCCGGATCTGCGGCACT ACTCTGGTGTTTCTGGCTGTT	440 (3,276)	AE014075	*E. coli* CFT073	63
DNA probe						
tsh		GAAGTCTTGCGGGAACGGTCAATA TTATCCAGATATGCGCCAGCCACT	406 (4,134)	AF218073	*E. coli* χ7122	10
vat		AACGGTTGGTGGCAACAATCC AGCCCTGTAGAATGGCGAGTA	420 (4,134)	AY151282	*E. coli* CFT073	44

[a] CDS, full-length coding sequence.
[b] GenBank database accession number.
[c] Primers are modified from Ngeleka et al. (35).

DNA was prepared using a standard protocol (1). Restriction enzymes (HindIII and BamHI from Promega) were used according to the manufacturer's guidelines. Digested genomic DNA was electrophoresed in a 0.7% agarose gel. Native plasmids were isolated as described by Kado and Liu (27) and separated by pulsed-field gel electrophoresis. After electrophoresis, digested genomic DNA and plasmids were transferred to positively charged nylon membranes (Roche, Mannheim, Germany) by capillary transfer and probes were hybridized according to the Roche Applied Science DIG Application Manual for Filter Hybridization (50a). Chemiluminescent detection was performed using CDP-Star reagent (Roche, Mannheim, Germany) according to the manufacturer's guidelines. The light emission signal was captured and analyzed with a Chemigenius[2] bioimaging system (Syngene).

Statistical analyses. Statistical analyses were performed with the Prism version 4.0 software package (GraphPad Software, Inc) using a two-tailed Fisher's exact test. A cluster analysis of the isolates was performed using the between-groups linkage method (version 12.0; SPSS, Inc.).

RESULTS

Validation of multiplex PCR assays. The validation of the primer sets and optimization of the primer concentrations used for the assays were done using positive and negative amplification controls (presented in Table 1). Following optimization of the conditions for the assay, all primers used in the six different multiplex reaction sets amplified specific products of the expected sizes (Fig. 1). Each of the amplified products was sequenced, and each corresponded to the expected product. In addition, no nonspecific bands were visualized with any of the control strains used.

Distribution of autotransporters among the ECOR collection. To investigate whether the distribution of specific auto-

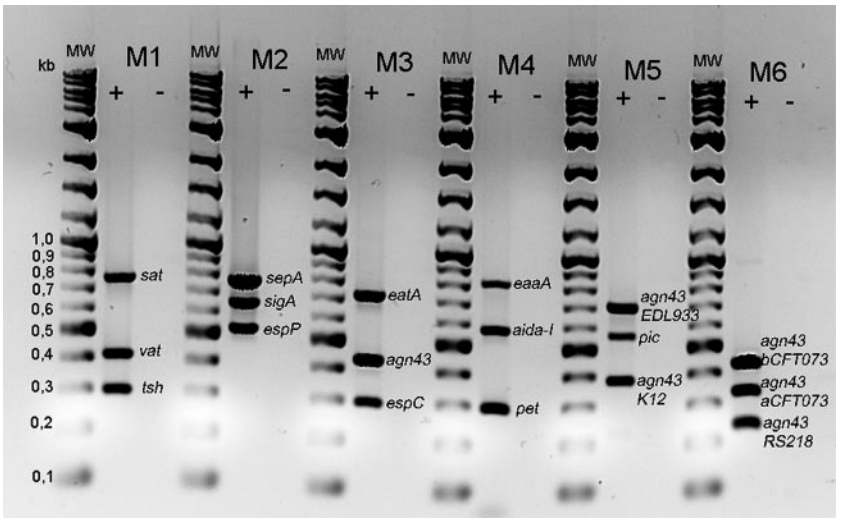

FIG. 1. Validation of PCR assay sets. Agarose gel showing PCR products amplified simultaneously from multiplex PCR assays in six separate reactions using primer pools M1 to M6, described in Materials and Methods. MW, molecular weight standard; kb, kilobases; +, positive control DNA; −, negative control DNA.

transporter sequences was associated with certain phylogenetic groups, we screened strains belonging to the ECOR collection (Fig. 2). The results indicated that certain autotransporter sequences were more common among particular clonal groups than among others. Most notably, *vat* sequences were present in all of the B2 group strains, with the exception of strain ECOR-65, and were found in only two non-B2 strains (ECOR-13 and ECOR-40). In addition, *sat* sequences were present in one cluster of the D group strains and in one cluster of the B2 group strains, but were found in only two other strains, belonging to phylogenetic group A (ECOR-11 and ECOR-24). *pic* sequences were also present mostly among one cluster of six strains belonging to phylogenetic group B2. The only other SPATE-encoding sequence that was phylogenetically distributed was *eaaA* (*eaaC*), which was limited to four strains belonging to a cluster in phylogenetic group A. None of the strains in the ECOR collection was positive for amplification of *pet*- or *aida-I*-specific sequences. All other SPATE-encoding sequences identified were absent from phylogenetic groups B2 and D, but were identified in at least one of the ECOR strains.

Ag43-encoding sequences represented the most common autotransporter sequences identified among the ECOR collection. Forty-four of the ECOR strains (61.1%) contained *agn43* sequences, as assessed by using the Ag43 primer set that amplifies all five allelic variants of *agn43*. The absence of *agn43* alleles was mostly among clusters of strains belonging to phylogenetic group A or B1 (Fig. 2). In addition, a cluster of four D group strains (ECOR-38 to ECOR-41) were *agn43* negative. The *agn43K12* (*flu*) allele was the most common *agn43* allele among the ECOR strains and was present in 37 strains dispersed among the different phylogenetic groups. The *agn43aCFT073* allele was identified in 12 strains, most of which belonged to phylogenetic group B2. The *agn43bCFT073* allele was identified in only eight strains, which all belonged to phylogenetic group B2. The *agn43RS218* allele was identified in nine strains dispersed among phylogenetic groups but was ab-

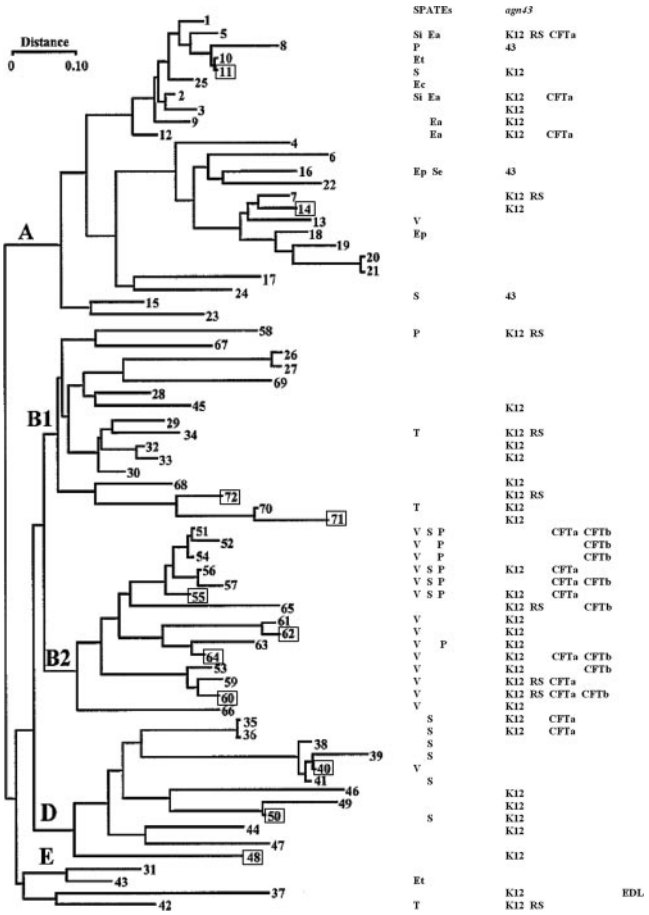

FIG. 2. Distribution of SPATEs and *agn43* and allelic variants among the ECOR collection. T, *tsh*; V, *vat*; S, *sat*; P, *pic*; Ep, *espP*; Si, *sigA*; Se, *sepA*; Ec, *espC*; Ea, *eaa*; Et, *eatA*; 43, *agn43*; K12, *agn43K12*; RS, *agn43RS218*; CFTa, *agn43aCFT073*; CFTb, *agn43bCFT073*; EDL, *agn43EDL933*. Framed strains were isolated from human urinary tract infections (ECOR-71 is from an asymptomatic bacteriuria case). Figure adapted from Herzer et al., with permission from the author (22).

TABLE 2. Distribution of most common SPATE autotransporter sequences and *agn43* allelic variants among commensal, urinary tract and avian pathogenic *E. coli* isolates categorized by phylogenetic group[a]

Source group and phylogenetic group	No. of strains in group	No. (%) of strains positive for presence of autotransporter sequence:									
		tsh[b]	*vat*[c]	*sat*[d]	*pic*[e]	*agn43*[f]	*agn43K12*[f]	*agn43RS218*	*agn43aCFT073*[g]	*agn43bCFT073*[h]	*agn43EDL933*
ECOR commensal strains											
A	23	0	1 (4)	1 (4)	1 (4)	9 (39)	6 (26)	2 (9)	3 (13)	0	0
B1	14	2 (14)	0	0	1 (4)	7 (50)	7 (50)	2 (14)	0	0	0
B2	11	0	10 (91)§	3 (27)	6 (55)*	11 (100)*	7 (64)	2 (18)	4 (36)	6 (55)¶	0
D	9	0	0	5 (56)	0	5 (56)	5 (56)	0	2 (22)	0	0
E (non)	4	1 (25)	0	0	0	2 (50)	2 (50)	1 (25)	0	0	1 (25)
Total	61	3 (5)	11 (18)	9 (15)	8 (13)	34 (56)	27 (44)	7 (11)	9 (15)	6 (10)[b]	1 (2)
UTI isolates											
A	8	2 (25)	0	1 (13)	0	5 (63)	5 (63)	0	0	0	0
B1	3	0	0	1 (33)	1 (33)	3 (100)	2 (67)	1 (33)	1 (33)	0	0
B2	37	1 (3)	34 (97)§	16 (43)	16 (43)*	37 (100)	29 (78)	7 (19)	22 (60)*	16 (43)*	0
D	36	0	1 (3)	29 (81)*	1 (3)	34 (94)	34 (94)	1 (3)	1 (3)	0	0
Total	84	3 (4)	35 (42)	47 (56)	18 (21)	79 (94)	70 (83)[e]	9 (11)	24 (29)[f]	16 (19)[g]	0
APEC isolates											
A	138	65 (47)	8 (6)	0	0	86 (62)	77 (56)	15 (11)	0	0	0
B1	30	9 (30)	0	0	1 (3)	18 (60)	12 (40)	6 (20)	0	0	1 (3)
B2	83	46 (55)	71 (86)§	0	12 (14)	61 (73)	54 (65)	15 (18)	3 (4)	2 (2)	0
D	72	40 (56)	28 (39)	0	17 (24)	49 (68)	44 (61)	7 (10)	2 (3)	2 (3)	0
Total	323	160 (50)	107 (33)	0	30 (9)	214 (66)	187 (58)	43 (13)	5 (2)	4 (1)	1 (0)

[a] For each of the source groups, association of a sequence with a single phylogenetic group compared to each of the other phylogenetic groups is indicated as follows: *, $P < 0.05$; ¶, $P < 0.005$; and §, $P < 0.0005$. The number of strains containing sequences encoding other autotransporters was too low to perform statistical analyses.
[b] *tsh* was associated with APEC isolates compared to UTI, commensal, or diarrheagenic *E. coli* strains ($P < 0.0001$).
[c] *vat* was associated with UTI isolates ($P < 0.004$) and APEC strains ($P < 0.03$) compared to commensal or diarrheagenic *E. coli* strains.
[d] *sat* was associated with UTI isolates compared to APEC, commensal, or diarrheagenic *E. coli* strains ($P < 0.0001$).
[e] *pic* was associated with UTI isolates compared to diarrheagenic *E. coli* strains ($P < 0.01$).
[f] *agn43* and *agn43K12* were associated with UTI isolates compared to APEC or commensal *E. coli* strains ($P < 0.0001$).
[g] *agn43aCFT073* was associated with UTI isolates compared to APEC or diarrheagenic *E. coli* strains ($P < 0.002$).
[h] *agn43bCFT073* was associated with UTI and commensal strains compared to APEC strains ($P < 0.002$).

sent from strains belonging to phylogenetic group D. The *agn43EDL933* allele was present only in strain ECOR-37. In addition, three strains, all of which belong to phylogenetic group A, were positive for the presence of *agn43* sequences by amplification using the Ag43 "universal" primers, but were negative for amplification using any of the allelic-variant-specific primers. Two or more *agn43* allelic variants were present in 20 ECOR strains (Fig. 2), and the presence of multiple *agn43* alleles was significantly associated with strains belonging to phylogenetic group B2 when compared to strains that belonged to other phylogenetic groups ($P = 0.0034$).

Within the ECOR collection, 11 strains are urinary tract isolates, and the remaining 61 strains are commensal fecal isolates from humans (29 strains) and other mammals (32 strains). When the distribution of different autotransporter sequences among the UTI isolates and human fecal isolates of the ECOR collection was examined, the only autotransporter sequence that was significantly associated with the UTI isolates compared to the human fecal isolates was *agn43K12* ($P < 0.01$). Specifically, 10 of 11 (91%) of the urinary tract isolates contained *agn43K12* sequences, whereas only 11 of the 29 (45%) human fecal isolates contained *agn43K12* sequences.

Distribution of autotransporter sequences among *E. coli* isolates from commensal and clinical sources. To more extensively investigate the specific presence of autotransporter genes among a diversity of *E. coli* isolates, the distributions of autotransporter sequences among commensal isolates from the ECOR collection, clinical isolates from human urinary tract infections or urosepsis cases, diarrheagenic *E. coli* strains, and APEC strains were assessed and compared. The distribution of autotransporter genes by source indicated the association of particular autotransporter genes with certain groups of clinical isolates, and none of the autotransporter sequences tested were specifically associated with commensal strains when compared to the clinical source groups comprised of clinical isolates (Tables 2 and 3).

In UTI isolates, the SPATE sequences *vat*, *sat*, and *pic* and the *agn43* sequences *agn43* (generic), *agn43K12*, *agn43aCFT073*, and *agn43bCFT073* were significantly associated (Table 2). Specifically, *vat* was associated with UTI isolates compared to commensal or diarrheagenic *E. coli* strains ($P < 0.004$), *sat* was associated with UTI isolates compared to all other groups ($P < 0.001$), and *pic* was associated with UTI isolates compared to diarrheagenic *E. coli* strains ($P < 0.01$). The most common

TABLE 3. Distribution of most common SPATE autotransporter sequences and *agn43* allelic variants among diarrheagenic *E. coli* strains categorized by pathotypes and phylogenetic group[a]

Pathotype and phylogenetic group	No. of strains in group	No. (%) of strains positive for presence of autotransporter sequence:									
		tsh	*sepA*	*espC*	*espP*	*eatA*	*agn43*	*agn43K12*	*agn43RS218*	*agn43bCFT073*	*agn43EDL933*
ETEC											
A	11	2 (18)	6 (55)	0	0	1 (9)	10 (91)	9 (82)	5 (45)	0	2 (18)
B1	1	0	0	0	0	0	1 (100)	1 (100)	0	0	0
Total	12	2 (17)	6 (50)	0	0	1 (8)	11 (92)	10 (83)	5 (42)	0	2 (17)
EPEC											
A	1	0	0	1 (100)	0	0	0	0	0	0	0
B1	4	1 (25)	1 (25)	0	0	0	4 (100)	4 (100)	3 (75)	0	1 (25)
B2	1	0	0	1 (100)	0	0	1 (100)	1 (100)	0	0	0
Total	6	1 (17)	1 (17)	2 (33)	0	0	5 (83)	5 (83)	3 (50)	0	1 (17)
EHEC											
B1	3	0	0	0	1 (33)	0	3 (100)	3 (100)	1 (33)	1 (33)	2 (67)
D	2	0	0	0	2 (100)	0	2 (100)	0	0	0	2 (100)
Total	5	0	0	0	3 (60)	0	5 (100)	3 (60)	1 (20)	1 (20)	4 (80)
Overall total	23	3 (13)	7 (30)[b]	2 (9)[c]	3 (13)[c]	1 (4)	21 (91)[d]	18 (78)	9 (39)[b]	1 (4)	7 (30)[b]

[a] Diarrheagenic isolates were negative for *vat*, *sat*, *pic*, *sigA*, *pet*, *aida-1*, *eaa*, and *agn43aCFT073* sequences. The number of strains tested was too low for statistical analyses between pathotypes or phylogenetic groups among diarrheagenic strains.
[b] Significantly associated with diarrheagenic *E. coli* compared to APEC, UPEC, or commensal strains ($P < 0.005$).
[c] Significantly associated with diarrheagenic *E. coli* compared to APEC or UPEC strains ($P < 0.05$).
[d] Significantly associated with diarrheagenic *E. coli* compared to APEC or commensal strains ($P < 0.02$).

SPATE-encoding sequence identified among the UTI isolates was *sat*, which was present in 56% of UTI isolates tested. Sequences specific to *vat* and *pic* were present in 42% and 21% of UTI isolates, respectively. *agn43* and *agn43K12* sequences were associated with UTI isolates compared to APEC or commensal *E. coli* strains ($P < 0.0001$). *agn43aCFT073* was associated with UTI isolates compared to APEC or diarrheagenic *E. coli* strains ($P < 0.002$), and *agn43bCFT073* was associated with UTI compared to APEC strains ($P < 0.002$) (Table 2).

In APEC isolates, the SPATE sequences *tsh* and *vat* were significantly associated. The *tsh*-specific sequence was significantly associated with APEC strains ($P < 0.0001$) compared to all other groups, and was present in 50% of APEC isolates (Table 2). *vat* sequences were also associated with APEC compared to commensal or diarrheagenic *E. coli* strains ($P < 0.03$). Notably, compared to UTI isolates, which commonly contained *sat* sequences, none of the APEC isolates were positive for *sat* (Table 2). Among the APEC strains tested, 295 isolates were previously classed for virulence based on lethality for 1-day-old chickens (10). The presence of autotransporter genes among APEC strains belonging to different lethality classes demonstrated that *tsh* and *vat* sequences were highly associated with high-virulence lethality class 1 (LC1) isolates. In addition, *agn43* and the specific *agn43K12* allele were associated with LC1 and moderately virulent LC2 APEC strains compared to isolates of low virulence (LC3) (Fig. 3).

In diarrheagenic *E. coli*, the SPATE sequences *sepA*, *espC*, and *espP*, generic *agn43* sequences, and the *agn43RS218* and *agn43EDL933* alleles were associated (Table 3). Specifically, *espC* and *espP* were associated with diarrheagenic *E. coli* strains compared to APEC or UTI isolates ($P < 0.05$), whereas *sepA* was associated with diarrheagenic *E. coli* strains compared to all other source groups ($P < 0.005$) (Table 3). *agn43* generic sequences were associated with diarrheagenic *E. coli* strains compared to APEC or commensal strains ($P < 0.02$),

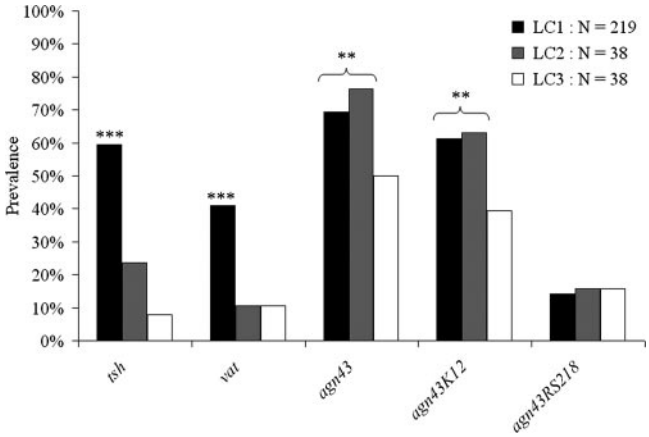

FIG. 3. Prevalence of autotransporters in 295 APEC strains previously classified for virulence based on lethality for 1-day-old chickens (10). Lethality classes are defined as follows: LC1, 50% lethal dose $< 10^8$ CFU; LC2, 50% lethal dose $\geq 10^8$ CFU; LC3, not lethal at $\geq 10^8$ CFU. The sequences encoding autotransporters were detected by multiplex PCR. Prevalence is indicated as the percentage of the 295 APEC strains associated with each LC. ***, $P < 0.0005$; **, $P < 0.005$.

whereas *agn43RS218* and *agn43EDL933* sequences were associated with diarrheagenic *E. coli* strains compared to all other source groups ($P < 0.005$). (Table 3). Among diarrheagenic *E. coli* strains, the *agn43EDL933* allele was mostly present in EHEC strains (Table 3) and was almost exclusive to diarrheagenic isolates.

Among all clinical isolates, Ag43-encoding sequences represented the most common autotransporter sequence identified. The presence of *agn43* sequences was most common among UTI isolates (94%), high among diarrheagenic *E. coli* strains (91%), and less common among APEC isolates (66%) (Tables 2 and 3). By contrast, 56% of the commensal fecal isolates contained *agn43* sequences (Table 2). The *agn43K12* (*flu*) allele was the most common allelic variant among isolates for each of the different groups of *E. coli* strains tested (Tables 2 and 3). Among clinical isolates, 25 strains, most of them classed in phylogenetic group A or B1, were positive for the presence of *agn43* sequences by amplification using the Ag43 "universal" primers but were negative for amplification using any of the allelic-variant-specific primers.

In the results for all clinical isolates tested, in addition to data presented in Tables 2 and 3, four APEC isolates were *sigA* positive, one APEC isolate was positive for the presence of *aida-I*, one diarrheagenic strain was positive for *eatA*, and no clinical isolates were positive for amplification of *pet*- or *eaaA*-specific sequences.

Phylogenetic associations of autotransporter sequences among *E. coli* strains from different sources. The analysis of strains belonging to the ECOR collection for the distribution of autotransporter sequences suggested a possible association of certain autotransporters with particular phylogenetic groups (Fig. 2). The analysis of commensal, urinary tract, and APEC isolates classified by phylogenetic group further illustrated the phylogenetic associations of some autotransporter sequences in *E. coli* isolates (Table 2). Among commensal, APEC, and UTI isolates, *vat* sequences were always highly associated with strains belonging to phylogenetic group B2. In contrast to the results for commensal and urinary tract isolates, *vat* sequences were also identified in 39% of APEC isolates that were classed as phylogenetic group D (Table 2). For the commensal strains and urinary tract isolates, *pic* sequences were more prevalent among phylogenetic group B2 (Table 2). For both commensal strains and urinary tract isolates, *sat* sequences were associated with phylogenetic group B2. By contrast, the *tsh* sequences were dispersed among different phylogenetic groups of APEC strains (Table 2). Among commensal strains, the *agn43* allele in general was associated with strains belonging to phylogenetic group B2 (Table 2). However, *agn43* was not specifically associated with a particular phylogenetic group among urinary tract or APEC isolates. The *agn43* allelic variants *agn43bCFT073* and *agn43aCFT073* were more prevalent among commensal strains and urinary tract isolates from phylogenetic group B2 (Table 2). Among APEC strains, no specific *agn43* allelic variants were associated with strains belonging to particular phylogenetic groups.

Genomic locations of *vat* and *tsh* sequences among avian pathogenic *E. coli* strains. Tsh and Vat are two closely related SPATEs that were originally identified in APEC (44, 48). These proteins share 78% identity at the amino acid level and have regions of high nucleotide identity. Thus far, *tsh* has been characterized as a plasmid-encoded SPATE in either APEC or human isolates (10, 26, 40). By contrast, *vat* is present on genomic islands in human UTI (9, 63) and APEC (44) isolates. Within the ECOR collection, which contains *E. coli* strains isolated from humans and other mammals, and among *E. coli* urinary tract isolates, *vat* sequences are associated predominantly with strains belonging to phylogenetic group B2 (Fig. 2 and Table 2) (42). However, in contrast to *E. coli* isolates from mammals, APEC strains classed as phylogenetic group D also commonly contained *vat* sequences. The localization of both *vat*- and *tsh*-specific sequences to either chromosomal or episomal locations in APEC strains was assessed to determine if there were any differences in the localization of these SPATE genes in APEC strains from different phylogenetic backgrounds. Regardless of phylogenetic group, among all APEC strains tested, *tsh*-hybridizing sequences were always associated with plasmids, whereas *vat*-hybridizing sequences never hybridized to plasmid DNA, but hybridized to total genomic DNA (see the supplemental material).

DISCUSSION

The development of a multiplex PCR assay permitted us to evaluate the presence of 13 specific autotransporter sequences in *E. coli* strains. Our results demonstrate that the presence of autotransporter sequences among *E. coli* strains is associated with certain phylogenetic groups and/or pathotypes (Tables 2 and 3). The most common autotransporter sequence we identified among the strains was *agn43*. The presence of *agn43* is thought to be well conserved in *E. coli*, and various studies have shown that some clinical isolates possess several copies of this gene (29, 50). *agn43* sequences were widely distributed among both commensal and pathogenic strains belonging to different phylogenetic groups. However, *agn43* sequences were present in only 56% of the commensal isolates from the ECOR collection (Table 2). By contrast, *agn43* sequences were present in over 90% of the urinary tract isolates as well as the diarrheagenic *E. coli* strains tested. These results suggest a higher prevalence of *agn43* sequences in pathogenic *E. coli* strains than in commensal *E. coli* strains. Of further note, pathogenic strains commonly contained multiple allelic variants of *agn43* sequences. Different *agn43* alleles confer distinct surface properties to *E. coli*, such as increased adherence to host cells or increased intercellular aggregation and biofilm production (29, 59). It may therefore be beneficial for pathogenic *E. coli* to express different variants of Ag43 to provide more-adaptive advantages for colonization and survival during infection of the host or persistence in other environments.

Since antigen 43 is present in *E. coli* K-12, and a number of allelic variants encoding this surface protein have been identified among different *E. coli* strains (29, 50), it is not surprising that *agn43* sequences are widely distributed among a diversity of *E. coli* strains. The *agn43K12* variant was the most common allele and was distributed among all phylogenetic groups. By contrast, other allelic variants were less common and some were restricted to specific phylogenetic groups. The *agn43bCFT073* allele was identified only in phylogenetic group B2 strains among the ECOR collection and UTI isolates (Table 2). Further, the *agn43EDL933* allele was almost exclusive to diarrheagenic *E. coli* strains (Table 3) and was detected only

in one APEC strain and one member of the ECOR collection (ECOR-37). Interestingly, ECOR-37, a commensal O55:H7 strain, is phylogenetically related to EHEC O157:H7 strain EDL933 (12) and it has been postulated that EHEC evolved from an O55:H7 enteropathogenic *E. coli* strain (14). Therefore, the presence of certain *agn43* allelic variants, such as *agn43EDL933* and *agn43bCFT073*, may be phylogenetically restricted and of epidemiological value for determining phylogenetic relationships among *E. coli* strains.

The presence of SPATE-encoding sequences in *E. coli* strains was also associated with certain phylogenetic groups and/or pathotypes. Interestingly, only *vat*, *sat*, and *pic* sequences, and none of the other SPATE-encoding sequences, were identified among all of the strains belonging to phylogenetic group B2 or D of the ECOR collection. In addition, the presence of *eaaA* (*eaaC*) sequences was specific to a cluster of strains from phylogenetic group A. EaaA was originally identified in human commensal fecal strain ECOR-9 (56). *eaaA* appears to be the only SPATE-encoding sequence identified that is specifically associated with human commensal isolates, as all four ECOR strains containing *eaaA* sequences were human commensal fecal isolates, and none of the tested isolates from clinical cases in humans or poultry contained *eaaA* sequences.

Our results are globally in accordance with those of Parham et al. (42). By using degenerate primers designed to target conserved regions encoding the β domains from various SPATEs, Parham et al. (42) determined that SPATE sequences were present mainly in strains belonging to phylogenetic groups B2 and D, as well as one cluster of strains from phylogenetic group A (42). By using multiplex PCR detection of specific SPATE genes, we have extended these studies to determine which specific SPATE sequences are present among strains of the ECOR collection. Some discrepancies were observed between our results and those of Parham et al. (42). By using specific primers, we identified SPATE sequences in 38 of the 72 ECOR strains compared to 29 SPATE-positive strains identified by Parham et al. (42). These differences may be due to a greater specificity of detection from use of the gene-specific primers than the degenerate primers or, in some cases, may be the result of nucleotide differences or the presence of partial or disrupted genes that would cause conflicting results.

The results globally suggest a pathotype-specific distribution of certain SPATEs among the *E. coli* strains tested. Indeed, *pic*, *vat*, and *sat* sequences have previously been shown to be associated with UTI isolates compared to commensal fecal isolates (16, 18, 42, 43). However, based on our results, the increased presence of some SPATE sequences among human UTI isolates may, at least in part, be attributed to the phylogenetic background of the strains. In general, ExPEC isolates belong to group B2 or D (30, 47), and most (87%) of the UTI isolates tested in our study belong to these phylogenetic groups (Table 2). In fact, no significant differences in the presence of *sat*, *pic*, or *vat* were found between UTI isolates belonging to group B2 or D compared with commensal isolates belonging to these same phylogenetic groups. Multivariate cluster analysis also indicated a similar combined presence of autotransporter genes among a subset of both clinical and commensal isolates that were classed as phylogenetic group B2. Among the B2 isolates, 24/37 (65%) UPEC isolates and 7/11 (63%) commensal isolates were clustered into this distinct group (data not shown). These data further support the idea that the phylogenetic origin of UTI isolates, the majority of which belong to group B2, may at least in part explain a greater prevalence of certain autotransporter genes in ExPEC isolates than in commensal isolates which belong to a more phylogenetically diverse population. In addition to being identified among UTI isolates, *pic* and *sat* are also present in certain diarrheagenic *E. coli* pathotypes or *Shigella* spp. Pic is prevalent among enteroaggregative *E. coli* (62) and in *Shigella flexneri* 2a (38). Sat, in addition to being associated with UTI isolates, is also present in some *Shigella* spp. (37, 53) and diarrheagenic *E. coli* (17, 61). Hence, *sat* and *pic* appear to be more widely distributed among commensal *E. coli* strains and different pathotypes of *E. coli* and *Shigella* than some other SPATEs. However, the presence of *sat* and *pic* sequences among commensal fecal isolates also suggests that these SPATEs are not likely to play a major role in directly eliciting enteric disease by diarrheagenic pathotypes.

A probable explanation for the phylogenetic- or pathotype-specific distribution of some of the autotransporter sequences is the inheritance or conservation of these genes on genomic islands, prophages, or conjugative plasmids. Most of the autotransporter sequences are encoded on genomic islands, including *vat* (42, 44), *pic* (19, 63), *sat* (63), *espC* (33), and *sigA* (49). Among all of the SPATE sequences, *vat* was clearly phylogenetically distributed, as nearly all strains classed as group B2, irrespective of their source, were *vat* positive. Similarly, Parham et al. also observed an almost exclusive association of *vat* with strains belonging to the ECOR B2 phylogenetic cluster (42). *vat* genes are located on pathogenicity islands that are adjacent to the tRNA *thrW* gene and may have been acquired and retained by strains that comprise phylogenetic group B2. By contrast, other genes, such as *espC*, which is present on a genomic island in a subset of EPEC (33), and *eae* sequences, which are associated with some Atlas prophages (56), may have a limited distribution, as they may have been acquired or retained by only a limited group of strains. Other autotransporters, including *tsh* and *sepA*, are encoded on large conjugative virulence plasmids (3, 10, 26, 40). Due to the increased genetic mobility of such plasmids, such sequences may have been more readily distributed among strains. This may explain, for instance, the similar distribution of *tsh* among APEC strains regardless of their phylogenetic group (Table 2). However, even plasmid-encoded genes may be largely limited to certain groups or pathotypes of strains, due to incompatability and selective pressure for the retention of other plasmids, which may contribute to virulence or fitness.

Both *tsh* and *vat* were found to be associated with APEC strains of high virulence in a day-old-chicken infection model (Fig. 3). This is in accordance with the roles of both of these SPATEs for virulence or the development of lesions in chickens (10, 44). Unlike *vat*, the presence of *tsh* sequences was limited in UTI isolates and the ECOR collection but was common among APEC strains. In addition to the greater prevalence of *tsh* among APEC strains (Table 2), it is also interesting to note the complete absence of *sat* sequences and low prevalence of *agn43CFT073* alleles among APEC strains. Hence, despite the fact that some APEC strains share many characteristics with ExPEC strains associated with human dis-

ease (51, 52), our results suggest that, in general, the APEC strains we tested also exhibit some traits that are distinct from those of *E. coli* UTI isolates. Similarly, unlike some common UTI pathotypes which contain genes coding for hemolysin (*hly*) and cytotoxic necrotizing factor (*cnf1*) (32), APEC strains nearly always lack these traits (51). It is important to note that some APEC and human UTI isolates shared the same serogroup and autotransporter gene profiles; notably, O1 or O2 strains lacked *tsh* but were *vat*$^+$, *agn43K12*$^+$, and *agn43RS218*$^+$. This is in accordance with other studies that have demonstrated a clonal relationship between APEC and human ExPEC strains. However, further characterization, such as multilocus sequence typing and more-expansive virulence gene typing, would be required to more firmly establish clonality between the APEC and UTI isolates in this study. Overall, the possibility that APEC strains are clonally related to human ExPEC and that APEC may be a transmissible source of *E. coli* strains that could cause urinary tract infection or other extraintestinal infections in humans is plausible in certain cases, but is only applicable to a limited subset of strains.

In summary, a multiplex PCR assay to screen for autotransporter genes in *E. coli* was developed and validated. With this method, we confirmed that autotransporter genes are phylogenetically and pathospecifically distributed among the 491 *E. coli* strains studied. These results are in accordance with the finding that, in *E. coli*, the presence of certain virulence genes is dependent on the genetic background and/or clinical origin of the strains (12). In the current work, the analysis of strains focused on the ECOR collection and mostly extraintestinal isolates from human UTIs or avian extraintestinal disease. It would be of further interest to examine a set of phylogenetically characterized *E. coli* isolates belonging to different pathotypes of diarrheagenic *E. coli* to further determine the distribution of autotransporter genes among these *E. coli* pathotypes.

ACKNOWLEDGMENTS

Kind thanks to J. R. Johnson, J. M. Fairbrother, L. Sperati Ruffoni, and J. P. Nataro for providing strains or plasmids. We thank S. Houle for assistance.

C.R. was funded by a Fondation Armand-Frappier scholarship. Funding for this project was provided by the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Foundation for Innovation, and the Canada Research Chairs programs.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. M. Moore, J. G. Seidman, J. A. Smith, and J. B. E. Neilands.** 1991. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, NY.
2. **Benz, I., and M. A. Schmidt.** 1992. AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesized via a precursor molecule. Mol. Microbiol. **6:**1539–1546.
3. **Boerlin, P., R. Travis, C. L. Gyles, R. Reid-Smith, N. Janecko, H. Lim, V. Nicholson, S. A. McEwen, R. Friendship, and M. Archambault.** 2005. Antimicrobial resistance and virulence genes of *Escherichia coli* isolates from swine in Ontario. Appl. Environ. Microbiol. **71:**6753–6761.
4. **Chaudhuri, R. R., A. M. Khan, and M. J. Pallen.** 2004. coliBASE: an online database for *Escherichia coli*, Shigella and Salmonella comparative genomics. Nucleic Acids Res. **32:**D296–D299.
5. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl. Environ. Microbiol. **66:**4555–4558.
6. **Danese, P. N., L. A. Pratt, S. L. Dove, and R. Kolter.** 2000. The outer membrane protein, antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. Mol. Microbiol. **37:**424–432.
7. **Dho-Moulin, M., and J. M. Fairbrother.** 1999. Avian pathogenic *Escherichia coli* (APEC). Vet. Res. **30:**299–316.
8. **Diderichsen, B.** 1980. *flu*, a metastable gene controlling surface properties of *Escherichia coli*. J. Bacteriol. **141:**858–867.
9. **Dobrindt, U., G. Blum-Oehler, T. Hartsch, G. Gottschalk, E. Z. Ron, R. Fünfstück, and J. Hacker.** 2001. S-fimbria-encoding determinant sfa_I is located on pathogenicity island III_{536} of uropathogenic *Escherichia coli* strain 536. Infect. Immun. **69:**4248–4256.
10. **Dozois, C. M., M. Dho-Moulin, A. Bree, J. M. Fairbrother, C. Desautels, and R. Curtiss III.** 2000. Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. Infect. Immun. **68:**4145–4154.
11. **Dutta, P. R., R. Cappello, F. Navarro-García, and J. P. Nataro.** 2002. Functional comparison of serine protease autotransporters of *Enterobacteriaceae*. Infect. Immun. **70:**7105–7113.
12. **Escobar-Paramo, P., O. Clermont, A. B. Blanc-Potard, H. Bui, C. Le Bouguenec, and E. Denamur.** 2004. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. Mol. Biol. Evol. **21:**1085–1094.
13. **Eslava, C., F. Navarro-Garcia, J. R. Czeczulin, I. R. Henderson, A. Cravioto, and J. P. Nataro.** 1998. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. Infect. Immun. **66:**3155–3163.
14. **Feng, P., K. A. Lampel, H. Karch, and T. S. Whittam.** 1998. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. J. Infect. Dis. **177:**1750–1753.
15. **Giovanardi, D., E. Campagnari, L. S. Ruffoni, P. Pesente, G. Ortali, and V. Furlattini.** 2005. Avian pathogenic *Escherichia coli* transmission from broiler breeders to their progeny in an integrated poultry production chain. Avian Pathol. **34:**313–318.
16. **Guyer, D. M., I. R. Henderson, J. P. Nataro, and H. L. Mobley.** 2000. Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. Mol. Microbiol. **38:**53–66.
17. **Guyer, D. M., S. Radulovic, F. E. Jones, and H. L. Mobley.** 2002. Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. Infect. Immun. **70:**4539–4546.
18. **Heimer, S. R., D. A. Rasko, C. V. Lockatell, D. E. Johnson, and H. L. Mobley.** 2004. Autotransporter genes *pic* and *tsh* are associated with *Escherichia coli* strains that cause acute pyelonephritis and are expressed during urinary tract infection. Infect. Immun. **72:**593–597.
19. **Henderson, I. R., J. Czeczulin, C. Eslava, F. Noriega, and J. P. Nataro.** 1999. Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. Infect. Immun. **67:**5587–5596.
20. **Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, and D. Ala'Aldeen.** 2004. Type V protein secretion pathway: the autotransporter story. Microbiol. Mol. Biol Rev. **68:**692–744.
21. **Henderson, I. R., and P. Owen.** 1999. The major phase-variable outer membrane protein of *Escherichia coli* structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and OxyR. J. Bacteriol. **181:**2132–2141.
22. **Herzer, P. J., S. Inouye, M. Inouye, and T. S. Whittam.** 1990. Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. J. Bacteriol. **172:**6175–6181.
23. **Jin, Q., Z. Yuan, J. Xu, Y. Wang, Y. Shen, W. Lu, J. Wang, H. Liu, J. Yang, F. Yang, X. Zhang, J. Zhang, G. Yang, H. Wu, D. Qu, J. Dong, L. Sun, Y. Xue, A. Zhao, Y. Gao, J. Zhu, B. Kan, K. Ding, S. Chen, H. Cheng, Z. Yao, B. He, R. Chen, D. Ma, B. Qiang, Y. Wen, Y. Hou, and J. Yu.** 2002. Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. Nucleic Acids Res. **30:**4432–4441.
24. **Johnson, J. R., T. A. Russo, P. I. Tarr, U. Carlino, S. S. Bilge, J. C. Vary, Jr., and A. L. Stell.** 2000. Molecular epidemiological and phylogenetic associations of two novel putative virulence genes, *iha* and $iroN_{E.\ coli}$, among *Escherichia coli* isolates from patients with urosepsis. Infect. Immun. **68:**3040–3047.
25. **Johnson, J. R., A. L. Stell, T. T. O'Bryan, M. Kuskowski, B. Nowicki, C. Johnson, J. N. Maslow, A. Kaul, J. Kavle, and G. Prats.** 2002. Global molecular epidemiology of the O15:K52:H1 extraintestinal pathogenic *Escherichia coli* clonal group: evidence of distribution beyond Europe. J. Clin. Microbiol. **40:**1913–1923.
26. **Johnson, T. J., K. E. Siek, S. J. Johnson, and L. K. Nolan.** 2006. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. J. Bacteriol. **188:**745–758.
27. **Kado, C. I., and S. T. Liu.** 1981. Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. **145:**1365–1373.
28. **Kaper, J. B., J. P. Nataro, and H. L. Mobley.** 2004. Pathogenic *Escherichia coli*. Nat. Rev. Microbiol. **2:**123–140.
29. **Klemm, P., L. Hjerrild, M. Gjermansen, and M. A. Schembri.** 2004. Structure-function analysis of the self-recognizing antigen 43 autotransporter protein from *Escherichia coli*. Mol. Microbiol. **51:**283–296.
30. **Lecointre, G., L. Rachdi, P. Darlu, and E. Denamur.** 1998. *Escherichia coli*

molecular phylogeny using the incongruence length difference test. Mol. Biol. Evol. **15:**1685–1695.

31. **Makino, K., K. Ishii, T. Yasunaga, M. Hattori, K. Yokoyama, C. H. Yutsudo, Y. Kubota, Y. Yamaichi, T. Iida, K. Yamamoto, T. Honda, C. G. Han, E. Ohtsubo, M. Kasamatsu, T. Hayashi, S. Kuhara, and H. Shinagawa.** 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. DNA Res. **5:**1–9.
32. **Marrs, C. F., L. Zhang, and B. Foxman.** 2005. *Escherichia coli* mediated urinary tract infections: are there distinct uropathogenic *E. coli* (UPEC) pathotypes? FEMS Microbiol. Lett. **252:**183–190.
33. **Mellies, J. L., F. Navarro-Garcia, I. Okeke, J. Frederickson, J. P. Nataro, and J. B. Kaper.** 2001. *espC* pathogenicity island of enteropathogenic *Escherichia coli* encodes an enterotoxin. Infect. Immun. **69:**315–324.
34. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
35. **Ngeleka, M., J. Pritchard, G. Appleyard, D. M. Middleton, and J. M. Fairbrother.** 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. Investig. **15:**242–252.
36. **Niewerth, U., A. Frey, T. Voss, C. Le Bouguenec, G. Baljer, S. Franke, and M. A. Schmidt.** 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.
37. **Niyogi, S. K., M. Vargas, and J. Vila.** 2004. Prevalence of the *sat*, *set* and *sen* genes among diverse serotypes of *Shigella flexneri* strains isolated from patients with acute diarrhoea. Clin. Microbiol. Infect. **10:**574–576.
38. **Noriega, F. R., F. M. Liao, S. B. Formal, A. Fasano, and M. M. Levine.** 1995. Prevalence of Shigella enterotoxin 1 among Shigella clinical isolates of diverse serotypes. J. Infect. Dis. **172:**1408–1410.
39. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:**690–693.
40. **Otto, B. R., S. J. van Dooren, J. H. Nuijens, J. Luirink, and B. Oudega.** 1998. Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. J. Exp. Med. **188:**1091–1103.
41. **Parham, N. J., S. J. Pollard, R. R. Chaudhuri, S. A. Beatson, M. Desvaux, M. A. Russell, J. Ruiz, A. Fivian, J. Vila, and I. R. Henderson.** 2005. Prevalence of pathogenicity island II_{CFT073} genes among extraintestinal clinical isolates of *Escherichia coli*. J. Clin. Microbiol. **43:**2425–2434.
42. **Parham, N. J., S. J. Pollard, M. Desvaux, A. Scott-Tucker, C. Liu, A. Fivian, and I. R. Henderson.** 2005. Distribution of the serine protease autotransporters of the *Enterobacteriaceae* among extraintestinal clinical isolates of *Escherichia coli*. J. Clin. Microbiol. **43:**4076–4082.
43. **Parham, N. J., U. Srinivasan, M. Desvaux, B. Foxman, C. F. Marrs, and I. R. Henderson.** 2004. PicU, a second serine protease autotransporter of uropathogenic *Escherichia coli*. FEMS Microbiol. Lett. **230:**73–83.
44. **Parreira, V. R., and C. L. Gyles.** 2003. A novel pathogenicity island integrated adjacent to the *thrW* tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. Infect. Immun. **71:**5087–5096.
45. **Patel, S. K., J. Dotson, K. P. Allen, and J. M. Fleckenstein.** 2004. Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. Infect. Immun. **72:**1786–1794.
46. **Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409:**529–533.
47. **Picard, B., J. S. Garcia, S. Gouriou, P. Duriez, N. Brahimi, E. Bingen, J. Elion, and E. Denamur.** 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. Infect. Immun. **67:**546–553.
48. **Provence, D. L., and R. Curtiss III.** 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. Infect. Immun. **62:**1369–1380.
49. **Rajakumar, K., C. Sasakawa, and B. Adler.** 1997. Use of a novel approach, termed island probing, identifies the *Shigella flexneri she* pathogenicity island which encodes a homolog of the immunoglobulin A protease-like family of proteins. Infect. Immun. **65:**4606–4614.
50. **Roche, A., J. McFadden, and P. Owen.** 2001. Antigen 43, the major phase-variable protein of the *Escherichia coli* outer membrane, can exist as a family of proteins encoded by multiple alleles. Microbiology **147:**161–169.

50a. **Roche Diagnostics.** 2000. DIG application manual for filter hybridization. Roche Molecular Biochemicals, Mannheim, Germany.

51. **Rodriguez-Siek, K. E., C. W. Giddings, C. Doetkott, T. J. Johnson, M. K. Fakhr, and L. K. Nolan.** 2005. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. Microbiology **151:**2097–2110.
52. **Ron, E. Z.** 2006. Host specificity of septicemic *Escherichia coli*: human and avian pathogens. Curr. Opin. Microbiol. **9:**28–32.
53. **Ruiz, J., M. M. Navia, J. Vila, and J. Gascon.** 2002. Prevalence of the *sat* gene among clinical isolates of *Shigella* spp. causing travelers' diarrhea: geographical and specific differences. J. Clin. Microbiol. **40:**1565–1566.
54. **Russo, T. A., and J. R. Johnson.** 2003. Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. Microbes Infect. **5:**449–456.
55. **Russo, T. A., and J. R. Johnson.** 2000. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. J. Infect. Dis. **181:**1753–1754.
56. **Sandt, C. H., and C. W. Hill.** 2000. Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of *Escherichia coli*. Infect. Immun. **68:**2205–2214.
57. **Schembri, M. A., L. Hjerrild, M. Gjermansen, and P. Klemm.** 2003. Differential expression of the *Escherichia coli* autoaggregation factor antigen 43. J. Bacteriol. **185:**2236–2242.
58. **Schneider, G., U. Dobrindt, H. Bruggemann, G. Nagy, B. Janke, G. Blum-Oehler, C. Buchrieser, G. Gottschalk, L. Emody, and J. Hacker.** 2004. The pathogenicity island-associated K15 capsule determinant exhibits a novel genetic structure and correlates with virulence in uropathogenic *Escherichia coli* strain 536. Infect. Immun. **72:**5993–6001.
59. **Sherlock, O., U. Dobrindt, J. B. Jensen, R. Munk Vejborg, and P. Klemm.** 2006. Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. J. Bacteriol. **188:**1798–1807.
60. **Sherlock, O., M. A. Schembri, A. Reisner, and P. Klemm.** 2004. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. J. Bacteriol. **186:**8058–8065.
61. **Taddei, C. R., A. C. Moreno, A. Fernandes Filho, L. P. Montemor, and M. B. Martinez.** 2003. Prevalence of secreted autotransporter toxin gene among diffusely adhering *Escherichia coli* isolated from stools of children. FEMS Microbiol. Lett. **227:**249–253.
62. **Vila, J., M. Vargas, I. R. Henderson, J. Gascon, and J. P. Nataro.** 2000. Enteroaggregative *Escherichia coli* virulence factors in traveler's diarrhea strains. J. Infect. Dis. **182:**1780–1783.
63. **Welch, R. A., V. Burland, G. Plunkett III, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. Mobley, M. S. Donnenberg, and F. R. Blattner.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. USA **99:**17020–17024.