JCM
Journals.ASM.org

Discrimination of Enterohemorrhagic *Escherichia coli* (EHEC) from Non-EHEC Strains Based on Detection of Various Combinations of Type III Effector Genes

Sabine Delannoy,[a] Lothar Beutin,[b] Patrick Fach[a]

Anses (French Agency for Food, Environmental and Occupational Health and Safety), Food Safety Laboratory, Maisons-Alfort, France[a]; National Reference Laboratory for *Escherichia coli*, Division of Microbial Toxins, Federal Institute for Risk Assessment (BfR), Berlin, Germany[b]

Enterohemorrhagic *Escherichia coli* (EHEC) strains comprise a subgroup of Shiga-toxin (Stx)-producing *E. coli* (STEC) and are characterized by a few serotypes. Among these, seven priority STEC serotypes (O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7) are most frequently implicated in severe clinical illness worldwide. Currently, standard methods using *stx*, *eae*, and O-serogroup-specific gene sequences for detecting the top 7 EHEC serotypes bear the disadvantage that these genes can be found in non-EHEC strains as well. Here, we explored the suitability of *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* genes and combinations thereof as candidates for a more targeted EHEC screening assay. For a very large panel of *E. coli* strains ($n = 1{,}100$), which comprised EHEC ($n = 340$), enteropathogenic *E. coli* (EPEC) ($n = 392$), STEC ($n = 193$), and apathogenic strains ($n = 175$), we showed that these genetic markers were more prevalent in EHEC (67.1% to 92.4%) than in EPEC (13.3% to 45.2%), STEC (0.5% to 3.6%), and apathogenic *E. coli* strains (0 to 2.9%). It is noteworthy that 38.5% of the EPEC strains that tested positive for at least one of these genetic markers belonged to the top 7 EHEC serotypes, suggesting that such isolates might be Stx-negative derivatives of EHEC. The associations of *espK* with either *espV*, *ureD*, or *Z2098* were the best combinations for more specific and sensitive detection of the top 7 EHEC strains, allowing detection of 99.3% to 100% of these strains. In addition, detection of 93.7% of the EHEC strains belonging to other serotypes than the top 7 offers a possibility for identifying new emerging EHEC strains.

Shiga toxin-producing *E. coli* (STEC) may cause food-borne infections leading to life-threatening diseases in humans. The outcomes of STEC infections may range from asymptomatic carriage to uncomplicated diarrhea to the severe symptoms of hemorrhagic colitis (HC) and hemolytic-uremic syndrome (HUS). Owing to their human pathogenicity, some STEC strains are also designated as enterohemorrhagic *E. coli* (EHEC) (1, 2). EHEC strains comprise a subgroup of STEC and are characterized by certain serotypes, which are frequently associated with outbreaks and severe clinical illness (3, 4). Since the early 1980s, numerous cases of HC and HUS were attributed to EHEC O157:H7 (CDC report of microbiological results of raw ground beef products analyzed for *E. coli* O157:H7, http://www.fsis.usda.gov/wps/portal/fsis/topics/data-collection-and-reports/microbiology/ec/summary-data/summary-data-2011/ct_index). Consequently, public health and regulatory responses have been focused largely on this serotype. In recent years, cumulative evidence from numerous countries has indicated that up to 30 to 60% of human EHEC infections are caused by non-O157 EHEC (5, 6; U.S. Department of Agriculture [USDA] risk profile for pathogenic non-O157 Shiga-toxin producing *E. coli*, http://www.fsis.usda.gov/wps/wcm/connect/92de038d-c30e-4037-85a6-065c3a709435/Non_O157_STEC_Risk_Profile_May2012.pdf?MOD=AJPERES). Recent studies have shown that the number of non-O157 STEC infections sometimes surpasses the number of STEC O157 infections (7, 8). Accordingly, the list of STEC associated with HUS has been extended to include serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7 and their non-motile derivatives. These are the seven priority STEC serotypes most frequently implicated in outbreaks and sporadic cases of HC and HUS worldwide (9; USDA report on detection and identification of non-O157 STEC from meat products, http://www.fsis.usda.gov/wps/portal/fsis/topics/science/laboratories-and-procedures/guidebooks-and-methods/microbiology-laboratory-guidebook/microbiology-laboratory-guidebook), which are referred to as the top 7 EHEC serogroups. In the United States, regulatory testing for these STEC serogroups in meat started in June 2012. Although regulations are disparate throughout the world, many food inspection programs aim at detecting STEC strains that pose a significant threat to human health in kinds of foods that are the most likely to disseminate EHEC and to be consumed raw or undercooked. Some beef products are thus of particular interest in that aspect. The U.S. regulations consider these 7 EHEC serogroups to be adulterants in beef trim (USDA report on detection and identification of non-O157 STEC from meat products). On March 2013, a new regulation on sprouts and seeds was published in Europe that introduced for the first time in European Union legislation a microbiologic criterion for certain STEC serogroups (namely O157, O26, O103, O111, O145 and O104:H4) recognized to be those causing most cases of HC and HUS occurring in the European Union. Therefore, microbiological criteria should be considered for these six serogroups. The European Union legislation also stipulates that the possibility cannot be excluded that

Received 6 June 2013 Returned for modification 9 July 2013
Accepted 16 July 2013
Published ahead of print 24 July 2013
Address correspondence to Patrick Fach, patrick.fach@anses.fr.

doi:10.1128/JCM.01471-13

other STEC serogroups may be pathogenic to humans as well. In fact, such STEC may cause less severe forms of disease, such as diarrhea and/or bloody diarrhea, and may also cause HUS and therefore represent a hazard for the consumer's health.

Despite the publication of this legislation, detection of non-O157 EHEC remains particularly challenging because they have no phenotypical characteristics that distinguish them from the large number of non-STEC that share the same habitats. The basis of STEC virulence is not fully understood, but a greater risk of HC and HUS is associated with STEC serotypes carrying additional virulence factors to Stx, such as the intimin (encoded by the *eae* gene of the locus of enterocyte effacement [LEE]), which is directly involved in the attaching and effacing (A/E) process of the enterocyte microvilli and in bacterial colonization of the gut, thus contributing to the disease process (2). The LEE also encodes regulatory elements, a type III secretion system (TTSS), secreted effector proteins, and their cognate chaperon (10, 11). However, the LEE is not highly specific for EHEC strains and is found in Stx-negative *E. coli* strains as well. LEE is a hallmark of enteropathogenic *E. coli* (EPEC), which may cause diarrhea but not HC and HUS. Hence, current methods based on PCR detection of *stx* and *eae* genes for screening non-O157 EHEC in complex samples bear the disadvantage that a large number of non-EHEC strains are detected (12, 13, 14). Identification of genetic markers allowing a more targeted screening of EHEC strains is needed for rapidly testing food, fecal, and environmental samples.

It is likely that a variable repertoire of virulence determinants, including non-LEE-encoded effector (*nle*) genes that encode translocated substrates of the type III secretion system are present in highly pathogenic STEC and are carried by genetic mobile elements (15, 16, 17, 18). A molecular risk assessment approach based on the evaluation of the *nle* gene content has been used to predict which STEC strains pose a significant risk to human health (4, 16, 17, 18). However, the genetic targets which best support such an approach have not yet been defined. Monitoring EHEC in foods requires, in particular, selection of genetic markers that clearly discriminate EHEC from EPEC strains.

In an attempt to identify such factors, we explored the suitability of certain *nle* genes derived from the genomic O islands OI-43, OI-44, OI-50, OI-57, and OI-71 as candidates to distinguish STEC strains that constitute a severe risk for human health from EPEC and STEC strains that are not associated with severe and epidemic disease. We focused on *ureD* (urease activity) encoded by OI-43 and OI-48, *espN* (EspN), and *espK* (EspK), two genes carried by OI-50, a locus involved in the persistence of EHEC O157:H7 in the intestines of orally inoculated calves (19). Also, we focused on *espM1* (EspM1), which is derived from OI-71, and on *Z2098*, a sequence derived from OI-57, a genomic island that may be associated with increased virulence of STEC strains in humans (14, 20). Genome sequencing of EHEC strains (EHEC O157:H7, O111, O103, and O26) has also revealed other genetic markers, such as *espV*, whose role in disease has not been evaluated. This gene is located on OI-44 of EHEC O157:H7 but its prevalence in other *E. coli* pathogroups has not been documented. In this study, we evaluated the distribution of *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* in various *E. coli* pathogroups to assess their association with STEC strains with high virulence for humans and to test their suitability for clearly distinguishing EHEC from other *E. coli* pathogroups.

TABLE 1 *E. coli* gene targets for the real-time PCR array

Gene (ORF name if chromosomal)[a]	Encoded protein or family effector	Genetic support (mobile element)[a]
*ure*D (*Z1142*)	Urease-associated protein UreD	OI-43 and OI-48
espV (*Z1387*)	AvrA family effector	OI-44
espK (*Z1829*)	Leucine-rich repeats	OI-50 (prophage CP-933N)
espN (*Z1824*)	Hypothetical protein	OI-50 (prophage CP-933N)
(*Z2098*)	Hypothetical protein	OI-57
espM1 (*Z2565*)	Non-LEE-encoded type III effector	OI-71

[a] Nomenclature of ORFs and mobile elements refers to the sequence of *E. coli* O157:H7 EDL933 (GenBank accession no. AE005174).

MATERIALS AND METHODS

Bacterial strains. The origin and properties of the *E. coli* strains (n = 1,100) used in this study were previously described (13, 14, 21). The EHEC type strains (n = 340) were defined by the presence of both *stx* and *eae* genes. STEC strains (n = 193) harbor *stx* only. EPEC strains (n = 392) harbor *eae* only. Apathogenic *E. coli* (n = 175) were defined as *stx*- and *eae*-negative strains. Cultivation of bacteria and preparation of DNA were performed as previously described (13, 14, 21).

High-throughput real-time PCR. *E. coli* gene targets used for the real-time PCR amplification are reported in Table 1. The genes stx_1, stx_2, and *eae* were used as internal controls and for group assignment purposes. Primers and probes for the detection of stx_1, stx_2, *eae*, *espK*, and *Z2098* were described previously (14, 22), whereas those for the detection of *ureD*, *espV*, *espN*, and *espM1* were designed for this work. Primers and probes are listed in Table 2. A LightCycler 1536 (Roche, Meylan, France) was used to perform high-throughput real-time PCR amplifications as described previously (13, 14, 21).

RESULTS

Distribution of *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* and combinations thereof among *E. coli* pathogroups. Distribution of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* among the different *E. coli* pathogroups is shown in Table 3. Overall, the genetic markers investigated were mostly detected in EHEC strains with frequencies ranging from 67.1% (*espM1*) to 92.4% (*espK*). These markers were less commonly associated with EPEC strains with frequencies ranging from 13.3% (*espN*) to 45.2% (*espV*) and rarely detected in STEC (0.5 to 3.6%) and non-pathogenic *E. coli* (0 to 2.9%). Overall, we observed that 38.5% of the EPEC strains which tested positive for at least one of the investigated genetic markers belonged to the top 7 EHEC serotypes (Table 4). Interestingly, among the EPEC strains positive for at least one of the investigated genetic markers were EPEC cluster 1 strains (45.25%) that were previously reported to be frequently involved in severe illness and outbreaks (22). These included in particular EPEC serotypes that have already been associated with *stx* production, such as O100:H25, O103:H25, O111:H2, O119:H25, O126:H27, O128:H2, O55:H7, O70:H11, O76:H7, O118:H8, O119:H8, O156:H8, O2:H40, O3:H8, O80:H2, O86:H11, O66:H8, O111:H25, O114:H2, and O119:H2 (data not shown).

None of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, or *espM1* is, by itself, suitable for reliable identification of all EHEC strains. Combinations of the genetic markers were explored to identify those which detect EHEC with best specificity. The results are presented in Table 3. The combined presence of these genetic markers was highly associated with EHEC, with frequencies ranging from 93.8% (*espN*/*espM1*) to 98.8% (*espK*/*ureD*). The same combinations detected EPEC strains with frequencies ranging from 26.5% (*espN*/*Z2098*) to 54% (*espK*/*espV*), STEC strains with

TABLE 3 Distribution of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* and combinations thereof among *E. coli* pathogroups

Genetic markers and combinations thereof[a]	Distribution (%) within each pathogroup[b]			
	EHEC	EPEC	STEC	EC
Z2098	87.4	23.2	3.6	1.1
espK	92.4	28.8	0.5	1.1
espN	86.5	13.3	0.5	0.0
espV	84.4	45.2	1.6	0.6
ureD	89.4	18.1	3.1	2.9
espM1	67.1	34.4	0.5	1.7
espK/*espV*	98.5	54.1	1.6	1.1
espK/*ureD*	98.8	33.4	3.6	3.4
espK/*Z2098*	97.9	36.7	3.6	2.3
espN/*Z2098*	96.5	26.5	3.6	1.1
espN/*espM1*	93.8	38.3	1.0	1.7

[a] *espK*/*espV*, strains giving a positive result for *espK* and/or *espV*; *espK*/*ureD*, strains giving a positive result for *espK* and/or *ureD*; *espK*/*Z2098*, strains giving a positive result for *Z2098* and/or *espK*; *espN*/*Z2098*, strains giving a positive result for *espN* and/or *Z2098*; *espN*/*espM1*, strains giving a positive result for *espN* and/or *espM1*.
[b] *E. coli* pathogroups are defined in Materials and Methods.

frequencies of 1% to 3.6%, and nonpathogenic *E. coli* strains with frequencies between 1.1% and 3.4%.

Distribution of *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* and combinations thereof among EHEC serotypes. The distribution of each genetic marker, *ureD*, *espV*, *espK*, *espN*, *Z2098*, or *espM1*, was different according to EHEC serotypes. Distribution of each genetic marker in various EHEC serotypes is reported in Table 5. Except for *espV*, which was not detected in any EHEC O45:H2, and *espM1*, which was absent from EHEC O45:H2 and from 48 out of the 49 O103:H2 strains tested, all the other genetic markers investigated were found to be highly prevalent in EHEC strains of the top 7 serotypes, with frequencies ranging from 71.4% (prevalence of *ureD* in O103:H2) to 100%.

Detection of the top 7 EHEC serotypes based on different combinations of these genetic markers is reported in Table 6. With the combined detection of *espN* with *espM1*, all EHEC O157:H7, O26:H11, O111:H8, O121:H19, O45:H2, and O145:H28 strains gave a positive result for *espN* and/or *espM1*, and 95.9% of O103:H2 strains tested positive. Likewise, detection of *espK* and/or *Z2098* allowed detection of most of the EHEC serotypes associated with human infections. Thus, all EHEC O111:H8, O26:H11, O45:H2, O103:H2, and O145:H28 strains, 98.5% of O157:H7 strains, and 95% of O121:H19 strains gave a positive result for *espK* and/or *Z2098*. With the combined detection of *espN* with *Z2098* all EHEC O26:H11, O103:H2, O145:H28, O45:H2, and O111:H8 strains, 98.5% of O157:H7 strains, and 85% (17/20) of O121:H19 strains tested positive for *espN* and/or *Z2098*. The association of *espK* with either *espV* or *ureD* allowed detection of most of the strains of the top 7 EHEC serotypes as well. Hence, all strains of serotypes O157:H7, O145:H28, O111:H8, O103:H2, O45:H2, and O121:H19 gave positive results for *espK* and/or *espV*, and 97.7% of O26:H11 strains gave positive result for *espK* and/or *espV*. Data were very similar when we tested *espK* in association with *ureD*. In this case, all strains of the top 7 EHEC serotypes gave a positive result for *espK* and/or *ureD* (Tables 6).

DISCUSSION

In many countries the emergence of O157 and non-O157 EHEC in severe and epidemic human disease is of great concern. Accord-

TABLE 2 Primers and probes for the real-time PCR array

Target gene (ORF name if chromosomal)[a]	Primer		Probe sequence (5′ to 3′)[b]	Location within GenBank sequence AE005174 (nt range)
	Forward	Reverse		
ureD (*Z1142*)	GCAATAATTGACTCTGATTGCC	GCTGCTGCGGTAAAATTTACT	[FAM]-TAGGCTGATCACCATGCCTGGTGC-[BHQ1]	1078824–1078845
				1078892–1078872
espV (*Z1387*)	TCAGGTTCCTCGTCTGATGCCGC	CTGGTTCAGGCCTGGAGCAGTCC	[FAM]-CTTGCAACAGGTTACGCTGCCGAGTATT-[BHQ1]	1078847–1078870
				1295446–1295424
espK (*Z1829*)	GCAGRCATCAAAAGCGAAATCACACC	TCGTTTGGTAACTGTGGCAGATACTC	[FAM]-ATTCAGATAGAAGAAGCGCGGGCCAG-[BHQ1]	1295360–1295382
				1295422–1295395
espN (*Z1824*)	GACATATTTGTTTATGTCATCAGGAGCGG	CCTCAGGATATGGATGGCCTACTGGC	[FAM]-AATGCTCTCGGCAATCGAATCCTTGACTC-[BHQ1]	1673422–1673397
				1673312–1673338
Z2098	CTGAAAAGAGCCAGAAACGTGC	TGCCTAAGATCATTACCCGGAC	[HEX]-TAACTGCTATACCTCGCGCGCG-[BHQ1]	1673395–16673370
				1666890–1666918
espM1 (*Z2565*)	GCGCTCTATCCGCTTTAATGTTAAC	CCATCCATGAATATCTTTAGTACTCTGC	[FAM]-TGCTTACGCTCTCCAGTATACAGCGCCT-[BHQ1]	1667017–1666992
				1666991–1666963
				1888173–1888193
				1888308–1888287
				1888286–1888265
				2275378–2275354
				2275291–2275318
				2275320–2275347

[a] Numbering as in *E. coli* O157:H7 EDL933 (GenBank accession no. AE005174).
[b] FAM, 6-carboxyfluorescein; BHQ1, black hole quencher 1; HEX, 5′-hexachlorofluorescein.

TABLE 4 Distribution of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* and combinations thereof among EPEC serotype strains

Genetic marker[a]	No. of strains detected/total no. of strains (%) for:					
	O103:H2	O121:H19	O145:H28	O157:H7	O26:H11	Other EPEC[b]
Z2098	7/11 (63.6)	0/5 (0.0)	8/8 (100)	9/9 (100)	31/71 (43.7)	36/288 (12.5)
espK	9/11 (81.8)	3/5 (60.0)	8/8 (100)	7/9 (77.8)	30/71 (42.3)	56/288 (19.4)
espN	6/11 (54.5)	0/5 (0.0)	8/8 (100)	6/9 (66.7)	13/71 (18.3)	19/288 (6.6)
espV	6/11 (54.5)	5/5 (100)	8/8 (100)	9/9 (100)	31/71 (43.7)	118/288 (41.0)
ureD	1/11 (9.1)	5/5 (100)	5/8 (62.5)	6/9 (66.7)	20/71 (28.2)	34/288 (11.8)
espM1	0/11 (0.0)	5/5 (100)	8/8 (100)	9/9 (100)	64/71 (90.1)	49/288 (17.0)
espK/*espV*	10/11 (90.9)	5/5 (100)	8/8 (100)	9/9 (100)	43/71 (60.6)	137/288 (47.6)
espK/*ureD*	9/11 (81.8)	5/5 (100)	8/8 (100)	9/9 (100)	30/71 (42.3)	70/288 (24.3)
espK/*Z2098*	10/11 (90.9)	3/5 (60)	8/8 (100)	9/9 (100)	33/71 (46.5)	81/288 (28.1)
espN/*Z2098*	7/11 (63.6)	0/5 (0)	8/8 (100)	9/9 (100)	31/71 (43.7)	49/288 (17.0)
espN/*espM1*	6/11 (54.5)	5/5 (100)	8/8 (100)	9/9 (100)	65/71 (91.5)	57/288 (19.8)

[a] *espK*/*espV*, strains giving a positive result for *espK* and/or *espV*; *espK*/*ureD*, strains giving a positive result for *espK* and/or *ureD*; *espK*/*Z2098*, strains giving a positive result for *Z2098* and/or *espK*; *espN*/*Z2098*, strains giving a positive result for *espN* and/or *Z2098*; *espN*/*espM1*, strains giving a positive result for *espN* and/or *espM1*.
[b] Most of the EPEC strains tested in this study were described by Bugarel et al. (25).

ingly, rapid and specific detection of EHEC strains has become a priority for public health authorities (5, 6; Table 7 of the 2009 CDC Foodnet report of the food-borne diseases active surveillance network, http://www.cdc.gov/foodnet/factsandfigures/Top10SalmonellaSerotypes.pdf.). The availability of rapid and specific methods for testing EHEC strains and their virulence markers is also a prerequisite for establishing monitoring programs to follow EHEC contamination in animals and foodstuffs which are part of food inspection programs. Detection of non-O157 EHEC is particularly challenging, because unlike EHEC O157, these strains have few characteristics that distinguish them from the large number of harmless commensal *E. coli* strains that share the same niches.

The current approach for detecting EHEC in food and stool samples is to screen first for the presence of the stx_1/stx_2 genes and the *eae* gene. The CEN/ISO TS 13136 (9) and MLG 5B.01 (USDA report on detection and identification of non-O157 STEC from meat products) standard methods request the presence of both the stx_1/stx_2 and *eae* genes for further investigation of specific sequences derived from the O-antigen genes associated with the seven priority serogroups. This approach bears the disadvantage that a large number of non-EHEC strains can produce cross-reactivity with the targets used in these tests. This is of particular interest in complex and polymicrobial samples such as food, fecal, and environmental specimens (23, 24, 25). Accordingly, the simple detection of the stx_1/stx_2, *eae*, and O-antigen genes in a food product may often result in its rejection even though it does not contain EHEC strains of the top 7 serotypes. For food inspection purposes, tests are needed that not only identify the targeted EHEC serogroups but also target the most salient distinguishing features of the priority STEC associated with human illness.

In an attempt to identify discriminative genetic markers associated with STEC strains related to the world's most frequent clinical cases, we evaluated the distribution of the genes *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* in various *E. coli* pathogroups. Based on a high-throughput real-time PCR approach, a very large panel of *E. coli* strains (n = 1,100) that comprised EHEC (n = 340), EPEC (n = 392), STEC (n = 193), and apathogenic *E. coli* (n = 175) was examined for these genetic markers. Distributions of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* were significantly different among the various *E. coli* pathogroups. These genetic markers were highly prevalent in EHEC (67.1% to 92.4%) but were individually insufficient to identify all EHEC isolates. Even though none of the genetic markers was, by itself, capable of reliably detecting all EHEC isolates, when used in association with each other they identified EHEC strains with high confidence. Hence, only a few EHEC strains did not react with the combinations of the genetic markers tested here. These might be aberrant strains, not representative for the classical EHEC types. Looking at other genes in these anecdotal strains or sequencing

TABLE 5 Distribution of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* among EHEC serotypes

Genetic marker	No. of strains detected/total no. of strains (%) for:									
	Top 7 serotypes	O103:H2	O111:H8	O121:H19	O145:H28	O157:H7	O26:H11	O45:H2	Other EHEC[a] (new emerging EHEC)	Total EHEC
Z2098	250/277 (90.3)	49/49 (100)	47/51 (92.2)	17/20 (85.0)	30/30 (100)	49/66 (74.2)	44/44 (100)	14/17 (82.4)	47/63 (74.6)	297/340 (87.4)
espK	269/277 (97.1)	48/49 (98.0)	51/51 (100)	19/20 (95.0)	29/30 (96.7)	62/66 (93.9)	43/44 (97.7)	17/17 (100)	45/63 (71.4)	314/340 (92.4)
espN	261/277 (94.2)	47/49 (95.9)	51/51 (100)	17/20 (85.0)	30/30 (100)	59/66 (89.4)	40/44 (90.9)	17/17 (100)	33/63 (52.4)	294/340 (86.5)
espV	248/277 (89.5)	48/49 (98.0)	51/51 (100)	20/20 (100)	30/30 (100)	65/66 (98.5)	34/44 (77.3)	0/17 (0)	39/63 (61.9)	287/340 (84.4)
ureD	257/277 (92.8)	35/49 (71.4)	51/51 (100)	16/20 (80.0)	30/30 (100)	64/66 (97.0)	44/44 (100)	17/17 (100)	47/63 (74.6)	304/340 (89.4)
espM1	206/277 (74.4)	1/49 (2.0)	51/51 (100)	20/20 (100)	30/30 (100)	64/66 (97.0)	40/44 (90.9)	0/17 (0)	22/63 (34.9)	228/340 (67.1)

[a] O103:H25 (n = 2), O118:H16 (n = 4), O118:H2, O119:H25 (n = 5), O123:H11, O127:H8s, O145, O145:H25 (n = 5), O156:H21, O156:H25 (n = 11), O165:H25 (n = 2), O172:H25 (n = 2), O172:NM, O177 (n = 2), O177:H25, O182:H25, O3, O49:H16, O5 (n = 11), O55:H7 (n = 2), O76:H51, O84:H2, Ont:H2, Ont:H25 (n = 2), Or:H16, OX186:H2.

TABLE 6 Detection of EHEC according to different combinations of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1*

Gene association[a]	No. of strains detected/total no. of strains (%) for: EHEC of the top 7	O103:H2	O111:H8	O121:H19	O145:H28	O157:H7	O26:H11	O45:H2	Other EHEC[b] (new emerging EHEC)	Total EHEC
espK/*espV*	276/277 (99.6)	49/49 (100)	51/51 (100)	20/20 (100)	30/30 (100)	66/66 (100)	43/44 (97.7)	17/17 (100)	59/63 (93.7)	335/340 (98.5)
espK/*ureD*	277/277 (100)	49/49 (100)	51/51 (100)	20/20 (100)	30/30 (100)	66/66 (100)	44/44 (100)	17/17 (100)	59/63 (93.7)	336/340 (98.8)
espK/*Z2098*	275/277 (99.3)	49/49 (100)	51/51 (100)	19/20 (95.0)	30/30 (100)	65/66 (98.5)	44/44 (100)	17/17 (100)	59/63 (93.7)	334/340 (98.2)
espN/*Z2098*	273/277 (98.6)	49/49 (100)	51/51 (100)	17/20 (85.0)	30/30 (100)	65/66 (98.5)	44/44 (100)	17/17 (100)	55/63 (87.3)	328/340 (96.5)
espN/*espM1*	275/277 (99.3)	47/49 (95.9)	51/51 (100)	20/20 (100)	30/30 (100)	66/66 (100)	44/44 (100)	17/17 (100)	44/63 (69.8)	319/340 (93.8)

[a] *espK*/*espV*, strains giving a positive result for *espK* and/or *espV*; *espK*/*ureD*, strains giving a positive result for *espK* and/or *ureD*; *espK*/*Z2098*, strains giving a positive result for *Z2098* and/or *espK*; *espN*/*Z2098*, strains giving a positive result for *espN* and/or *Z2098*; *espN*/*espM1*, strains giving a positive result for *espN* and/or *espM1*.
[b] O103:H25 ($n = 2$), O118:H16 ($n = 4$), O118:H2, O119:H25 ($n = 5$), O123:H11, O127:H8s, O145, O145:H25 ($n = 5$), O156:H21, O156:H25 ($n = 11$), O165:H25 ($n = 2$), O172:H25 ($n = 2$), O172:NM, O177 ($n = 2$), O177:H25, O182:H25, O3, O49:H16, O5 ($n = 11$), O55:H7 ($n = 2$), O76:H51, O84:H2, Ont:H2, Ont:H25 ($n = 2$), Or:H16, OX186:H2.

their genomes might reveal more differences, which may make things clearer regarding their status. We should assume, on principle, that it is not necessarily the case that all members of a particular serotype would be EHEC.

The genetic markers examined in this study were also detected in some EPEC strains (13.3% to 45.2%) and very rarely in STEC (0.5% to 3.6%) and apathogenic *E. coli* (0 to 2.9%) strains. It is noteworthy that 38.5% of the EPEC strains that tested positive for at least one of the genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* belonged to the top 7 EHEC serotypes. Interestingly, other EPEC strains having a known EHEC serotype, such as O55:H7 and O103:H25, were also found positive for at least one of these genetic markers. Additionally, some of the EPEC strains that were found positive for at least one of these genetic markers were previously found to be closely related to EHEC strains according to their virulence gene content (22). These findings indicate that such isolates might be Stx–negative derivatives of EHEC that are also designated as EHEC-like strains (14). We assumed these isolates were EHEC derivatives according to their serotypes and *nle* gene content but they might also be EPEC strains that we are as yet unable to discriminate from EHEC derivatives. Further investigation using whole-genome sequencing may clarify the exact designation of these strains in the future.

The genetic markers *ureD*, *espV*, *espK*, *espN*, *Z2098*, and *espM1* were detected at different frequencies among the EHEC serotypes. We explored the various associations of these genetic markers to search for the best combinations of markers giving higher specificity and sensitivity for detecting EHEC. The genetic markers *espK*, *espV*, *ureD*, and *Z2098* were shown to be the best candidates as genetic markers for detecting EHEC. Taken individually, they were not detected in all strains of the top 7 EHEC serotypes, while in association they were detected in 98.6% to 100% of the top 7 EHEC strains. The association of *espK* with either *espV*, *ureD*, or *Z2098* proved to be the best combinations for more specific and sensitive detection of EHEC strains. Hence, a positive result for *espK* and/or *espV* was observed in 99.6% of EHEC strains belonging to the seven major serotypes of EHEC reported worldwide in human infections (only one EHEC O26:H11 isolate tested negative). Also, 93.7% of EHEC strains with serotypes other than those of the top 7 serotypes tested positive for *espK* and/or *espV*. Only a subset (54.1%) of EPEC strains tested positive for *espK* and/or *espV*. Most STEC and avirulent *E. coli* strains were negative for both *espK* and *espV*. Another interesting approach was to associate *espK* with *Z2098*. This combination of genetic markers resulted in the detection of 99.3% of EHEC strains belonging to the 7 major EHEC serotypes and 93.7% of EHEC strains with serotypes other than the top 7. The presence of *espK* and/or *Z2098* was found in only 36.7% of EPEC, 3.6% of STEC and 2.3% of apathogenic *E. coli* strains. The best approach for detecting EHEC with the highest specificity and sensitivity was to combine *espK* with *ureD*. This association allowed detection of 100% of EHEC strains of the top 7 serotypes and 93.7% of EHEC strains belonging to other serotypes. Detection of *espK* and/or *ureD* was also reported for only 33.4% of EPEC, 3.6% of STEC, and 3.4% of apathogenic *E. coli* strains.

These findings showed that combining detection of *espK* with either *espV*, *ureD*, or *Z2098* is a highly sensitive and specific approach for identifying with ≥99% confidence EHEC serotypes related to the world's most frequent clinical cases. Detection of these genetic markers in combination with *stx* in complex samples (food or fecal specimens) would provide a more EHEC-targeted diagnostic approach than combining only *stx* and *eae*. Given the rapidity of these PCR assays, this approach should have a major impact on top-7 EHEC surveillance and outbreak investigations and is likely to be of benefit to public health. Moreover, detection of these sets of genetic markers in 93.7% of EHEC strains having serotypes other than the top 7 may be helpful for identifying new emerging EHEC strains. The number of strains and the diversity of serotypes and pathogroups that were investigated in this study provide a solid basis for future utilization of these tests for the development of analytical methods and risk characterization of STEC. This should be confirmed with further evaluation of these assays on spiked and naturally contaminated samples. A complete evaluation of these tests in real samples will be crucial to definitively know how they can be applied in EHEC surveillance and outbreak investigations.

ACKNOWLEDGMENTS

We are grateful to F. Scheutz (WHO Collaborating Centre for Reference and Research on *Escherichia* and *Klebsiella*, Copenhagen, Denmark), A. Gill (Health Canada, Ottawa, Canada), M. Rivas (INEI-ANLIS, Buenos Aires, Argentina), G. H. Loneragan (Texas Tech University, Lubbock, TX), C. DebRoy (Pennsylvania State University, State College, PA), Roxane M. F. Piazza (Instituto Butantan, Sao Paulo, Brazil), P. Feng (U.S. FDA, College Park, MD), and N. Strockbine (CDC, Atlanta, GA) for providing some *E. coli* isolates or DNA extracts from *E. coli*.

REFERENCES

1. **Levine MM.** 1987. *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. J. Infect. Dis. **155:**377–389.
2. **Nataro JP, Kaper JB.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
3. **Karmali MA, Mascarenhas M, Shen S, Ziebell K, Johnson S, Reid-Smith R, Isaac-Renton J, Clark C, Rahn K, Kaper JB.** 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.
4. **Bugarel M, Beutin L, Martin A, Gill A, Fach P.** 2010. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with hemorrhagic colitis and hemolytic uremic syndrome in humans. Int. J. Food Microbiol. **142:**318–329.
5. **European Food Safety Authority (EFSA).** 2013. Scientific opinion on VTEC-seropathotype and scientific criteria regarding pathogenicity assessment. EFSA J. **11:**3138. http://www.efsa.europa.eu/en/efsajournal/pub/3138.htm.
6. **European Food Safety Authority (EFSA).** 2007. Scientific opinion of the panel on biological hazards on a request from EFSA on monitoring of verotoxigenic *Escherichia coli* (VTEC) and identification of human pathogenic types. EFSA J. **579:**1–61.
7. **Fey PD, Wickert RS, Rupp ME, Safranek TJ, Hinrichs SH.** 2000. Prevalence of non-O157:H7 shiga toxin-producing *Escherichia coli* in diarrheal stool samples from Nebraska. Emerg. Infect. Dis. **6:**530–533.
8. **Vally H, Hall G, Dyda A, Raupach J, Knope K, Combs B, Desmarchelier P.** 2012. Epidemiology of Shiga toxin producing *Escherichia coli* in Australia, 2000–2010. BMC Public Health **12:**63. doi:10.1186/1471-2458-12-63.
9. **European Food Safety Authority (EFSA).** 2009. Scientific report of EFSA: technical specifications for the monitoring and reporting of verotoxigenic *Escherichia coli* (VTEC) on animals and food (VTEC surveys on animals and food). EFSA J. **7:**1–43.
10. **Elliott SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng YK, Lai LC, McNamara BP, Donnenberg MS, Kaper JB.** 1998. The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. Mol. Microbiol. **28:**1–4.
11. **Perna NT, Mayhew GF, Pósfai G, Elliott S, Donnenberg MS, Kaper JB, Blattner FR.** 1998. Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. **66:**3810–3817.
12. **Beutin L, Jahn S, Fach P.** 2009. Evaluation of the 'GeneDisc' real-time PCR system for detection of enterohaemorrhagic *Escherichia coli* (EHEC) O26, O103, O111, O145 and O157 strains according to their virulence markers and their O- and H-antigen-associated genes. J. Appl. Microbiol. **106:**1122–1132.
13. **Delannoy S, Beutin L, Fach P.** 2012. Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic *Escherichia coli* (EHEC) strains of serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28 and O157:H7 by real time PCR. J. Clin. Microbiol. **50:**4035–4040.
14. **Delannoy S, Beutin L, Fach P.** 2013. Towards a molecular definition of enterohemorrhagic *Escherichia coli* (EHEC): detection of genes located on O island 57 as markers to distinguish EHEC from closely related enteropathogenic *E. coli* strains. J. Clin. Microbiol. **51:**1083–1088.
15. **Wickham ME, Lupp C, Mascarenhas M, Vazquez A, Coombes BK, Brown NF, Coburn BA, Deng W, Puente JL, Karmali MA, Finlay BB.** 2006. Bacterial genetic determinants of non-O157 STEC outbreaks and hemolytic-uremic syndrome after infection. J. Infect. Dis. **194:**819–827.
16. **Coombes BK, Wickham ME, Mascarenhas M, Gruenheid S, Finlay BB, Karmali MA.** 2008. Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing *Escherichia coli* strains. Appl. Environ. Microbiol. **74:**2153–2160.
17. **Konczy P, Ziebell K, Mascarenhas M, Choi A, Michaud C, Kropinski AM, Whittam TS, Wickham M, Finlay B, Karmali MA.** 2008. Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing *Escherichia coli*. J. Bacteriol. **190:**5832–5840.
18. **Bugarel M, Beutin L, Fach P.** 2010. Low-density macroarray targeting non-locus of enterocyte effacement effectors (nle genes) and major virulence factors of Shiga toxin-producing *Escherichia coli* (STEC): a new approach for molecular risk assessment of STEC isolates. Appl. Environ. Microbiol. **76:**203–211.
19. **Vlisidou I, Marchés O, Dziva F, Mundy R, Frankel G, Stevens MP.** 2006. Identification and characterization of EspK, a type III secreted effector protein of enterohaemorrhagic *Escherichia coli* O157:H7. FEMS Microbiol. Lett. **263:**32–40.
20. **Imamovic L, Tozzoli R, Michelacci V, Minelli F, Marziano ML, Caprioli A, Morabito S.** 2010. OI-57, a genomic island of *Escherichia coli* O157, is present in other seropathotypes of Shiga toxin-producing *E. coli* associated with severe human disease. Infect. Immun. **78:**4697–4704.
21. **Delannoy S, Beutin L, Ylanna Burgos Fach P.** 2012. Specific detection of enteroaggregative hemorrhagic *Escherichia coli* O104:H4 strains using the CRISPR locus as target for a diagnostic real-time PCR. J. Clin. Microbiol. **50:**3485–3492.
22. **Bugarel M, Martin A, Fach P, Beutin L.** 2011. Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) *Escherichia coli* strains: a basis for molecular risk assessment of typical and atypical EPEC strains. BMC Microbiol. **11:**142. doi:10.1186/1471-2180-11-142.
23. **Perelle S, Dilasser F, Grout J, Fach P.** 2007. Screening food raw materials for the presence of the world's most frequent clinical cases of Shiga toxin-encoding *Escherichia coli* O26, O103, O111, O145 and O157. Int. J. Food Microbiol. **113:**284–288.
24. **Fratamico PM, Bagi LK, Cray WC, Naragan N, Yan X, Medina M, Liu Y.** 2011. Detection by multiplex real-time polymerase chain reaction assays and isolation of Shiga toxin producing *Escherichia coli* serogroups O26, O45, O103, O111, O121 and O145 in ground beef. Foodborne Pathog. Dis. **8:**601–607.
25. **DebRoy C, Roberts E, Valadez AM, Dudley EG, Cutter CN.** 2011. Detection of Shiga toxin producing *Escherichia coli* O26, O45, O103, O111, O113, O121, O145, and O157 serogroups by multiplex polymerase chain reaction of the *wzx* gene of the O-antigen gene cluster. Foodborne Pathog. Dis. **8:**651–652.