JOURNAL OF CLINICAL MICROBIOLOGY, Dec. 2005, p. 6098–6107 Vol. 43, No. 12
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.12.6098–6107.2005

Association of Virulence Genotype with Phylogenetic Background in Comparison to Different Seropathotypes of Shiga Toxin-Producing *Escherichia coli* Isolates

Jean Pierre Girardeau,[1]* Alessandra Dalmasso,[1]† Yolande Bertin,[1] Christian Ducrot,[2] Séverine Bord,[2] Valérie Livrelli,[3] Christine Vernozy-Rozand,[4] and Christine Martin[1]

Unité de Microbiologie, Centre de Recherche INRA de Clermont-Ferrand-Theix, 63122 Saint-Genes Champanelle, France[1]*; Groupe de Recherche Pathogénie Bactérienne Intestinale, Faculté de Pharmacie, Université d'Auvergne Clermont 1, 63001 Clermont-Ferrand Cedex, France*[3]*; Unité d'Epidémiologie Animale, INRA, Centre de Recherche INRA de Clermont-Ferrand-Theix, 63122 Saint Genès Champanelle, France*[2]*; and Unité de Microbiologie Alimentaire et Prévisionnelle, Ecole Nationale Vétérinaire de Lyon, Marcy l'Etoile, France*[4]

Received 26 May 2005/Returned for modification 1 August 2005/Accepted 7 September 2005

The distribution of virulent factors (VFs) in 287 Shiga toxin-producing *Escherichia coli* (STEC) strains that were classified according to Karmali et al. into five seropathotypes (M. A. Karmali, M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper, J. Clin. Microbiol. 41:4930–4940, 2003) was investigated. The associations of VFs with phylogenetic background were assessed among the strains in comparison with the different seropathotypes. The phylogenetic analysis showed that STEC strains segregated mainly in phylogenetic group B1 (70%) and revealed the substantial prevalence (19%) of STEC belonging to phylogenetic group A (designated STEC-A). The presence of virulent clonal groups in seropathotypes that are associated with disease and their absence from seropathotypes that are not associated with disease support the concept of seropathotype classification. Although certain VFs (*eae*, $stx_{2\text{-EDL933}}$, $stx_{2\text{-vha}}$, and $stx_{2\text{-vhb}}$) were concentrated in seropathotypes associated with disease, others (*astA*, HPI, stx_{1c}, and $stx_{2\text{-NV206}}$) were concentrated in seropathotypes that are not associated with disease. Taken together with the observation that the STEC-A group was exclusively composed of strains lacking *eae* recovered from seropathotypes that are not associated with disease, the "atypical" virulence pattern suggests that STEC-A strains comprise a distinct category of STEC strains. A practical benefit of our phylogenetic analysis of STEC strains is that phylogenetic group A status appears to be highly predictive of "nonvirulent" seropathotypes.

Shiga toxin-producing *Escherichia coli* (STEC) causes a spectrum of human illness, including hemorrhagic colitis (HC) and hemolytic-uremic syndrome (HUS) (21, 24, 29, 30). STEC isolates that cause human infections belong to a large number of O:H serotypes. *E. coli* O157:H7 is the most prevalent serotype associated with large outbreaks and sporadic cases of HC and HUS in many countries (24, 29, 30). The O157:H7 strains harbor a large pathogenicity island (PAI) termed the locus for enterocyte effacement (LEE) encoding the different determinants necessary for the development of the characteristic attaching-and-effacing lesion on enterocytes (12, 20, 29). LEE-positive serotypes are commonly referred to as enterohemorrhagic *E. coli* (EHEC) (19, 22). The LEE seems to confer enhanced virulence since LEE-positive STEC are more commonly associated with outbreaks and HUS than LEE-negative serotypes (5, 19). However, LEE-negative STEC strains are also associated with severe human disease (17, 29), and some serotypes of LEE-positive STEC isolated from cattle have never been associated with human disease (50). These observations suggest that other unknown factors, possibly PAIs or genomic islands, may enhance the virulence potential of STEC strains (19, 24, 31).

Shiga toxins (Stx1 and Stx2) are the most critical virulence factors responsible for the principal manifestations of HUS and HC (11, 26, 30, 41). STEC isolates producing Stx2 are more commonly associated with severe disease (e.g., HUS) compared to isolates producing Stx1 alone or Stx1 and Stx2 (5, 27, 41, 42). Although only three stx_1 alleles were described, stx_2 comprises at least 11 distinct subtypes (8, 30). Recently, we described in bovine STEC strains a new stx_2 subtype ($stx_{2\text{-NV206}}$) showing a high cytotoxicity for Vero cells (2). In addition, several other virulence factors, including adhesins and plasmid-encoded virulence factors, contribute to the pathogenicity of STEC strains (12, 19, 24). Potential virulence genes such as the enteroaggregative *E. coli* heat-stable enterotoxin (EAST1) gene (*astA*) associated with diarrheogenic *E. coli* strains (3, 45, 51, 52) and PAIs such as the high pathogenicity island (HPI) of *Yersinia* spp. were also detected in STEC strains (2, 20).

A large variety of STEC serotypes have been implicated in disease. However, certain STEC serotypes recovered from animals and food have never been associated with severe human disease. For a better understanding of the apparent differences in virulence between groups of STEC serotypes, STEC strains were classified into five seropathotypes (A to E) by Karmali et al. (22), according to incidence and association with HUS and outbreaks. Recent studies have demonstrated that the deter-

* Corresponding author. Mailing address: Institut National de Recherches Agronomiques, Laboratoire de Microbiologie, Centre de Recherche INRA de Clermont Ferrand Theix, 63122 Saint-Genès Champanelle, France. Phone: (33) 4 73 62 42 42. Fax: (33) 4 73 63 45 81. E-mail: girard@clermont.inra.fr.

† Present address: Facoltà di Medicina Vétérinaria Università degli Studi di Torino, Turin, Italy.

mination of the seropathotype distribution of virulent factors (VFs) allows identification of DNA targets for selective detection of strains that present a risk to public health (22, 43). Such an approach has highlighted the association between the genomic island EDL933 OI-122 and seropathotypes linked to epidemic and/or severe disease (22).

To study their evolutionary relationship, different authors (7, 13, 37, 49) have studied the clonal relationship of STEC strains. Based on multilocus enzyme electrophoresis analysis and multilocus sequence typing, Whittam and coworkers have studied the clonal relationships of STEC strains (STEC Reference Center [http://www.shigatox.net/stec/index.html]). Four clonal groups have been identified: EHEC 1, EHEC 2, STEC 1, and STEC 2 (dendrograms showing these clonal groups may be viewed at the STEC Reference Center Web site).

A strategy for investigating the evolutionary origins of pathogenic *E. coli* is to determine the phylogenetic distribution of the virulence determinants (14, 33, 37). Phylogenetic analyses have shown that most *E. coli* strains belonged to four main phylogenetic groups, A, B1, B2, and D (25, 37). Whereas most commensal and diarrheogenic strains belong to groups A and B1, extraintestinal *E. coli* strains belong mainly to group B2 and group D (14). Recent phylogenetic studies indicated that STEC and EHEC strains fell into phylogenetic groups A, B1, and D (4, 13, 14). However, there was a paucity of information regarding the phylogenetic distribution of the virulence factors of STEC strains.

In the present study, the distribution of selected VFs (stx_1 and stx_2 subtypes, *eae*, *astA*, and HPI) in a collection of 287 well-defined STEC isolates was determined and analyzed. The first aim of the present study was to analyze the distribution of VFs among isolates classified by seropathotypes. The second aim was to examine the phylogenetic structure of the different seropathotypes. Finally, we sought to analyze the association between VFs and the genetic background of the strains based on classification by phylogenetic and clonal groups.

MATERIALS AND METHODS

Bacterial strains. A total of 287 STEC isolates classified into four different sets were used in the present study. The first set includes 172 bovine isolates which are part of a well-characterized bacterial collection obtained during a 1-year prospective study in the same geographic area in France (6, 36). The bovine strains belonged to 74 serotypes. The second set comprised 50 STEC strains originating from food samples collected in France (36). The third set is composed of 11 environmental isolates collected in France from dairy cattle herd manure, a wastewater treatment plant, and pig farm manure (46). The fourth set, including 54 STEC strains of diverse geographic origins isolated from human disease, was provided by the STEC Center, National Food Safety and Toxicology Center, Michigan State University. Of the 287 STEC isolates, 274 belong to non-O157 serotypes. They were found to be positive for the presence of stx_1 variants ($n = 141$), stx_2 variants ($n = 243$), or both ($n = 88$) by PCR and Southern hybridization (1, 2, 36, 46). Each strain was isolated from different patients, animals, or foods and from environmental samples. Reference strains for genotypic detection were as previously reported (2, 34, 46).

Seropathotype classification. Based on their clinical and epidemiological features, STEC strains are classified into the five seropathotypes described by Karmali et al. (22). Seropathotypes A and B (29 strains) included strains with serotypes known to be associated with HUS and outbreaks. Seropathotype C included 96 isolates associated with sporadic HUS but not with outbreaks. Seropathotype D included 70 isolates that were assigned according to the criteria of low incidence in humans and no association with HUS. The 92 isolates included in seropathotype E belong to serotypes that have not been found in humans. Assignment of O:H serotypes to seropathotype groups was based on published reports (17, 29, 50) and on three internet databases (available at http://www.microbionet.com.au/vtec2u.htm, http://www.who.int/emcdocuments/zoonoses/docs/whocsraph988.html, and http://www.lugo.usc.es/ecoli).

Nomenclature of Shiga toxin types. In the literature, different designations have been used for the same genes, causing confusion. The stx_1 abbreviation was used interchangeably to designate all stx_1-related genes or to designate the original stx_1 subtype found in the strain EDL933. Similarly, the term stx_2 was used to designate all stx_2-related genes or only the stx_2 subtype found in the EDL933 strain. In the present study, to avoid confusion, the term stx_1 variants was attributed to all stx_1-related genes and stx_2 variants refers to all stx_2-related genes. The terms stx_1 and stx_2 were reserved for the original subtypes found in the *E. coli* EDL933 reference strain. Toxins of Shiga toxin type 2d include the genetically closely related toxins ($stx_{2d\text{-}Ount}$, $stx_{2d\text{-}OX3a}$, and $stx_{2d\text{-}O111}$). Shiga toxin type 2c is reserved for the mucus-activatable toxins ($stx_{2\text{-}vha}$ and $stx_{2\text{-}vhb}$) described by Melton-Celsa et al. (28), and toxins of type 2e include the porcine edema disease-associated toxins.

Subtyping of stx_1 genes. In the present study, we developed a restriction fragment length polymorphism (RFLP)-PCR system to discriminate the three stx_1 variants (stx_1, stx_{1c}, and stx_{1d}). The three stx_1 gene variants were first amplified by using the VT1-A (ACACTGGATGATCTCAGTGG) and VT1-B (CTG AATCCCCCTCCATTATG) oligonucleotide primers at an annealing temperature of 55°C to obtain a 603-bp amplified PCR product. The amplicons were then digested with the restriction endonucleases BglI, HaeI, and RsaI as recommended by the manufacturer (Roche Applied Science). The amplified product from STEC possessing stx_1, stx_{1c}, and stx_{1d} generated fragments of 215 and 387 bp with BglI, 220 and 382 bp with RsaI, and 415 and 187 bp with HaeI, respectively. The PCR products were separately digested with each restriction endonuclease and incubated at 37°C for 4 h. Agarose gel (1.4%) electrophoresis was used to separate the restricted fragments. The restriction enzymes were selected by using the NEBcutter V2.0 program (http://tools.neb.com/NEBcutter2/index.php). To validate the RFLP-PCR procedure, the results obtained were confirmed by PCR amplification as previously described using the VT1AvarF/VT1AvarR and Lin-up/1_{OX3} primer pairs specific to the stx_{1d} and stx_{1c} variants, respectively (9, 23).

Subtyping of stx_2 genes. An RFLP-PCR system using the VT2c-VT2d primer pair was used to distinguish stx_2, $stx_{2\text{-}vha}$, $stx_{2\text{-}vhb}$, and $stx_{2\text{-}NV206}$ subtypes (2, 35, 44). The VT2 cm-VT2f primer pair was used to detect stx_{2d} (34). Using the VTea-VTeb primer pair, the detection of the stx_{2e} variant was performed as described previously (18, 48).

Detection of the genes encoding intimin and enteroaggregative heat-stable enterotoxin 1 and marker genes for the HPI of *Yersinia* spp. The intimin gene (*eae*) included in the LEE pathogenicity island was detected by PCR amplification as previously described (2, 36, 46). The *astA* gene encoding the EAST1 enteroaggregative *E. coli* heat-stable enterotoxin 1 was detected by Southern hybridization with a DNA probe amplified using the east11a/east11b primer pair (39, 51). The probe was labeled with alkali-labile DIG-dUTP (PCR DIG Probe Synthesis Kit; Roche Diagnostics). The hybridization was performed at 42°C. Chemiluminescence detection with CDP-Star (NEN) was done by exposure of membranes to Hyper film ECL (Amersham Pharmacia Biotech). PCR detection of the marker genes *irp1*, *irp2*, and *fyuA* specific to the HPI of *Yersinia* was performed as described previously (40).

Phylogenetic group determination. The main phylogenetic groups (A, B1, B2, and D) of the *E. coli* strains were determined by triplex PCR amplification as described by Clermont et al. (10).

Statistical analysis. Statistical analyses were performed with SAS for Unix Windows (version 8.01; SAS Institute, Cary, N.C.). Comparison of the prevalence for a particular characteristic in different populations was evaluated with the chi-square test and odds ratios (ORs) with 95% confidence intervals (95%CI) were determined. The threshold for statistical significance was P values of ≤ 0.01.

RESULTS

Strain characterization. Serotypes, sources of isolation, and virulence genotypes of the studied strains, sorted by seropathotype, are shown in Table 1. The 287 STEC strains included in the five reported seropathotypes belonged to 107 different O:H serotypes. All O157:H7 and O157:NM strains (13 isolates) were included in seropathotype A. The seropathotype B comprised 16 strains belonging to serotypes O26:H11, O26:NM, O103:H2, O111:NM, and O111:H2. The 96 strains in seropathotype C belonged to 24 serotypes. The best-known sero-

TABLE 1. Seropathotypes, serotypes, sources, genotypes, and phylogenetic status of the studied strains

Seropathotype	Serotype	No. of isolates	No. of isolates containing the indicated virulence determinant(s)												Phylogeny	
			eae	*astA*	HPI	stx_1 variants	stx_2 variants	stx_2	$stx_{2\text{-vha}}$	$stx_{2\text{-vhb}}$	$stx_{2\text{-NV206}}$	stx_{2d}	stx_1 variants + stx_2 variants	stx_2 + $stx_{2\text{-vhb}}$	Phylogenetic group	Clonal group
A (13)	O157:H7	12	12	1	0	8	12	9	4	0	0	0	8	0	D	EHEC1
	O157:NM	1	1	0	0	0	1	1	0	0	0	0	0	0	D	EHEC1
B (16)	O26:H11	5	5	0	5	4	2	2	0	0	0	0	2	0	B1	EHEC2
	O26:H?	1	1	0	1	1	0	0	0	0	0	0	0	0	B1	EHEC2
	O103:H2	6	6	0	0	5	1	0	1	0	0	1	1	0	B1	STEC2
	O111:NM	3	3	1	2	3	3	3	0	0	0	0	3	0	B1	EHEC2
	O111:H2	1	1	1	1	0	1	1	0	0	0	0	0	0	B1	EHEC2
C (96)	ON:NM	3	2	0	0	0	3	1	3	2	0	1	0	1	B1	EHEC-B1
	ON:H2	4	0	2	0	2	2	2	1	4	0	0	2	2	B1	STEC-B1
	OR:H16	1	0	0	0	0	1	1	0	0	0	1	0	0	B1	STEC-B1
	OR:H25	2	2	0	0	0	2	0	2	0	0	2	0	0	B1	EHEC-B1
	O5:NM	2	2	0	0	2	0	0	0	2	0	0	0	0	B1	EHEC-B1
	O8:H2	1	0	0	0	0	1	1	0	1	0	0	0	1	B1	STEC-B1
	O8:H19	3	0	0	1	3	3	3	0	1	0	0	3	1	B1	STEC-B1
	O18ac:H?	1	0	0	0	1	0	0	0	0	0	0	0	0	D	STEC-D
	O22:H8	5	0	0	0	1	4	1	1	2	1	1	1	1	B1	STEC-B1
	O84:NM	2	1	0	0	2	1	0	0	1	0	0	1	0	B1	EHEC-B1
	O91:H?	2	0	0	0	1	2	0	0	1	0	1	1	0	B1	STEC-B1
	O91:H10	10	0	0	1	0	8	3	6	5	0	0	0	2	B1	STEC-B1
	O91:H21	11	0	0	0	4	11	5	4	10	0	0	4	4	B1	STEC1
	O98:NM	1	1	0	0	1	0	0	0	0	0	0	0	0	B1	EHEC-B1
	O105:H18	2	0	0	0	2	2	2	0	2	0	0	2	2	D	STEC-D
	O105:H?	1	0	0	0	1	1	1	0	0	0	0	1	0	A	STEC-A
	O112ac:H?	1	0	0	0	0	1	0	1	0	0	0	0	0	B1	STEC-B1
	O113:H21	16	0	0	0	2	16	6	8	13	0	0	2	5	B1	STEC1
	O128ab:H2	1	0	0	1	1	1	0	0	1	0	1	1	0	B1	STEC-B1
	O157:H26	1	1	0	1	1	1	1	0	1	0	0	1	0	B1	EHEC-B1
	O163:H19	1	0	0	0	0	1	1	0	1	0	0	0	0	B1	STEC-B1
	O174:H2	13	0	0	0	13	5	0	0	0	0	0	11	0	B1	STEC-B1
	O174:H21	10	0	1	1	0	10	2	2	7	0	0	0	0	B1	STEC1
	O174:NM	2	0	0	0	0	1	0	1	0	0	0	0	0	B1	STEC-B1
D (70)	OR:NM	3	1	0	1	2	2	0	0	0	0	1	2	0	B1	EHEC-B1
	OR:H21	1	0	0	0	0	1	1	0	1	0	1	0	1	B1	STEC1
	ON:H8	3	0	0	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	ON:H21	5	0	0	0	2	4	2	2	2	0	0	2	1	B1	STEC1
	O6:H?	1	0	0	0	0	1	1	0	0	0	0	0	0	B1	STEC-B1
	O8:H9	1	0	0	1	1	0	0	0	0	0	0	0	0	A	STEC-A
	O22:NM	2	0	0	0	0	2	0	0	0	2	1	0	0	B1	STEC-B1
	O22:H16	6	0	0	1	2	6	0	0	2	1	4	2	0	A	STEC-A
	O39:H?	1	0	0	0	0	1	0	0	1	0	0	0	0	B1	STEC-B1
	O39:H8	1	0	0	0	0	1	1	0	1	0	0	0	1	B1	STEC-B1
	O49:NM	1	0	0	1	1	0	0	0	0	0	0	0	0	A	STEC-A
	O49:H?	1	1	0	0	1	1	0	1	1	0	0	0	0	B1	EHEC-B1
	O75:H8	1	0	0	1	1	1	0	0	0	1	1	1	0	B1	STEC-B1
	O76:H19	2	1	0	1	2	2	0	0	0	0	2	2	0	B1	EHEC-B1
	O77:H18	1	0	0	0	0	1	1	0	0	0	0	0	0	D	STEC-D

	O83:K1:H?	1	0	0	1	1	0	0	0	0	0	0	0	0	B2	ECOR B2
	O110:H?	2	0	2	0	2	0	0	0	0	0	0	0	0	A	STEC-A
	O112ac:H19	2	0	0	0	1	2	0	1	1	0	0	1	0	B1	STEC-B1
	O113:H4	10	0	0	1	0	10	0	1	2	7	1	0	0	A	STEC-A
	O113:H?	3	0	2	0	1	2	1	0	0	3	1	1	0	A	STEC-A
	O113:H28	1	0	1	0	1	0	0	0	0	0	0	0	0	A	STEC-A
	O116:H21	1	0	0	0	0	1	1	0	0	0	0	0	0	B1	STEC-B1
	O117:H7	1	0	0	0	1	1	0	1	0	0	0	1	0	B1	STEC-B1
	O120:H?	1	0	0	0	1	1	1	0	0	0	0	1	0	B1	STEC-B1
	O127:H?	1	0	0	0	1	0	0	0	0	0	0	0	0	A	STEC-A
	O141:NM	1	1	0	0	0	1	1	0	0	0	0	0	0	B1	EHEC-B1
	O150:NM	1	0	1	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	O166:H28	1	0	0	0	0	1	0	0	0	0	1	0	0	B2	EPEC1
	O171:H2	7	0	0	0	0	7	1	3	5	0	2	0	1	B1	STEC-B1
	O172:H?	3	0	0	0	2	3	0	3	0	0	0	2	0	B1	STEC-B1
	O174:H?	1	0	0	0	0	1	0	1	1	0	1	0	0	B1	STEC-B1
	O175:H16	3	0	0	0	0	3	0	3	0	0	0	0	0	A	STEC-A
E (92)	ON:H27	1	0	1	0	0	1	1	0	0	0	0	0	0	A	STEC-A
	ON:H38	1	0	0	0	0	1	0	0	0	0	0	0	0	A	STEC-A
	ON:H42	1	1	0	0	1	1	0	0	1	0	0	1	0	B1	EHEC-B1
	ON:K84:H19	1	0	0	0	0	1	1	0	1	0	0	0	1	B1	STEC-B1
	O1:H18	1	0	0	0	0	1	1	0	1	0	0	0	1	D	STEC-D
	O1:H20	1	0	0	0	1	0	0	0	0	0	0	0	0	D	STEC-D
	O2:H27	1	0	1	0	0	1	1	0	0	0	1	0	0	A	STEC-A
	O2:H45	1	0	1	0	1	0	0	0	0	1	0	0	0	A	STEC-A
	O3:H?	2	0	2	1	2	0	0	0	0	0	0	0	0	A	STEC-A
	O6:H?	3	0	0	3	1	3	1	0	1	3	0	1	1	A	STEC-A
	O6:H10	8	0	3	7	1	7	1	0	0	6	0	1	0	A	STEC-A
	O6:H49	1	0	0	0	1	1	1	0	0	0	0	1	0	B1	STEC-B1
	O15:H16	2	0	2	0	0	2	1	0	0	1	0	0	0	A	STEC-A
	O15:H45	2	0	1	0	1	2	1	0	1	0	0	1	1	D	STEC-D
	O15:H21	1	0	0	0	0	1	1	0	1	0	1	0	0	B1	STEC-B1
	O20:H16	1	0	1	0	0	1	0	0	0	0	0	0	0	B1	STEC-B1
	O23:H?	1	0	0	0	1	1	0	0	1	0	0	1	0	D	STEC-D
	O23:H15	2	0	0	0	2	2	0	0	2	0	0	2	0	D	STEC-D
	O25:H21	1	0	1	0	1	0	0	0	0	0	0	0	0	A	STEC-A
	O39:NM	1	0	0	0	0	1	0	0	1	0	0	0	0	B1	STEC-B1
	O46:H38	7	0	0	0	7	7	6	0	1	0	0	7	0	B1	STEC-B1
	O54:H2	2	0	0	0	2	1	0	0	0	0	1	1	0	B1	STEC-B1
	O70:H?	1	0	0	0	1	0	0	0	0	1	0	0	0	A	STEC-A
	O74:H?	2	0	1	0	1	2	1	0	1	0	0	1	1	B1	STEC-B1
	O76:H?	1	0	0	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	O87:H16	1	0	0	0	0	1	0	0	0	0	1	0	0	A	STEC-A
	O96:H19	2	0	0	0	0	2	2	0	0	0	0	0	0	B1	STEC-B1
	O102:H21	1	0	0	0	0	1	0	0	0	0	0	0	0	B1	STEC-B1
	O103:H14	1	0	0	0	1	0	0	0	0	1	0	0	0	A	STEC-A
	O109:NM	4	0	3	4	4	1	0	0	0	0	0	1	0	A	STEC-A
	O113:NM	3	0	1	0	0	3	0	0	0	3	0	0	0	A	STEC-A
	O116:H28	1	0	1	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	O117:H?	4	0	0	1	2	3	1	1	0	0	0	1	0	B1	STEC-B1
	O130:H43	1	0	0	0	1	1	1	0	1	0	0	1	1	D	STEC-D
	O132:H18	2	0	0	0	0	2	0	0	0	0	0	0	0	D	STEC-D

Continued on following page

TABLE 1—*Continued*

Seropathotype	Serotype	No. of isolates	No. of isolates containing the indicated virulence determinant(s)											Phylogeny	
			eae	*astA*	HPI	stx_1 variants	stx_2 variants	stx_2	$stx_{2\text{-vha}}$	$stx_{2\text{-vhb}}$	$stx_{2\text{-NV206}}$	stx_{2d}	stx_1 variants + stx_2 variants	stx_2 + $stx_{2\text{-vhb}}$	Phylogenetic group	Clonal group
	O136:H12	3	0	3	0	3	0	0	0	0	0	0	0	0	A	STEC-A
	O140:H32	2	0	0	0	2	0	0	0	0	0	0	0	0	A	STEC-A
	O150:H8	1	0	0	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	O159:H28	1	0	0	0	1	0	0	0	0	0	0	0	0	A	STEC-A
	O168:H8	1	0	0	0	0	1	0	0	1	0	0	0	0	B1	STEC-B1
	O172:H16	1	0	0	0	0	1	0	0	1	0	0	0	0	B1	STEC-B1
	O172:H21	6	0	0	0	5	6	0	6	0	0	0	5	0	B1	STEC-B1
	O174:H43	1	0	0	0	1	0	0	0	0	0	0	0	0	B1	STEC-B1
	O174:H49	1	0	1	0	1	1	0	0	0	0	0	1	0	B1	STEC-B1
	OX177:NM	1	1	0	0	0	1	0	1	1	0	1	0	0	B1	EHEC-B1
	OX178:H19	2	0	0	1	2	2	1	0	1	1	0	2	1	A	STEC-A
	OX178:H19	6	0	0	0	4	6	3	0	3	0	0	4	0	B1	STEC-B1

types that conform to the features of seropathotype C are O91:H21, O113:H21, and O174:H21, but serotypes O8:H19, O22:H8, O91:H10, O105:H18, and O174:H2, which were associated with HUS, were also included in seropathotype C. The 70 strains in seropathotype D belonged to 32 serotypes. The most commonly isolated serotypes included in seropathotype D are O22:H16, O113:H4, and O171:H2. The 92 strains in seropathotype E belonged to 47 serotypes. The most commonly isolated serotypes included in seropathotype E are O6:H10, O46:H38, O172:H21, and OX178:H19.

Distribution of studied virulent determinants among the different seropathotypes. The studied virulent determinants differed with respect to their distribution among the different seropathotypes (Table 2). Compared to seropathotypes that are not associated with disease (D and E, combined), seropathotypes that are associated with disease (A, B, and C, combined) exhibited a significant higher prevalence of various VFs analyzed (specifically, *eae*, stx_2, $stx_{2\text{-vha}}$ and $stx_{2\text{-vhb}}$). Inversely, *astA*, stx_{1c}, and $stx_{2\text{-NV206}}$ were significantly more prevalent among seropathotypes D and E.

Of the 287 STEC strains, only 44 (15%) had the intimin gene (*eae*) considered as a stable marker of the LEE. The prevalence of *eae* was significantly higher ($P < 0.0001$) in seropathotypes A, B, and C (linked to severe disease) than in seropathotypes D and E (no link with disease) with an OR of 11.6, revealing *eae* as a strong predictor of seropathotypes associated with severe disease.

The *astA* gene encoding the EAST1 enteroaggregative *E. coli* heat-stable enterotoxin-1 was detected in 35 isolates (12%). A substantial part (40%) of the *astA* gene was recovered from strains which possessed only the stx_1 subtype. The prevalence of *astA* was significantly higher in seropathotypes D and E (not linked to severe disease) than in seropathotypes A, B, and C (linked with disease). The HPI also exhibited a significant nonrandom distribution among the seropathotypes. However, the difference in the prevalence of HPI between seropathotypes associated or not with disease was not significant.

The differences in the prevalence of stx_1 variants or stx_2 variants (alone or combined) between seropathotypes associated or not with disease were not significant. Subtyping revealed the wide distribution of the stx_1 among the five seropathotypes. Consistent with previous studies indicating that isolates that possess stx_{1c} were recovered either from asymptomatic patients or from healthy sheep (23), stx_{1c} was more frequent in seropathotypes D and E, which are not associated with disease. The stx_{1d} subtype that was commonly recovered from sheep was not encountered in the present study.

Analysis of the stx_2 genotype showed a significant nonrandom distribution of the different subtypes among the five seropathotypes. The differences in the prevalence of stx_2, $stx_{2\text{-vha}}$, and $stx_{2\text{-vhb}}$ between seropathotypes associated with disease (A, B, and C, combined) and seropathotypes not associated with disease (D and E, combined) was significant, favoring seropathotypes that have been associated with disease. Analysis of the distribution of $stx_{2\text{-vhb}}$ in relation to the serotypes indicates that $stx_{2\text{-vhb}}$ was restricted to *eae*-negative strains. Interestingly, among these strains, the prevalence of $stx_{2\text{-vhb}}$ was more than twice higher (56% versus 22%) among isolates in seropathotype C than among isolates in seropathotypes D

TABLE 2. Seropathotype distribution of virulence factors and *stx* subtypes[a]

Virulence genotype	Total (n = 287)	No. (%) of isolates					P^a	OR (95%CI)[b]
		A (n = 13)	B (n = 16)	C (n = 96)	D (n = 70)	E (n = 92)		
eae	44 (15)	13 (100)	16 (100)	9 (9)	4 (6)	2 (2)	<0.0001	11.6 (4.6–28.1)
astA	35 (12)	1 (8)	2 (12)	3 (3)	6 (9)	23 (25)	<0.001	0.23 (0.09–0.5)
HPI	39 (13)	0	9 (56)	5 (5)	8 (12)	17 (19)		
stx_1 variants	136 (47)	8 (61)	13 (81)	36 (38)	25 (31)	54 (59)		
Only stx_1 variants	49 (17)	0	9 (56)	7 (7)	10 (14)	23 (25)		
stx_1	117 (41)	8 (61)	13 (81)	31 (32)	16 (23)	49 (53)		
stx_{1c}	18 (6)	0	0	3 (3)	7 (10)	8 (10)	0.01	0.24 (0.07–0.8)
stx_2 variants	236 (82)	13 (100)	7 (43)	87 (90)	60 (85)	69 (76)		
Only stx_2 variants	151 (53)	5 (38)	3 (19)	58 (60)	48 (68)	37 (40)		
stx_1 variatns + stx_2 variants	86 (30)	8 (61)	5 (31)	29 (30)	12 (17)	32 (35)		
stx_2	96 (33)	10 (77)	6 (37)	38 (39)	16 (23)	26 (28)	0.002	2.17 (1.2–3.6)
$stx_{2\text{-vha}}$	58 (20)	4 (31)	1 (6)	29 (30)	16 (23)	8 (9)	0.01	2.15 (1.2–3.8)
$stx_{2\text{-vhb}}$	91 (32)	0	0	54 (56)	17 (24)	20 (21)	0.0002	2.61 (1.5–4.2)
$stx_{2\text{-nv206}}$	32 (11)	0	0	1 (1)	14 (20)	17 (19)	<0.0001	0.03 (0.004–0.2)
stx_{2d}	29 (10)	0	1 (6)	7 (7)	16 (23)	5 (5)		
stx_{2e}	4 (1)	0	0	0	3 (4)	1 (1)		
$stx_2/stx_{2\text{-vhb}}$	30 (10)	0	0	19 (20)	4 (6)	7 (7)		
$stx_{2\text{-vha}}/stx_{2\text{-vhb}}$	21 (7)	0	0	16 (17)	4 (6)	1 (1)	0.002	4.61 (1.6–12.9)

[a] P values (from chi-square test) are shown only if P was <0.01.
[b] Seropathotypes A, B, and C, combined (associated with disease) are compared to seropathotypes D and E, combined (not associated with disease) and taken as the reference.

plus E (combined). This finding reveals $stx_{2\text{-vhb}}$ as a significant predictor of the "virulent" status among the *eae*-negative strains (P = 0.0002; OR = 4.2 [95%CI = 2.3 to 7.3]). Similarly, a significant higher prevalence of $stx_{2\text{-vha}}$ was found in seropathotype C. Strains of several serotypes predominantly found in seropathotype C (particularly O91:H10, O91:H21, O113:H21, and O174:H2), possessed two toxin type 2 variants ($stx_{2\text{-vhb}}/stx_2$ or $stx_{2\text{-vhb}}/stx_{2\text{-vha}}$).

The stx_{2d} subtype was infrequently detected among isolates included in this report (10%). Consistent with previous reported findings suggesting that the Stx2d-producing strains might be less pathogenic for humans (34, 53), the stx_{2d} subtype was more frequent in seropathotype D (low incidence in humans and no association with severe disease) and rare (only one isolate) in seropathotypes A and B (associated with HUS and outbreaks). The stx_{2e} subtype associated with porcine edema disease (18) was only detected on four isolates (from seropathotypes D and E).

Interestingly, the newly identified subtype $stx_{2\text{-NV206}}$ was detected among 11% of the studied strains. All but one $stx_{2\text{-NV206}}$-positive strain carried only one *stx* gene. Analysis of the distribution of $stx_{2\text{-NV206}}$ indicates that the prevalence of $stx_{2\text{-NV206}}$ differed considerably among the different seropathotypes. The difference in the prevalence of $stx_{2\text{-NV206}}$ between seropathotypes D and E (not linked with disease) and seropathotypes A, B, and C (linked to severe disease) was highly significant, revealing $stx_{2\text{-NV206}}$ as a significant predictor of seropathotypes not associated with severe disease ($P <$ 0.0001; OR = 12.1 [95%CI = 2.7 to 50.6]).

Phylogenetic distribution of the STEC strains. Phylogenetic analyses revealed that the studied STEC strains segregated mainly in phylogenetic group B1 (201 of 287 [70%]). Of the remaining strains, 53 (19%) and 29 (10%) segregated in phylogenetic groups A and D, respectively. As expected from previous studies (13, 14), phylogenetic group B2, which is predominant among extraintestinal strains, was rarely found in STEC (two isolates).

Among the phylogenetic group B1 strains, only 15% (31 of 201 isolates) were *eae* positive. All of the non-O157 strains with *eae* fell into this phylogenetic group. According to other studies (13, 16, 22), the O103:H2 strains (six isolates) were classified as STEC 2 and the strains of serotypes O26:H11, O26:NM, O111:H2, and O111:NM (ten isolates) were classified as EHEC 2. The 15 remaining *eae*-positive strains belonging to phylogenetic group B1 (of 12 different serotypes such as OR:H25, O5:NM, O49:NM, O84:NM, O98:NM, and O157:H26) were designated as EHEC-B1. Of the 170 *eae*-negative isolates of phylogenetic group B1, 41 isolates of serotypes ON:H21, O91:H21, O113:H21, and O174:H21 were classified according to Whittam and coworkers as STEC 1. The remaining 129 isolates lacking *eae* (of 57 different serotypes) were designated STEC-B1 (major serotypes were OR:H5, OR:H8, O22:H8, O22:H16, O46:H38, O74:H42, O91:H10, O171:H2, O172:H21, and O174:H2).

Of the phylogenetic group D (29 isolates), 13 strains with *eae* of serotypes O157:H7 and O157:NM were classified as EHEC 1. All of the 16 remaining group D strains lacked *eae*. These strains of 11 different serotypes (such as O1:H18, O1:H20, O15:H45, O23:H15, O77:H18, O105:H18, O130:H43, and O132:H18) were designated STEC-D.

All of the 53 isolates that belonged to the phylogenetic group A were found to lack *eae*. These strains, designated STEC-A, belonged to 25 different serotypes but only 6 accounted for 60% of the strains (O6:H10, O15:H16, O109:NM, O113:H4, O136:NM, and OX178:H19).

Phylogenetic origin in relation to seropathotypes. The different seropathotypes were clearly distinguished by the phylogenetic origin of their constituting isolates (Table 3). Most of the isolates in seropathotypes A, B, and C (associated with human disease) belonged to a major phylogenetic group

TABLE 3. Phylogenetic distribution of isolates among the different seropathotypes

Phylogenetic group	Clonal group or group of strains	No. (%) of isolates				
		A (n = 13)	B (n = 16)	C (n = 96)	D (n = 71)	E (n = 92)
B1	STEC 2	0	6 (33)	0	0	0
	EHEC 2	0	10 (67)	0	0	0
	EHEC-B1*	0	0	10 (10)	3 (4)	2 (2)
	STEC 1	0	0	36 (38)	5 (7)	0
	STEC-B1*	0	0	46 (48)	37 (53)	47 (51)
D	EHEC 1	13 (100)	0	0	0	0
	STEC-D*	0	0	3 (3)	4 (6)	9 (10)
A	STEC-A*	0	0	0	19 (28)	34 (37)
B2	B2*	0	0	0	2 (3)	0

[a] *, group of strains.

(group D for seropathotype A and group B1 for seropathotypes B and C). In contrast, isolates in seropathotypes D and E (not associated with disease) appeared to be genetically more heterogeneous. In seropathotype D, 63% of the isolates belonged to group B1, 28% belonged to group A, 6% belonged to group D, and 3% belonged to group B2. Nonhuman isolates in seropathotype E belonged either to phylogenetic group B1 (53%) or to phylogenetic group A (37%), and a small proportion (10%) belonged to phylogenetic group D. Virulent clonal groups (STEC 1, STEC 2, EHEC 1, and EHEC 2) were predominantly distributed among seropathotypes associated with disease (A, B, and C). In contrast, "atypical" groups (e.g., STEC-A and STEC-D) were concentrated within seropathotypes that are not associated with disease (D and E).

Of particular interest was the clear association observed between phylogenetic group A and the "non virulent" seropathotypes (D and E) ($P < 0.0001$; OR = 63.2 [95%CI = 8.6 to 464]). In view of this association, the phylogenetic group A status appears highly predictive of the "nonvirulent" seropathotypes.

Phylogenetic distribution of VFs. To investigate the relationships between genetic background and virulence genes, we assessed the phylogenetic distribution of the VFs. The distribution of VFs in the different phylogenetic and clonal groups is shown in Table 4. Compared to phylogenetic groups D and B1 (combined), the phylogenetic group A exhibited a significant higher prevalence of various VFs analyzed (specifically, *astA*, HPI, stx_{1c}, and $stx_{2\text{-NV206}}$). In contrast, *eae*, stx_2, $stx_{2\text{-vha}}$, and $stx_{2\text{-vhb}}$ were significantly more prevalent among phylogenetic groups D and B1 than in phylogenetic group A.

The difference in the prevalence of *eae* between phylogenetic group D and group B1 ($P = 0.004$) and group A ($P < 0.001$) was highly significant, favoring isolates in phylogenetic group D. The difference in the prevalence of HPI and *astA* genes between phylogenetic group A on the one hand and phylogenetic groups B1 and D on the other was highly signif-

TABLE 4. Phylogenetic distribution of virulence factors and *stx* subtypes

Virulence genotype	No. (%) of isolates								P[a]	OR (95%CI)[b]
	D			B1				A		
	Overall (n = 29)	EHEC 1 (n = 13)	STEC-D (n = 16)	Overall (n = 201)	STEC 1 (n = 41)	STEC-B1 (n = 129)	EHEC 2, STEC 2, EHEC-B1 (n = 31)	STEC-A (n = 53)		
eae	13 (42)	13 (100)	0	31 (15)	0	0	31 (100)	1 (2)	<0.01	0.08 (0.01–0.59)
astA	2 (6)	1 (7)	1 (5)	12 (6)	1 (2)	8 (6)	3 (10)	21 (40)	0.01	2.64 (1.4–4.9)
HPI	1 (3)	0	1 (5)	18 (9)	1 (2)	7 (5)	10 (34)	22 (41)	<0.0001	7.46 (3.6–15.3)
stx_1 variants	19 (61)	8 (61)	11 (61)	93 (46)	8 (20)	64 (47)	21 (71)	23 (42)		
Only stx_1 variants	2 (6)	0	2 (11)	28 (14)	1 (2)	12 (9)	15 (52)	18 (34)	0.001	3.04 (1.5–5.9)
stx_1	18 (58)	8 (61)	10 (56)	82 (41)	7 (17)	56 (43)	19 (61)	16 (30)		
stx_{1c}	1 (3)	0	1 (5)	7 (3)	1 (2)	6 (4)	0	7 (13)	0.01	4.08 (1.4–11.8)
stx_2 variants	29 (93)	13 (100)	16 (89)	172 (85)	39 (95)	118 (87)	15 (48)	37 (70)		
Only stx_2 variants	12 (38)	5 (38)	7 (39)	108 (54)	33 (80)	66 (51)	9 (29)	31 (58)		
stx_1 + stx_2 variants	17 (55)	8 (61)	9 (50)	65 (32)	7 (17)	52 (38)	6 (19)	5 (9)	<0.001	0.18 (0.07–044)
stx_2	16 (52)	11 (85)	5 (27)	73 (36)	14 (34)	51 (37)	8 (26)	7 (13)	0.001	0.23 (0.1–0.53)
$stx_{2\text{-vha}}$	4 (13)	4 (31)	0	53 (26)	16 (39)	31 (23)	6 (19)	2 (4)	0.001	0.12 (0.3–0.51)
$stx_{2\text{-vhb}}$	5 (16)	0	5 (29)	83 (41)	32 (78)	49 (36)	2 (6)	5 (9)	0.001	0.16 (0.6–0.42)
$stx_{2\text{-nv206}}$	0	0	0	5 (2)	0	5 (4)	0	23 (43)	<0.0001	32.6 (11.6–92)
stx_{2d}	1 (3)	0	1 (5)	24 (12)	3 (7)	15 (11)	6 (19)	4 (7)		
stx_2/$stx_{2\text{-vhb}}$	0	0	0	29 (14)	12 (29)	17 (13)	0	1 (2)		
$stx_{2\text{-vha}}$/$stx_{2\text{-vhb}}$	0	0	0	20 (10)	11 (27)	7 (5)	2 (6)	1 (2)		

[a] P values (from chi-square test) are shown only if P was <0.01.
[b] Phylogenetic group A compared to phylogenetic groups D and B1 combined and taken as the reference.

icant for HPI ($P < 0.0001$) and *astA* ($P = 0.01$), favoring phylogenetic group A (STEC-A). Consistent with vertical inheritance of HPI within lineages, all strains belonging to serotypes O6:H10, O26:H11, and O109:NM possessed the HPI determinants. In contrast, and consistent with horizontal transfer within lineages, the distribution pattern of *astA* showed variability among strains belonging to the same serotype. The substantial prevalence (35%) of HPI within the clonal group EHEC 2 was due to the constant presence of HPI in strains of serotypes O26:H11, O26:NM, O111:H2, and O111:NM.

Distribution of the *stx* genes was analyzed in relation to the phylogenetic origin of the isolate (Table 4). The difference in the prevalences of stx_1 or stx_2 variants between the different phylogenetic groups was not significant. It is well documented that isolates producing Stx1 are more rarely associated with serious disease than isolates producing Stx2 alone. Consistent with this finding, a substantial prevalence (34%) of strains possessing only the stx_1 variants was observed in phylogenetic group A. Among these strains a substantial part (28%) possess the stx_{1c} subtype which is predominantly recovered from healthy animals, asymptomatic infection or uncomplicated diarrhea in humans (53).

The distribution pattern of the different stx_2 subtypes differed considerably between the phylogenetic groups. The difference in the prevalence of stx_2 between phylogenetic groups D and B1 ($P = 0.01$) and between groups D and A ($P < 0.001$) was highly significant, favoring isolates in phylogenetic group D (concentrated into the clonal group EHEC 1). The difference in the prevalence of $stx_{2\text{-NV206}}$ between phylogenetic group A and the phylogenetic groups B1 and D (combined) was also highly significant ($P < 0.0001$; OR = 32.6 [95%CI = 11.6 to 92]), revealing a close association between $stx_{2\text{-nv206}}$ and phylogenetic group A. Indeed, among the phylogenetic group A, the predominant stx_2 genotype was the $stx_{2\text{-nv206}}$ genotype, accounting for 62% of the stx_2-positive isolates.

There was a significant difference in the prevalence of $stx_{2\text{-vha}}$ between phylogenetic groups B1 and A ($P < 0.001$) favoring isolates in phylogenetic group B1. The distribution patterns of $stx_{2\text{-vha}}$ in group B1 suggest a vertical inheritance within certain lineages (O91:H10, O91:H21, O113:H21, and O172:H21). A similar phylogenetic distribution was observed with $stx_{2\text{-vhb}}$, with a significant difference in the prevalence of $stx_{2\text{-vhb}}$ between phylogenetic group B1 and D ($P = 0.03$; OR = 2.4 [95%CI = 0.9 to 5.8]) and between groups B1 and A ($P < 0.001$; OR = 7.3 [95%CI = 2.7 to 18.4) also favoring isolates in phylogenetic group B1. The distribution patterns of $stx_{2\text{-vhb}}$ in group B1 suggest both vertical inheritance and horizontal transfer within lineages. Consistent with horizontal transfer, a strain-to-strain distribution of $stx_{2\text{-vhb}}$ was observed in serotypes O22:H16, O91:H10, and O174:H2. In contrast and consistent with vertical inheritance, $stx_{2\text{-vhb}}$ was detected in all strains belonging to serotypes O91:H21, O113:H21, and O174:H21, which are the archetypal strains of the clonal group STEC 1. The combination of stx_2 and $stx_{2\text{-vhb}}$ was restricted to isolates in phylogenetic group B1 (the most frequent serotypes are O74:H42, O91:H21, and O113:H21).

When the 170 strains lacking *eae* in the phylogenetic group B1 were stratified dichotomously as serotypes associated with disease (seropathotype C [81 isolates]) versus serotypes not associated with disease (seropathotypes D plus E [89 isolates]), the only significant difference in the prevalence of VF between the two groups of strains was observed with $stx_{2\text{-vhb}}$. The prevalence of $stx_{2\text{-vhb}}$ within group B1 was nearly twice as high (62% versus 32%) among isolates in the seropathotype C than among isolates in "nonvirulent" seropathotypes (D plus E, combined). The finding that an apparent association exists between $stx_{2\text{-vhb}}$ and seropathotype C revealed $stx_{2\text{-vhb}}$ as a significant predictor of the "virulent" status among group B1 isolates lacking *eae* ($P < 0.0001$; OR = 6.70 [95%CI = 3.6 to 12.0]). When only STEC 1 strains in seropathotype C were compared to isolates in seropathotypes D and E, $stx_{2\text{-vhb}}$ appeared to be an even stronger predictor of "virulent" status ($P = 0.0004$; OR = 8.2 [95%CI = 3.4 to 19.4]).

DISCUSSION

To our knowledge, our study is the first to provide an overview of characteristics of STEC strains classified in different seropathotypes by combining the prevalence of several known virulence determinants and their classification into the major phylogenetic groups of the *E. coli* species. Consistent with previous studies on EPEC and STEC strains (13, 14), our phylogenetic analysis shows that STEC strains segregate mainly in phylogenetic group B1 and confirms the rarity of the phylogenetic group B2. Of the remaining strains, phylogenetic groups A and D represent, respectively, 20 and 10% of the collection.

We provide here novel insights into the phylogenetic structure of the different seropathotypes. The presence in seropathotypes that are associated with disease (A, B, and C) of well-known virulent clonal groups (STEC 1, STEC 2, EHEC 1, and EHEC 2) and their concomitant replacement by "atypical" groups (e.g., STEC-A and STEC-D) in seropathotypes that are not associated with disease (D and E) support the concept of the seropathotype classification proposed by Karmali et al. (22).

From the whole data analysis, we show that there is a link between seropathotype classification, prevalence of various VFs, and phylogeny. The *astA* gene which is widely distributed among diarrheogenic *E. coli* strains in humans and animals may represent an additional virulent determinant of STEC strains (15, 38). Consistent with previous studies on human STEC infections (45), our study confirms the presence of the *astA* gene among strains carrying *eae* associated with HUS and outbreaks (O26:H11, O111:H2, and O157:H7). However, the *astA* gene was mostly (92%) recovered from strains lacking *eae* that belonged predominantly (70%) to phylogenetic group A (STEC-A). Of these strains, over half possessed only the stx_1 variants and were distributed in seropathotypes not associated with disease (D and E). Whether STEC-A strains harboring only the stx_1 variants and *astA* are true pathogens needs further elucidation.

Consistent with previous findings involving other STEC isolates (20), our data show the presence of the *Yersinia* HPI among both human and animal STEC isolates and confirm the absence of HPI from O157:H7 and O157:NM strains. In the present study a majority (61%) of the HPI-positive strains belong to seropathotypes that are not associated with disease (D and E). However, HPI was detected in all members of the clonal group EHEC 2 (O26:H11, O26:NM, O111:NM, and

O111:H2) indicating that the HPI is a common component of the genome of these strains and suggesting a high degree of stability of HPI in the genome of certain lineages. To explain the presence of HPI among STEC strains, Karch et al. (20) formed the hypothesis that HPI in STEC is a form of fitness island rather than a PAI. Similarly to the distribution of the *astA* gene, HPI was predominantly recovered from strains lacking *eae* distributed in seropathotypes D and E. Although the significance of these findings remains to be established, the present study reveals that *astA* and HPI were significantly predictive, among *eae*-negative isolates, of the "nonvirulent" seropathotypes.

There is considerable epidemiological evidence indicating that STEC isolates producing Stx2 alone are more commonly associated with serious disease than isolates producing Stx1 only or both Stx1 and Stx2 (27, 32, 41, 42). Many isolates produce two or more Stx2 variants, but the relative contribution of each variant to the pathogenesis is not known. In the present study, the 287 STEC isolates were subjected to stx_2 subtyping. The stx_2 subtype was prominent and was widely distributed among the different seropathotypes. Consistent with previous reports involving bovine non-O157 STEC (2, 8), $stx_{2\text{-vhb}}$ was also frequently detected, and the combination stx_2/$stx_{2\text{-vhb}}$ was the most frequent among STEC possessing more than one stx_2 subtype. Melton-Celsa et al. (28) demonstrated that the Stx2c variants ($Stx_{2\text{-vha}}$ and $Stx_{2\text{-vhb}}$) are activatable by intestinal mucus which causes a marked increase in toxicity and may compensate for a lack of other virulence components such as the LEE. Our data showing the confinement of $stx_{2\text{-vhb}}$ to *eae*-negative isolates support this hypothesis. In contrast to stx_2, the prevalence of $stx_{2\text{-vhb}}$ differed considerably among the different seropathotypes. Indeed, the prevalence of $stx_{2\text{-vhb}}$ was twice as high among isolates associated with HUS (seropathotype C) than among isolates that are not associated with disease (seropathotypes D plus E), revealing $stx_{2\text{-vhb}}$ as a significant predictor of "virulent" status. In addition, our phylogenetic analysis indicates that $stx_{2\text{-vhb}}$ segregates mainly within phylogenetic group B1 and is concentrated in the clonal group STEC 1. Moreover, the distribution pattern of $stx_{2\text{-vhb}}$ suggests a vertical inheritance of $stx_{2\text{-vhb}}$ in serotypes ON:H21, O91:H21, O113:H21, and O174:H21, which are the best known serotypes of the clonal group STEC 1. These findings indicate that $stx_{2\text{-vhb}}$ is a common component of the genome of certain lineages in STEC 1 and suggest an important role of the $stx_{2\text{-vhb}}$ subtype in the pathogenesis of these lineages. Consistent with this hypothesis, a major role of the Stx2-activatable variants ($Stx_{2\text{-vha}}$ and $Stx_{2\text{-vhb}}$) was recently demonstrated in the pathogenesis of the *eae*-negative strain B2F1 (O91:H21) in neonatal pigs (11). In that study, while the absence of intimin and plasmid-associated virulence determinants had little impact on the pathogenesis of STEC disease during infection, an apparent association between the severity of lesions in pigs and the production of the Stx2-activatable toxin was observed.

Our results demonstrate a close association between $stx_{2\text{-NV206}}$ and STEC-A and reveals $stx_{2\text{-NV206}}$ as a strong predictor of the "nonvirulent" seropathotypes E and D. Considering the high cytotoxic activity of the $Stx_{2\text{-NV206}}$ variant observed in vitro toward Vero cells (2), the substantial prevalence (42%) of $stx_{2\text{-NV206}}$ in STEC-A was unexpected. The observed link between $stx_{2\text{-NV206}}$ and STEC-A may result either from the preferential association of a given phage with a particular genetic background or from a coevolution of the bacteria and phage chromosomes (47). Although the role of $stx_{2\text{-NV206}}$ in pathogenesis of STEC disease remains to be established, the strong association observed between $stx_{2\text{-NV206}}$ and STEC-A supports the concept that arrival, retention, and/or expression of virulence factors require a particular genetic background (14).

One limitation of the present study results from the collection, which included multiple representatives of certain genotypes from different sources. This has resulted, for example, in a disproportionate representation of certain clones that inflate the apparent prevalence of VFs associated with these clones (for example, O113:H4, O113:H21, and O174:H2). As suggested by Johnson et al. (16), to protect against this bias, we used a more-stringent-than-usual criterion for statistical significance ($P \leq 0.01$).

One of the major findings of the present study is the demonstration of a striking phylogenetic distribution of various VFs. Specifically, when group A was compared to other phylogenetic groups combined, significant differences in prevalence favoring group A were seen for stx_{1c}, $stx_{2\text{-NV206}}$, *astA*, and HPI. Taken together with the observation that STEC-A comprises exclusively strains from seropathotypes that are not associated with disease, their "atypical" virulence pattern suggests that STEC-A isolates comprise a distinct category of STEC strains. The ability to identify STEC-A isolates is of clinical and epidemiological significance because these isolates are not commonly associated with human disease. Because of the geographically circumscribed study population, further epidemiological investigations on other STEC populations are needed to generalize these results.

REFERENCES

1. **Bertin, Y., K. Boukhors, V. Livrelli, and C. Martin.** 2004. Localization of the insertion site and pathotype determination of the locus of enterocyte effacement of Shiga toxin-producing *Escherichia coli* strains. Appl. Environ. Microbiol. **70:**61–68.
2. **Bertin, Y., K. Boukhors, N. Pradel, V. Livrelli, and C. Martin.** 2001. *stx2* subtyping of Shiga toxin-producing *Escherichia coli* isolated from cattle in France: detection of a new Stx2 subtype and correlation with additional virulence factors. J. Clin. Microbiol. **39:**3060–3065.
3. **Bertin, Y., C. Martin, J. P. Girardeau, P. Pohl, and M. Contrepois.** 1998. Association of genes encoding P fimbriae, CS31A antigen and EAST 1 toxin among CNF1-producing *Escherichia coli* strains from cattle with septicemia and diarrhea. FEMS Microbiol. Lett. **162:**235–239.
4. **Bidet, P., P. Mariani-Kurkdjian, F. Grimont, N. Brahimi, C. Courroux, P. Grimont, and E. Bingen.** 2005. Characterization of *Escherichia coli* O157:H7 isolates causing hemolytic uraemic syndrome in France. J. Med. Microbiol. **54:**71–75.
5. **Boerlin, P., S. A. McEwen, F. Boerlin-Petzold, J. B. Wilson, R. P. Johnson, and C. L. Gyles.** 1999. Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. J. Clin. Microbiol. **37:**497–503.
6. **Bonnet, R., B. Souweine, G. Gauthier, C. Rich, V. Livrelli, J. Sirot, B. Joly, and C. Forestier.** 1998. Non-O157 Stx2-producing *Escherichia coli* strains associated with sporadic cases of haemolytic-uremic syndrome in adults. J. Clin. Microbiol. **36:**1777–1780.
7. **Boyd, E. F., and D. L. Hartl.** 1998. Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. J. Bacteriol. **180:**1159–1165.
8. **Brett, K. N., M. A. Hornitzky, K. A. Bettelheim, M. J. Walker, and S. P. Djordjevic.** 2003. Bovine non-O157 Shiga toxin 2-containing *Escherichia coli* isolates commonly possess $stx_{\text{-EDL933}}$ and/or $stx_{2\text{-vhb}}$ subtypes. J. Clin. Microbiol. **41:**2716–2722.
9. **Bürk, C., R. Dietrich, G. Açar, M. Moravek, M. Bülte, and E. Märtlbauer.** 2003. Identification and characterization of a new variant of Shiga toxin 1 in *Escherichia coli* ONT:H19 of bovine origin. J. Clin. Microbiol. **41:**2106–2112.
10. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple deter-

mination of the *Escherichia coli* phylogenetic group. Appl. Environ. Microbiol. **66:**4555–4558.
11. **Dean-Nystrom, E. A., A. R. Melton-Celsa, J. F. Pohlenz, H. W. Moon, and A. D. O'Brien.** 2003. Comparative pathogenicity of *Escherichia coli* O157 and intimin-negative non-O157 Shiga toxin-producing *E. coli* strains in neonatal pigs. Infect. Immun. **71:**6526–6533.
12. **Donnenberg, M. S., S. Tzipori, M. L. McKee, A. D. O'Brien, J. Alroy, and J. B. Kaper.** 1993. The role of the *eae* gene of enterohemorrhagic *Escherichia coli* in intimate attachment in vitro and in a porcine model. J. Clin. Investig. **92:**1418–1424.
13. **Donnenberg, M. S., and T. S. Whittam.** 2001. Pathogenesis and evolution of virulence in enteropathogenic and enterohemorrhagic *Escherichia coli*. J. Clin. Investig. **107:**539–548.
14. **Escobar-Páramo, 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.
15. **Girardeau, J. P., L. Lalioui, A. M. Said, C. De Champs, and C. Le Bouguenec.** 2003. Extended virulence genotype of pathogenic *Escherichia coli* isolates carrying the *afa-8* operon: evidence of similarities between isolates from humans and animals with extraintestinal infections. J. Clin. Microbiol. **41:** 218–226.
16. **Johnson, J. R., P. Delavari, M. Kuskowski, and A. L. Stell.** 2001. Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. J. Infect. Dis. **183:**78–88.
17. **Johnson, R., R. C. Clarke, J. B. Wilson, S. C. Read, K. Rahn, S. A. Renwick, K. A. Sandhu, D. Alves, M. A. Karmali, H. Lior, S. A. McEwen, J. S. Spika, and C. L. Gyles.** 1996. Growing concerns and recent outbreaks involving non-O157:H7 serotypes of verotoxigenic *Escherichia coli*. J. Food Prot. **59:** 1112–1122.
18. **Johnson, W. M., D. R. Pollard, H. Lior, S. D. Tyler, and K. R. Rozee.** 1990. Differentiation of genes coding for *Escherichia coli* verotoxin 2 and the verotoxin associated with porcine edema disease (VTe) by the polymerase chain reaction. J. Clin. Microbiol. **28:**2351–2353.
19. **Kaper, J. B., J. L. Mellies, and J. Nataro.** 1999. Pathogenicity islands and other mobile genetic elements of diarrheagenic *Escherichia coli*, p. 33–58. *In* J. B. Kaper and J. Hacker (ed.), Pathogenicity islands and other mobile virulence elements. ASM Press, Washington, D.C.
20. **Karch, H., S. Schubert, D. Zhang, W. Zhang, H. Schmidt, T. Olschlager, and J. Hacker.** 1999. A genomic island, termed high-pathogenicity island, is present in certain non-O157 Shiga toxin-producing *Escherichia coli* clonal lineages. Infect. Immun. **67:**5994–6001.
21. **Karmali, M. A.** 1989. Infection by verocytotoxin-producing *Escherichia coli*. Clin. Microbiol. Rev. **2:**15–38.
22. **Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper.** 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. J. Clin. Microbiol. **41:**4930–4940.
23. **Koch, C., S. Hertwig, R. Lurz, B. Appel, and L. Beutin.** 2001. Isolation of a lysogenic bacteriophage carrying the $stx_{1_{OX3}}$ gene, which is closely associated with Shiga toxin-producing *Escherichia coli* strains from sheep and humans. J. Clin. Microbiol. **39:**3992–3998.
24. **Law, D.** 2000. Virulence factors of *Escherichia coli* O157 and other Shiga toxin-producing *E. coli*. J. Appl. Microbiol. **88:**729–745.
25. **Lecointre, G., J. Rachdi, P. Darlu, and E. Denamur.** 1998. *Escherichia coli* molecular phylogeny using the incongruence length difference test. Mol. Biol. Evol. **15:**1685–1695.
26. **Lindgren, S. W., A. R. Melton, and A. D. O'Brien.** 1993. Virulence of enterohemorrhagic *Escherichia coli* O91:H21 clinical isolates in an orally infected mouse model. Infect. Immun. **61:**3832–3842.
27. **Louise, C. B., and T. G. Obrig.** 1995. Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. J. Infect. Dis. **172:**1397–1401.
28. **Melton-Celsa, A. R., S. C. Darnell, and A. D. O'Brien.** 1996. Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic *Escherichia coli* O91:H21 isolates in orally infected, streptomycin-treated mice. Infect. Immun. **64:**1569–1576.
29. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
30. **Paton, J. C., and A. W. Paton.** 1998. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. Clin. Microbiol. Rev. **11:**450–479.
31. **Paton, A. W., and J. C. Paton.** 1998. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for stx_1, stx_2, *eaeA*, enterohemorrhagic *E. coli hlyA*, rfb_{O111}, and rfb_{O157}. J. Clin. Microbiol. **36:**598–602.
32. **Paton, A. W., M. C. Woodrow, R. M. Doyle, J. A. Lanser, and J. C. Paton.** 1999. Molecular characterization of a Shiga toxigenic *Escherichia coli* O113: H21 strain lacking *eae* responsible for a cluster of cases of hemolytic-uremic syndrome. J. Clin. Microbiol. **37:**3357–3361.
33. **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.
34. **Pierard, D., G. Muyldermans, L. Moriau, D. Stevens, and S. Lauwers.** 1998. Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal *Escherichia coli* isolates. J. Clin. Microbiol. **36:**3317–3322.
35. **Pollard, D. R., W. M. Johnson, H. Lior, S. D. Tyler, and K. R. Rozee.** 1990. Rapid and specific detection of verotoxin genes in *Escherichia coli* by polymerase chain reaction. J. Clin. Microbiol. **28:**540–545.
36. **Pradel, N., V. Livrelli, C. DeChamps, J.-B. Palcoux, A. Reynaud, F. Scheutz, J. Sirot, B. Joly, and C. Forestier.** 2000. Prevalence and characterization of Shiga toxin-producing *Escherichia coli* isolated from cattle, food, and children during a one-year prospective study in France. J. Clin. Microbiol. **38:**1023–1031.
37. **Reid, S. D., C. J. Herbelin, A. C. Bumbaugh, R. K. Selander, and T. S. Whittam.** 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. Nature **406:**64–67.
38. **Savarino, S. J., P. Fox, D. Yikang, and J. P. Nataro.** 1994. Identification and characterization of a gene cluster mediating enteroaggregative *Escherichia coli* aggregative adherence fimbria I biogenesis. J. Bacteriol. **156:**4949–4957.
39. **Savarino, S. J., A. Fasano, D. C. Robertson, and M. M. Levine.** 1991. Enteroaggregative *Escherichia coli* elaborate a heat-stable enterotoxin demonstrable in an in vitro rabbit intestinal model. J. Clin. Investig. **87:**1450–1455.
40. **Schubert, S., A. Rakin, H. Karch, E. Carniel, and J. Heesemann.** 1998. Prevalence of the "high-pathogenicity island" of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. Infect. Immun. **66:** 480–485.
41. **Siegler, R. L., T. G. Obrig, T. J. Pysher, V. L. Tesh, N. D. Denkers, and F. B. Taylor.** 2003. Response to Shiga toxin 1 and 2 in a baboon model of hemolytic-uremic syndrome. Pediatr. Nephrol. **18:**92–96.
42. **Tesh, V. L., J. A. Burris, J. W. Owens, V. M. Gordon, E. A. Wadolkowski, A. D. O'Brien, and J. E. Samuel.** 1993. Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. Infect. Immun. **61:**3392–3402.
43. **Toma, C., E. Martinez Espinosa, T. Song, E. Miliwebsky, I. Chinen, S. Iyoda, M. Iwanaga, and M. Rivas.** 2004. Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing *Escherichia coli*. J. Clin. Microbiol. **42:**4937–4946.
44. **Tyler, S. D., W. M. Johnson, H. Lior, G. Wang, and K. R. Rozee.** 1991. Identification of verotoxin type 2 variant B subunit genes in *Escherichia coli* by the polymerase chain reaction and restriction fragment length polymorphism analysis. J. Clin. Microbiol. **29:**1339–1343.
45. **Vaz, T. M. I., K. Irino, M. A. M. F. Kato, A. M. G. Dias, T. A. T. Gomes, M. I. C. Medeiros, M. M. M. Rocha, and B. E. C. Guth.** 2004. Virulence properties and characteristics of Shiga toxin-producing *Escherichia coli* in São Paulo, Brazil, from 1976 through 1999. J. Clin. Microbiol. **42:**903–905.
46. **Vernozy-Rozand, C. V., M. P. Montet, F. Lequerrec, E. Serillon, B. Tilly, C. Bavai, S. Ray-Gueniot, J. Bouvet, C. Mazuy-Cruchaudet, and Y. Richard.** 2002. Prevalence of verotoxin *Escherichia coli* (VTEC) in slurry, farmyard manure and sewage sludge in France. J. Appl. Microbiol. **93:**473–478.
47. **Wagner, P. L., D. W. K. Acheson, and M. K. Waldor.** 1999. Isogenic lysogens of diverse Shiga toxin 2-encoding bacteriophages produce markedly different amounts of Shiga toxin. Infect. Immun. **67:**6710–6714.
48. **Weinstein, D. L., M. P. Jackson, J. E. Samuel, R. K. Holmes, and A. D. O'Brien.** 1988. Cloning and sequencing of a Shiga-like toxin type II variant from *Escherichia coli* strain responsible for edema disease of swine. J. Bacteriol. **170:**4223–4230.
49. **Whittam, T. S., M. L. Wolfe, I. K. Wachsmuth, F. Orskov, I. Orskov, and R. A. Wilson.** 1993. Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. Infect. Immun. **61:**1619–1629.
50. **World Health Organization.** 1999. Zoonotic non-O157 Shiga toxin-producing *Escherichia coli* (STEC): report of a W.H.O. Scientific Working Group Meeting, Berlin, Germany, 22 to 23 June 1998, p. 1–30. World Health Organization, Geneva, Switzerland.
51. **Yamamoto, T., and P. Echeverria.** 1996. Detection of the enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 gene sequences in enterotoxigenic *E. coli* strains pathogenic for humans. Infect. Immun. **64:**1441–1445.
52. **Yatsuyanagi, J., S. Saito, Y. Miyajima, K. Amano, and K. Enomoto.** 2003. Characterization of atypical enteropathogenic *Escherichia coli* strains harboring the *astA* gene that were associated with a waterborne outbreak of diarrhea in Japan. J. Clin. Microbiol. **41:**2033–2039.
53. **Zweifel, C., J. E. Blanco, M. Blanco, J. Blanco, and R. Stephan.** 2004. Serotypes and virulence genes of ovine non-O157 Shiga toxin-producing *Escherichia coli* in Switzerland. Int. J. Food. Microbiol. **95:**19–27.