View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@University of Nebraska

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Public Health Resources

Public Health Resources

2005

Evolutionary Genetics of a New Pathogenic *Escherichia* Species: *Escherichia albertii* and Related *Shigella boydii* Strains

Katie E. Hyma
Michigan State University

David W. Lacher
Michigan State University

Adam M. Nelson
Michigan State University

Alyssa C. Bumbaugh
Michigan State University

J. Michael Janda
California Department of Health Services

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/publichealthresources

 Part of the Public Health Commons

Hyma, Katie E.; Lacher, David W.; Nelson, Adam M.; Bumbaugh, Alyssa C.; Janda, J. Michael; Strockbine, Nancy A.; Young, Vincent B.; and Whittam, Thomas S., "Evolutionary Genetics of a New Pathogenic *Escherichia* Species: *Escherichia albertii* and Related *Shigella boydii* Strains" (2005). *Public Health Resources*. 218.
https://digitalcommons.unl.edu/publichealthresources/218

This Article is brought to you for free and open access by the Public Health Resources at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Public Health Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Katie E. Hyma, David W. Lacher, Adam M. Nelson, Alyssa C. Bumbaugh, J. Michael Janda, Nancy A. Strockbine, Vincent B. Young, and Thomas S. Whittam

This article is available at DigitalCommons@University of Nebraska - Lincoln: https://digitalcommons.unl.edu/publichealthresources/218

JOURNAL OF BACTERIOLOGY, Jan. 2005, p. 619–628 Vol. 187, No. 2
0021-9193/05/$08.00+0 doi:10.1128/JB.187.2.619–628.2005
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Evolutionary Genetics of a New Pathogenic *Escherichia* Species: *Escherichia albertii* and Related *Shigella boydii* Strains

Katie E. Hyma,[1] David W. Lacher,[1] Adam M. Nelson,[1] Alyssa C. Bumbaugh,[1] J. Michael Janda,[2] Nancy A. Strockbine,[3] Vincent B. Young,[1,4] and Thomas S. Whittam[1]*

Microbial Evolution Laboratory, National Food Safety and Toxicology Center,[1] and Infectious Disease Unit, Department of Internal Medicine,[4] Michigan State University, East Lansing, Michigan; Microbial Diseases Laboratory, Division of Communicable Disease Control, California Department of Health Services, Richmond, California[2]; and Centers for Disease Control and Prevention, Atlanta, Georgia[3]

Received 30 August 2004/Accepted 8 October 2004

A bacterium originally described as *Hafnia alvei* induces diarrhea in rabbits and causes epithelial damage similar to the attachment and effacement associated with enteropathogenic *Escherichia coli*. Subsequent studies identified similar *H. alvei*-like strains that are positive for an intimin gene (*eae*) probe and, based on DNA relatedness, are classified as a distinct *Escherichia* species, *Escherichia albertii*. We determined sequences for multiple housekeeping genes in five *E. albertii* strains and compared these sequences to those of strains representing the major groups of pathogenic *E. coli* and *Shigella*. A comparison of 2,484 codon positions in 14 genes revealed that *E. albertii* strains differ, on average, at ~7.4% of the nucleotide sites from pathogenic *E. coli* strains and at 15.7% from *Salmonella enterica* serotype Typhimurium. Interestingly, *E. albertii* strains were found to be closely related to strains of *Shigella boydii* serotype 13 (*Shigella* B13), a distant relative of *E. coli* representing a divergent lineage in the genus *Escherichia*. Analysis of homologues of intimin (*eae*) revealed that the central conserved domains are similar in *E. albertii* and *Shigella* B13 and distinct from those of *eae* variants found in pathogenic *E. coli*. Sequence analysis of the cytolethal distending toxin gene cluster (*cdt*) also disclosed three allelic groups corresponding to *E. albertii*, *Shigella* B13, and a nontypeable isolate serologically related to *S. boydii* serotype 7. Based on the synonymous substitution rate, the *E. albertii*-*Shigella* B13 lineage is estimated to have split from an *E. coli*-like ancestor ~28 million years ago and formed a distinct evolutionary branch of enteric pathogens that has radiated into groups with distinct virulence properties.

In the early 1990s, *Hafnia alvei* was first implicated as a causative agent of diarrhea (2). At the International Center for Diarrheal Disease Research, Dhaka, Bangladesh, clinical isolates of *H. alvei* were found to differ from representative *H. alvei* strains by phenotypic and genotypic assays, including the presence of an intimin gene (*eae*) homologous to the gene critical for the attaching-effacing phenotype of enteropathogenic *Escherichia coli* (EPEC) (3). A subsequent study identified similar *H. alvei*-like strains that were positive for the *eae* probe and, based on partial 16S rRNA sequences, were found to be more closely related to EPEC than to *eae*-negative *H. alvei* (20).

Further investigation led to the recognition of these *H. alvei*-like strains as a new *Escherichia* species, *Escherichia albertii*, based on the results of DNA-DNA hybridization and 16S rRNA gene sequence data (18). DNA relatedness, assessed by DNA-DNA hybridization, was highest (ranging from 54 to 64%) among *E. albertii*, *E. coli*, and *Shigella flexneri*, lowest (9 to 17%) between *E. albertii* and *H. alvei*, and intermediate (16 to 45%) among *E. albertii* and the four *Escherichia* species *E. blattae*, *E. fergusonii*, *E. hermanii*, and *E. vulneris* (18). Analysis of 16S rRNA gene sequences reinforced the conclusion that *E. albertii* does not belong to *H. alvei* but is most closely related to *Escherichia* and *Shigella* (18). Additionally, the structure of the O chain of the lipopolysaccharide of *E. albertii* has been described and does not resemble that of any known lipopolysaccharide O chain from *E. coli* or *H. alvei*; however, based on the types of sugar residues and acidity, it could fit into either group (13).

Based on genotypic analysis, another bacterium, *Shigella boydii* serotype 13 (*Shigella* B13), has also been shown to be a distant relative of *E. coli* representing a divergent lineage in the genus *Escherichia* (5, 26, 40). *E. coli* and *Shigella* have been described as one genetic species based on DNA-DNA hybridization patterns, and it has been demonstrated that although the O antigens of *Shigella* B13 exhibit a reciprocal cross-reaction with *E. coli* O28 antigens, they are genetically distinct (5). Additionally, *Shigella* B13 is the only member of the *E. coli*-*Shigella* group that can be reliably separated by bulk DNA-DNA hybridization techniques (5).

Nucleotide sequencing of multiple conserved genes has supported the hypothesis that *E. coli* and *Shigella* are the same genetic species and has also provided evidence that *Shigella* B13 belongs to a divergent branch outside the diversity of *E. coli* and *Shigella*. Furthermore, three main groups of *Shigella* strains are apparent, each including serotypes from the traditional *Shigella* species: *S. boydii*, *S. flexneri*, *S. sonnei*, and *S. dysenteriae* (26, 40). An independent multilocus sequence typing (MLST) analysis of a variety of serotypes of *Shigella* and enteroinvasive *E. coli* (EIEC) strains confirmed the presence of these three groups and supported the placement of *Shigella*

* Corresponding author. Mailing address: Microbial Evolution Laboratory, 165 Food Safety & Toxicology Building, Michigan State University, East Lansing, MI 48824. Phone: (517) 432-3100, ext. 178. Fax: (517) 432-2310. E-mail: whittam@msu.edu.

This article is a U.S. government work, and is not subject to copyright in the United States.

TABLE 1. Sources of *E. albertii* and *Shigella* strains

Strain	Species or type	Host	Locale	Source[a]
9194	*E. albertii*	Human (child)	Bangladesh	J. M. Janda
10790	*E. albertii*	Human (child)	Bangladesh	J. M. Janda
10457	*E. albertii*	Human (child)	Bangladesh	J. M. Janda
12502	*E. albertii*	Human (child)	Bangladesh	J. M. Janda
19982	*E. albertii*	Human (child)	Bangladesh	J. M. Janda
2045-54	*Shigella* B13	Human (child)	Egypt	CDC
2046-54	*Shigella* B13	Human (child)	Egypt	CDC
1610-55	*Shigella* B13	Monkey	United States (Fla.)	CDC
5216-70	*Shigella* B13	Human	United States (Ohio)	CDC
603-73	*Shigella* B13	Human	United Kingdom	CDC
3552-77	*Shigella* B13	Human	United Kingdom	CDC
3553-77	*Shigella* B13	Human	United Kingdom	CDC
3554-77	*Shigella* B13	Human	United Kingdom	CDC
3555-77	*Shigella* B13	Human	United Kingdom	CDC
3556-77	*Shigella* B13	Human	United Kingdom	CDC
3557-77	*Shigella* B13	Human	United Kingdom	CDC
3052-94	*Shigella* B13	Human (infant)	New Zealand	CDC
3053-94	*Shigella* B13	Human (19 yr)	New Zealand	CDC
3103-99	*Shigella* B13	Human (40 yr)	United States (Ill.)	CDC
C-425 (ATCC 12032)	*Shigella* B13	Human	No data	ATCC
616	*Shigella* B13	Human	No data	CDC
K-694	*Shigella* B13	Human	Bangladesh	K. A. Talukder
K-1	*Shigella* B7	Human	Bangladesh	K. A. Talukder
3097-02	Nontypeable	Human	United States (Wis.)	CDC

[a] CDC, Centers for Disease Control and Prevention; ATCC, American Type Culture Collection.

serotypes into distinct groups within the diversity of the *E. coli* population; this study also placed nontypeable *Shigella* strains into these three groups (6).

Attachment and effacement appear to be critical components of the virulence of *E. albertii* (2, 3, 19). In *Citrobacter rodentium* and diarrheagenic *E. coli*, attaching-effacing lesion formation is mediated by intimin, an outer membrane protein encoded by the *eae* gene, and other components of the locus of enterocyte effacement pathogenicity island (34). Intimin is a highly polymorphic protein with more than 20 distinct subtypes and has been discovered in clinical isolates and animal sources (1, 25, 41, 54). The C-terminal region of the protein is the surface-exposed, immunogenic domain, which functions in binding to epithelial cells via the translocated bacterial receptor, Tir (30). In contrast, the central region is a highly conserved transmembrane domain with homology to the invasins of pathogenic *Yersinia* (53). Sequence comparisons have shown that intimin subtypes are encoded by distinct *eae* alleles with mosaic structures that presumably have arisen through lateral gene transfer and recombination (33, 48).

A second putative virulence factor discovered in *E. albertii* is cytolethal distending toxin (CDT) (38). In addition to strains of pathogenic *E. coli* and *Shigella* (22, 24), CDT is dispersed phylogenetically and has been found in *Campylobacter* spp. (23), *Haemophilus ducreyi* (8), *Actinobacillus actinomycetemcomitans* (46), and *Helicobacter* spp. (7, 50, 52). Cultured mammalian cell lines exposed to CDT become arrested in the G_1 or G_2 phase of the cell cycle, in association with cell distension and eventually death (9, 27, 28, 36). In all bacterial species in which CDT activity has been demonstrated, three closely linked genes (*cdtA*, *cdtB*, and *cdtC*) have been associated with cytotoxic activity (28, 36, 37, 44). Although genetic studies have indicated that all three genes are necessary to transfer full activity to a noncytotoxic *E. coli* strain (9, 28, 36, 37), *cdtB* has been implicated as encoding the enzymatically active subunit (27, 28). The CdtB subunit is homologous to a class of phosphodiesterases that includes nucleases, and it has been demonstrated that CdtB damages host cell DNA through DNase I-like activity in vitro (12, 17, 27, 32). This activity triggers the DNA damage checkpoint response (9), which results in cell cycle arrest (9, 12, 27, 28). The CdtA and CdtC subunits function together to mediate the binding of the AB_2 holotoxin to eukaryotic cells and the subsequent delivery of CdtB to the cytoplasm (10, 29, 35).

The purpose of the study reported here was to elucidate the phylogenetic relationships of *E. albertii* and its relatives that express *Shigella* B13- and *S. boydii* serotype 7 (*Shigella* B7)-related antigens. In addition, we characterize the genes encoding the putative virulence factors intimin and CDT, whose role, if any, in the pathogenesis of *E. albertii* and relatives is not known. We propose a model, based on MLST analysis and supported by virulence gene sequences, in which the divergence of the *E. albertii* lineage preceded the radiation of pathogenic *E. coli* and *Shigella*. We also demonstrate that the *E. albertii* lineage includes other unusual strains, such as divergent *Shigella* B13 (5) and two nontypeable clinical isolates.

MATERIALS AND METHODS

Bacterial strains. All strains were grown overnight in Luria-Bertani broth at 37°C. The strains examined included 5 *E. albertii* human clinical isolates (18, 20), 17 *Shigella* B13 isolates, 1 nontypeable strain, and 1 *Shigella* B7 isolate (Table 1). We use the nomenclature of Pupo et al. (40) in referring to the *Shigella* serotypes as B13 and B7.

DNA isolation. Genomic DNA was isolated from 2 ml of overnight culture by using a Puregene DNA isolation kit (Gentra Systems Inc., Minneapolis, Minn.). DNA preparations were quantified by using a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Wilmington, Del.). DNA preparations were diluted to a final concentration of 100 ng/μl and stored at 4°C.

MLST analysis. The nucleotide sequences of internal fragments of the following genes were obtained for MLST analysis: *arcA*, *aroE*, *aspC*, *clpX*, *cyaA*, *dnaG*, *fadD*, *grpE*, *icdA*, *lysP*, *mdh*, *mtlD*, *mutS*, *rpoS*, and *uidA*. The analysis included 34 *E. coli* and *Shigella* strains and 2 *Salmonella* strains for comparison. Information on primers and protocols for MLST can be found at website http://www.shigatox.net/mlst. Sequence analysis was performed with MEGA2 software (version 2.0; Pennsylvania State University, University Park).

16S rRNA gene sequencing. The forward primer S-D-Bact-0008-a-S-20 and the reverse primer S-*-Univ-1492-b-A-21 (45) were used to amplify bacterial 16S rRNA genes by PCR under the following conditions: 94°C for 10 min; 35 cycles of 92°C for 1 min, 58°C for 1 min, and 72°C for 1 min; and 72°C for 5 min. Additional internal sequencing primers included 16SintF1 (5′-CGG AGG GTG CAA GCG TTA AT-3′), 16SintF2 (5′-ACC CTT ATC CTT TGT TGC-3′), 16SintR1 (5′-CCA GGT AAG GTT CTT CGC GT-3′), and 16SintR2 (5′-AGT ACT TTA CAA CCC GAA GG-3′).

Diagnostic multiplex PCR. The following primers were designed and used for diagnostic multiplex PCR: clpX_28F (5′-TGG GCT CGA GTT GGG CA-3′), clpX_411R (5′-TCC TGC TGC GGA TGT TTA CG-3′), lysP_107F (5′-GGG CGC TGC TTT CAT ATA TTC TT-3′), lysP_358R (5′-TCC AGA TCC AAC CGG GAG TAT CAG GA-3′), mdh_50F (5′-CTG GAAGGC GCA GAT GTG GTA CTG ATT-3′), and mdh_164R (5′-CTT GCT GAA CCA GAT TCT TCA CAA TAC CG-3′).

***eae* sequencing.** The entire *eae* gene, which encodes intimin, was sequenced for representative strains 19982, K-694, C-425, 3555-77, 3097-02, and K-1. PCR and sequencing primers included cesT-F9 (5′-TCA GGG AAT AAC ATT AGA AA-3′), eae-F1 (5′-ACT CCG ATT CCT CTG GTG AC-3′), eae-R3 (5′-TCT TGT GCG CTT TGG CTT-3′), and escD-R1 (5′-GTA TCA ACA TCT CCC GCC CA-3′). Internal sequencing primers included eae-F01s (5′-AAG CAA CAT GAC CGA TGA CA-3′), eae-F02s (5′-GGC GAA TAC TGG CGA GAC TA-3′), eae-F05s (5′-AGT TAA ATG CAA ATG GCG GT-3′), eae-F06s (5′-GTA ACG GAC TTT ACG GCT GAT A-3′), eae-F07s (5′-GCC TAA TAG CAT AGT GAG CGT T-3′), eae-F08s (5′-GTG ATA ATC AAA CCG CAA CT-3′), eae-F09s (5′-CGA CCT TAG GGA CGT TAA GT-3′), eae-F10s (5′-GCT TGT TAG TGC CCG AGT TC-3′), eae-R02s (5′-GGG CGC TCA TCA TAG TCT TT-3′), eae-R05s (5′-GGT AAT AAT GCC TGT ATT AG-3′), eae-R06s (5′-AGT GAG CGA GAC TGA AGT TG-3′), eae-R08s (5′-GTC GCG GTA TAA GTA AT-3′), eae-R09s (5′-TTC CTG GGA ATA ATG TTG AT-3′), eae-R10s (5′-TGG CAT GAG CTT ACA GCA TC-3′), eae-R11s (5′-CGG ATA ACG GCT TAC CAT CT-3′), and eae-R12s (5′-GCA TTA AGC GCT GAA GTC AT-3′).

***cdt* sequencing.** Nucleotide sequences were obtained for an internal fragment of putative virulence locus *cdtB* by using the primers VAT2 and WMI1 (39). Sequences for the *cdtABC* operon were determined for the four representative *E. coli* strains 19982, K-694, 3097-02, and K-1. The following additional primers were used for PCR and sequencing: cdt_5_flank_F (5′-AGG AGA TTT GCC AAT GTG AT-3′), cdt_1101_B_R (5′-GTA TCT ACA GCC GTT GAT GG-3′), cdt_1248_B_R (5′-GGC AGG CCG ATG AGG TG-3′), cdt_2057_R (5′-CGT TTG TTG AAG GTA TGA GAG-3′), and cdt_1815_F (5′-TCT ATT CAC ATT CCG CA-3′). The following primers were used for internal sequencing: cdt_584_R (5′-CCC GGG AAG CTT GTC ATA-3′) and cdt_564_F (5′-CTT TCT ACT GCA CCT TGC CA-3′). Primer randR1 (5′-TGC CGG GCA A-3′) was used for PCR.

CDT activity assays. CDT activity assays were performed by the method of Young et al. (51, 52). Bacterial strains were grown overnight with shaking at 37°C in Luria-Bertani broth. Total bacterial proteins were solubilized with a nonionic detergent (bacterial protein extraction reagent; Pierce Chemical Co., Rockford, Ill.) according to the recommendations of the manufacturer. Bacteria were pelleted by centrifugation at 2,040 × *g* for 10 min, suspended in 300 μl of bacterial protein extraction reagent, and vortexed. Insoluble material was removed by centrifugation at 13,800 × *g* for 5 min, and soluble proteins contained in the supernatant were collected and stored at −80°C. HeLa cells were used to seed 13-mm-diameter circular glass coverslips in 24-well tissue culture plates at a density of $\sim 2 \times 10^3$ per well. Ten microliters of bacterial protein extract was added to each well, and the plates were incubated in 5% CO_2 at 37°C. At appropriate time points, coverslips were washed with phosphate-buffered saline and then fixed and stained with a Hema 3 stain set (Fisher Scientific, Hampton, N.H.). CDT-positive *E. coli* 493/89 (22) and CDT-negative *E. coli* K-12 were used as positive and negative controls, respectively.

To confirm the CDT-positive results, CDT activity assays were also performed with Shiga toxin-insensitive CHO cells by the method of Scott and Kaper (44). CHO cells were maintained at 37°C in Ham's F-12 medium plus glutamine and supplemented with 10% fetal bovine serum, penicillin, and streptomycin. Cells were trypsinized and suspended at 2×10^4 per ml in Ham's F-12 medium with 1% fetal bovine serum, penicillin, and streptomycin. The cell suspension was added to 96-well plates at 150 μl per well. Bacterial cultures were grown with shaking overnight in Trypticase soy broth with 0.6% yeast extract. Aliquots (15 μl per well) of each test supernatant were added in triplicate to freshly plated CHO cells. Assay mixtures were incubated for 5 days at 37°C in 5% CO_2 and then fixed and stained with the Hema 3 stain set.

Virulence gene profiling. Multilocus virulence gene profiling was performed by using a Beckman (Fullerton, Calif.) CEQ8000 genetic analysis system as described previously (47) with the following modification: additional labeled primer sets were used for more extensive profiling of strains to give a total of 30 genes for fragment analysis, including an *mdh* control. Several primer sets were obtained from IDT (Coralville, Iowa), Proligo (Boulder, Colo.), and Research Genetics (Huntsville, Ala.). Cycling conditions were modified to include an annealing temperature of 50°C and 40 amplification cycles. Additional primer pairs can be found at website http://www.shigatox.net/mlst.

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been submitted to GenBank under accession numbers AY696659 to AY697031. These include accession numbers AY696752 to AY696755 for sequences for the *cdtABC* operon from *E. coli* strains 19982, K-694, 3097-02, and K-1, respectively.

RESULTS

MLST analysis. In most cases, PCR primers used for amplifying and sequencing of *E. coli* and *Shigella* worked successfully for *E. albertii*, divergent *Shigella* B13, and related strains. The exceptions included the *uidA* gene, which encodes β-glucuronidase (a hallmark of typical *E. coli*) and was amplified in only five *Shigella* B13 strains: 5216-70, 3053-94, 3052-94, 3556-77, and 3557-77. In three strains, *E. albertii* 12502, *Shigella* B13 603-73, and *Shigella* B13 5216-70, a large (~1.3-kb) insertion sequence was present in the region of *rpoS* used for MLST analysis. The insertion sequence was removed for analysis. In *Shigella* B13 strain 616, *rpoS* could not be amplified with *E. coli* PCR primers.

The results of PCR and sequencing of housekeeping genes allowed for a comparison of 2,484 homologous codon positions in 14 genes (*uidA* not included). *E. albertii* strains were highly divergent from *E. coli* strains, differing on average at ~7.4% of the nucleotide sites from pathogenic *E. coli* strains (7.3% for EPEC strain E2348/69 and 7.5% for *E. coli* O157:H7) and 15.7% from *Salmonella enterica* serotype Typhimurium LT2. The *E. albertii* strains were most closely related to a *Shigella* B13 strain (ATCC 12032) (6), differing at only 0.72% of the nucleotide sites.

To assess the generality of these findings, an additional 16 *Shigella* B13 strains, 1 nontypeable strain, and 1 *Shigella* B7-related strain were examined. Five of the *Shigella* B13 strains (5216-70, 3053-94, 3052-94, 3556-77, and 3557-77) were phylogenetically more closely related to members of the *E. coli-Shigella* group than to other *Shigella* B13 strains (Fig. 1). These *Shigella* B13 strains were the same strains in which *uidA* could be amplified with *E. coli uidA* PCR primers. Strain 5216-70 (Fig. 1B) was most closely related to a group of EIEC strains (EIEC group 2) (6), whereas strains 3053-94, 3052-94, 3556-77, and 3557-77 were phylogenetically most closely related to each other and formed a distinct cluster (Fig. 1C); this cluster appears to have descended from an ancestor that also gave rise to all *Shigella* and EIEC groups. For clarification, these five strains are hereafter referred to as atypical *Shigella* B13 strains. The remaining *Shigella* B13 strains, the nontypeable strain, and the *Shigella* B7-related strain were most closely related to the

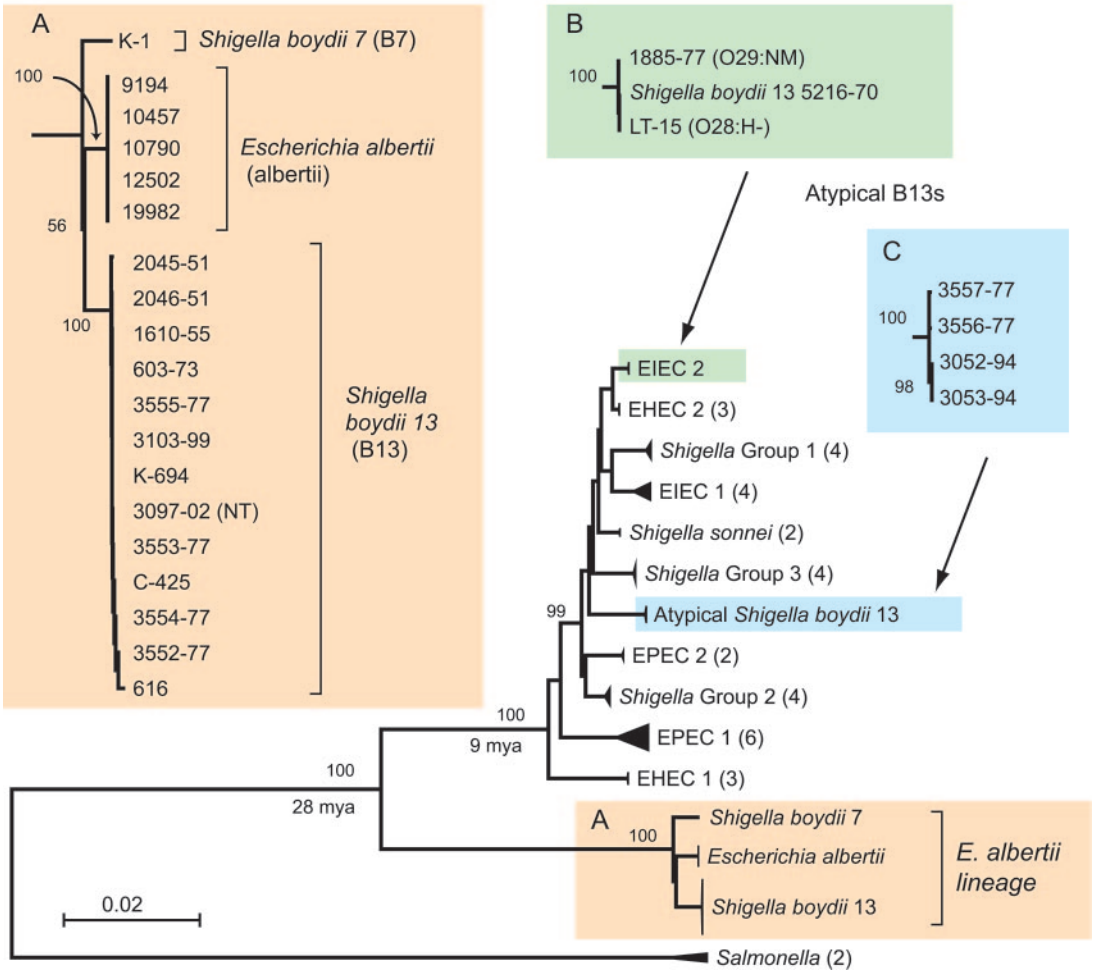


FIG. 1. Phylogenetic relationships of *Shigella* B13 strains and the *E. albertii* lineage. A neighbor-joining tree (MEGA2) was constructed by using the Tamura-Nei model with gamma correction for the distribution of rates based on 14 housekeeping genes. The *E. albertii* lineage (A) includes *E. albertii*, *Shigella* B13, and *Shigella* B7. On the basis of the rate of synonymous substitutions for *E. coli* and *S. enterica* (4.7×10^{-9} per site per year), the *E. albertii* lineage shared a common ancestor with *E. coli* and *Shigella* pathogenic groups ~28 million years ago (mya). Atypical *S. boydii* strains express the *Shigella* B13 antigen and do not cluster with the *E. albertii* lineage but are more similar in their MLST profile to *E. coli* and *Shigella* groups. *Shigella* B13 strain 5216-70 belongs to EIEC group 2 (B), and four atypical *Shigella* B13 strains belong to their own cluster within the diversity of *E. coli* and *Shigella* (C). The number of strains examined is in parentheses. NT, nontypeable; EHEC, enterohemorrhagic *E. coli*.

E. albertii strains and are hereafter collectively referred to as the *E. albertii* lineage (Fig. 1A).

Within the *E. albertii* lineage, ~1.1% (82 of 7,470) of nucleotide sites and ~0.2% (6 of 2,490) of amino acid sites are variable. Synonymous substitution rates per 100 nucleotide sites in housekeeping genes between *E. albertii* and *E. coli-Shigella* range from a low of 10.97 for *icdA* to a high of 37.99 for *cyaA* and average 25.34, compared to an average of 5.21 ± 0.26 (mean and standard deviation) within *E. coli-Shigella*. Because there is substantial divergence at every locus, there is no evidence for recent recombination or gene exchange events between the *E. albertii* lineage and the *E. coli* population. Nonsynonymous substitution rates are much lower in each comparison, indicating conservative selection of housekeeping loci and allowing us to measure sequence divergence independently of adaptive evolution (Table 2). It is noteworthy that the average divergence between *E. albertii* and *E. coli* represents about 40% of the evolutionary distance from *E. coli* to *S. enterica*.

We used the combined sequences of the 14 housekeeping genes to define distinct multilocus sequence types. There were six different sequence types in the *E. albertii* lineage, and three different groups of sequence types can be resolved. One of the groups contains four sequence types comprised of *Shigella* B13 strains and nontypeable strain 3097-02. Variability within this group is low: ~0.04% (3 of 7,470) of nucleotide sites and ~0.08% (2 of 2,490) of amino acid sites. The remaining two sequence types each represent their own branch; the *E. albertii* strains all belong to one, and *Shigella* B7-related strain K-1 belongs to the other. In pairwise comparisons, the *E. albertii* group varies from the *Shigella* B13 group at ~0.6% of nucleotide sites and ~0.01% of amino acid sites and from the *Shigella* B7-related strain at ~0.7% of nucleotide sites and ~0.08% of amino acid sites. The *Shigella* B7-related strain differs from the *Shigella* B13 group by ~0.7% of nucleotide sites and ~0.1% of amino acid sites (Fig. 1A).

16S rRNA gene sequencing. To assess further the genetic relatedness of the *E. albertii* and *Shigella* B13 strains, we se-

TABLE 2. Pairwise sequence divergence among strains of the *E. albertii-Shigella* B13 lineage (Ea), *E. coli* (Ec), and *S. enterica* (Se)[a]

Locus	d_S for:			d_N for:		
	Ea/Ec	Ea/Se	Ec/Se	Ea/Ec	Ea/Se	Ec/Se
arcA	14.71	44.43	47.09	0	0	0
aroE	18.94	112.57	104.86	1.23	7.98	7.65
aspC	32.02	89.38	93.63	0.31	4.05	4.26
clpX	12.48	48.27	44.19	0	0.77	0.76
cyaA	37.99	85.85	53.62	0.01	0.72	0.73
dnaG	41.17	78.98	82.81	0.33	11.83	11.6
fadD	48.61	120.56	98.68	0.77	3.81	3.17
grpE	12.43	53.97	46.81	1.02	5.07	4.71
icdA	10.97	61.53	57.87	0	1.23	1.00
lysP	31.19	80.85	73.44	1.84	3.77	4.09
mdh	32.00	73.63	58.66	0.27	2.61	2.34
mtlD	32.2	70.88	59.54	0.40	3.89	3.70
mutS	19.01	55.09	62.07	0.57	1.13	0.85
rpoS	20.84	28.5	26.44	0.01	0.01	0
Avg	25.34	67.00	60.65	0.43	3.03	2.86

[a] Sequence differences are expressed as the synonymous substitution rate (d_S) per 100 synonymous sites and the nonsynonymous substitution rate (d_N) per 100 nonsynonymous sites. Calculations were performed with MEGA2 software.

quenced and analyzed part of the 16S rRNA gene of all members of the *E. albertii* lineage and atypical *Shigella* B13 strains. There were 1.8% (27 of 1,464) variable sites within the lineage and five distinct sequence types divided into groups corresponding to MLST results: two groups of atypical *Shigella* B13 strains, one group of *E. albertii* strains, one group of *Shigella* B13 strains, and one group including the *Shigella* B7-related strain. In pairwise comparisons within the *E. albertii* lineage, the *Shigella* B13 and B7-related strains were the most similar (99.86%), and the *Shigella* B7-related and atypical *Shigella* B13 strains were the least similar (98.32%).

***eae* sequencing.** The close evolutionary relationships between *E. albertii* and *Shigella* B13 strains predicted that these bacteria should share additional chromosomal genes. We tested this prediction by attempting to detect, amplify, and sequence *eae* homologues in *Shigella* B13 and related *Shigella* strains of the *E. albertii* lineage. All strains tested were amplified, and an analysis of intimin (*eae*) sequences in six strains revealed three distinct *eae* variants, one in *E. albertii* strains, one in the *Shigella* B13 subcluster, and one in the *Shigella* B7-related strain. The central conserved domain of *eae* in six strains of the *E. albertii* lineage were most similar to each other, with the iota2 allele (GenBank accession number AF530553) being the most closely related *E. coli* allele. The C-terminal region of the intimin gene (Int280 region), in contrast to the central conserved region, which varied at 1.2% of nucleotide sites and 0.9% of amino acid sites within this lineage, differed at 37.6% of nucleotide sites and 42.3% of amino acid sites (Table 3). The Int280 region in strains of the *Shigella* B13 subcluster all have the same sequence, which was most similar to the iota2 allele sequence; *E. albertii* and *Shigella* B7-related strains, although possessing different alleles, are most closely related to each other, with the alpha and zeta alleles being the next most similar alleles (Fig. 2).

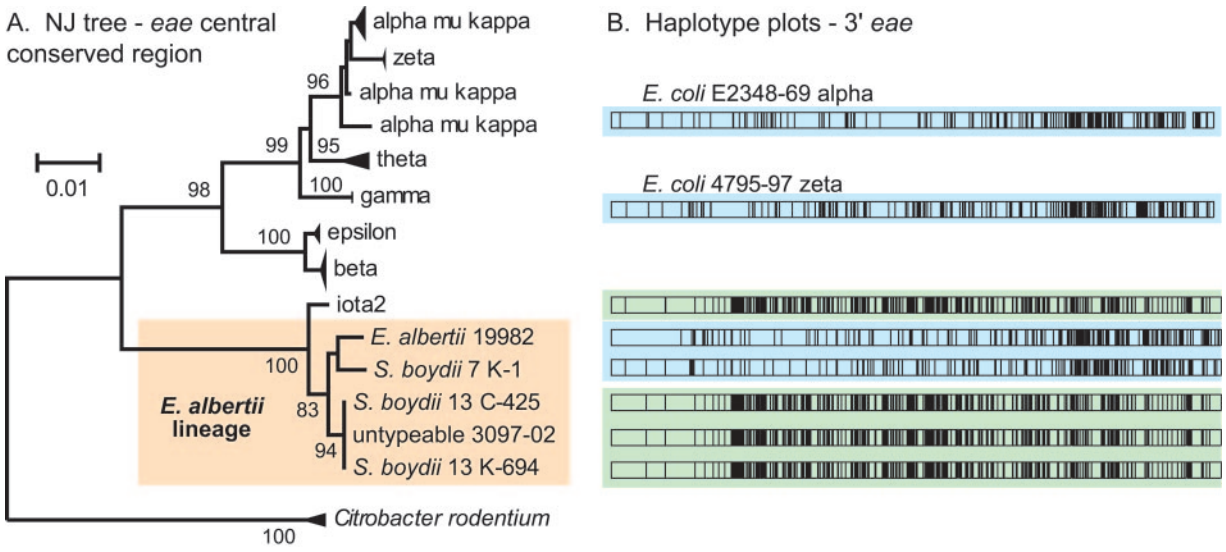
***cdt* sequencing and CDT expression.** Partial sequence data were determined for the *cdtB* subunit of members of the *E. albertii* lineage, and complete *cdtABC* sequences were obtained for representatives of the distinct subclusters of the *E. albertii* lineage. For atypical *Shigella* B13 strains and *Shigella* B13 strains 3553-77 and 3555-77, *cdtB* could not be PCR amplified. Sequence analysis of *cdtABC* revealed three allelic groups corresponding to the *E. albertii*, *Shigella* B13, and *Shigella* B7-like clusters within the lineage, consistent with evolutionary analysis of multiple housekeeping genes. The *cdt* sequences of the *E. albertii* lineage are most similar to the CDT V and CDT III variants (Fig. 3) originally found in *E. coli* strain 9142-88, in sorbitol-fermenting, Shiga toxin-producing *E. coli* O157:H$^-$ strain 493/89, and on *E. coli* plasmid pVIR. Although the amino acid residues in CdtB that are important for enzymatic activity have been conserved in the *E. albertii* lineage, all members of the *Shigella* B13 subcluster have a putative stop codon within the *cdtB* subunit (Fig. 3). Findings from subsequent experiments (see below) demonstrated that these strains do not produce stereotypical CDT activity.

All strains were tested for the ability to produce functional CDT. HeLa cells exposed to protein extracts from whole-cell lysates of *E. albertii* and *Shigella* B7-related strains exhibited stereotypical CDT cytopathic effects, including distended cytoplasm and nuclear abnormalities. HeLa cells exposed to protein extracts from *Shigella* B13 strains exhibited minimal cytopathic effects (Fig. 4), consistent with the hypothesis that the stop codon in the *Shigella* B13 subcluster has inactivated the gene. There were no instances in which CDT activity was de-

TABLE 3. Sequence divergence and properties of the domain and subunits of the intimin alleles (*eae*) and CDT gene (*cdt*) of the *E. albertii-Shigella* B13 lineage[a]

Locus	No. of		Mean ± SD		d_S/d_N	% GC	Codon adaptation index
	Alleles	Polymorphic sites	d_S	d_N			
Entire *eae*	3	513	16.06 ± 1.06	8.55 ± 0.63	1.9	42.0	0.282
eae 5′ region	3	8	2.30 ± 0.79	0.00 ± 0.00	0	41.8	0.286
eae central region	3	12	1.50 ± 0.51	0.20 ± 0.11	7.6	45.2	0.320
eae 3′ region	3	493	41.88 ± 3.30	21.34 ± 1.59	2.0	39.7	0.253
cdtABC	3	17	1.10 ± 0.36	0.29 ± 0.11	3.8	43.3	0.173
cdtA	3	7	1.06 ± 0.60	0.46 ± 0.21	2.3	45.9	0.170
cdtB	3	6	0.98 ± 0.54	0.22 ± 0.14	4.4	44.0	0.174
cdtC	2	4	1.27 ± 0.80	0.08 ± 0.08	16.1	38.1	0.172
MLST genes	6[b]	82	1.34 ± 0.19	0.03 ± 0.02	44.1	52.3	0.448

[a] See Table 2, footnote *a*, for definitions.
[b] Number of distinct 14-locus sequence types.



C. Amino acid similarity (%)

Intimin variant (GenBank)	alpha	zeta	theta	lambda	gamma	epsilon	beta	iota2	19982	K-1	C-425	C.rod
alpha (AF022236)		77	66	68	66	65	67	62	78	82	62	68
zeta (AJ271407)	99		65	67	67	63	67	62	76	81	62	68
theta (AF449418)	99	98		68	83	66	66	64	67	66	64	63
lambda (AF530557)	99	99	99		70	71	69	65	73	69	65	68
gamma (AE005595)	99	98	99	99		69	69	63	67	67	63	67
epsilon (AY186750)	99	97	99	99	98		69	63	64	67	63	64
beta (AF081187)	98	97	98	98	98	100		63	66	65	63	82
iota2 (AF530553)	97	96	97	97	96	97	96		63	62	100	61
E. albertii 19982	97	96	97	97	96	97	96	100		76	63	65
S. boydii 7 K-1	96	96	96	96	96	96	96	99	99		62	67
S. boydii 13 C-425	96	95	96	96	96	96	96	99	99	99		61
C. rodentium (AB040740)	94	93	94	94	94	95	94	94	94	93	93	

FIG. 2. Intimin variants of the *E. albertii-Shigella* B13 lineage. (A) A neighbor-joining tree (MEGA2) based on nucleotide sequences of the *eae* central conserved region of intimins of *E. coli* shows that the sequences of the intimin gene homologues of *E.albertii-Shigella* B13 strains are most similar to the sequence of the *E. coli* iota2 allele. (B) Haplotype plots of the nucleotide differences in the 3′ part of the *eae* gene encoding the Int280 region. The sequences of *E. albertii* and *Shigella* B7 strains are similar to the sequence of the iota2 allele of *E. coli* in the central conserved region but are most similar to the sequences of the alpha and zeta alleles of *E. coli* in the Int280 region. (C) Percent amino acid similarities among representative intimin variants. The lower triangular matrix shows pairwise similarity values for the central conserved region, and the upper triangular matrix shows values for the Int280 region. The shading highlights close similarities to the intimin domains of the *E. albertii-Shigella* B13 lineage (tan). C.rod, *C. rodentium*.

tected for a strain that was found to be negative for *cdtB* by PCR. Although none of the strains tested was found to be positive for Shiga toxin (*stx*) by multilocus virulence gene profiling, the CDT results were verified by the method of Scott and Kaper (44) with Shiga toxin-insensitive CHO cells (data not shown).

Virulence gene profiling. To identify other potential virulence factors, 30 known virulence genes of *E. coli* and *Shigella* were tested for their presence in members of the *E. albertii* lineage (Table 4). As expected, strains of the *E. albertii* lineage were all positive for *eae*. In addition, members of the lineage were positive for at least one locus of *E. coli* O island 43: *ureA* (urease), *terC* (tellurite resistance), or *iha* (IrgA homologue adhesion). Most *E. albertii* strains were also positive for *toxB*, a toxin gene present on *E. coli* O157 plasmids. None of the chromosomal *E. coli* virulence loci examined was present in the strains examined, but *pic*, a secreted autotransporter toxin gene, and set_{1a}, an enterotoxin gene present on the *Shigella* chromosome, were present in some strains of the *E. albertii* lineage. Although neither of these loci was found in atypical *Shigella* B13 strains, the following were present: chromosomal *Shigella* virulence factor gene *sigA*, encoding a secreted autotransporter toxin; *shuA*, a heme receptor gene on the outer membrane of *S. dysenteriae*; *E. coli* plasmid-carried *ehx*, an enterohemolysin gene; *pet*, an *E. coli* secreted autotransporter toxin gene; and *astA*, the EAST1 heat-stable toxin gene present in duplicate on both the chromosome and a plasmid of pathogenic *E. coli*. Atypical *Shigella* B13 strain 5216-70, which is genetically related to a group of EIEC strains, was not positive for *eae*.

Diagnostic multiplex PCR. With the sequence data from the conserved housekeeping genes, we devised a multiplex PCR to detect conserved sequences that distinguish members of the *E. albertii* lineage from *E. coli* and *Shigella* strains. Candidate sequences with fixed differences between lineages were identified in the *lysP* and *mdh* loci. Based on nucleotide polymorphisms, primers lysP_107F and lysP_358R and primers mdh_50F and mdh_164R were designed to selectively amplify 252- and 115-bp regions, respectively, of the housekeeping genes *lysP* and *mdh* in the *E. albertii* lineage. As a positive control, prim-

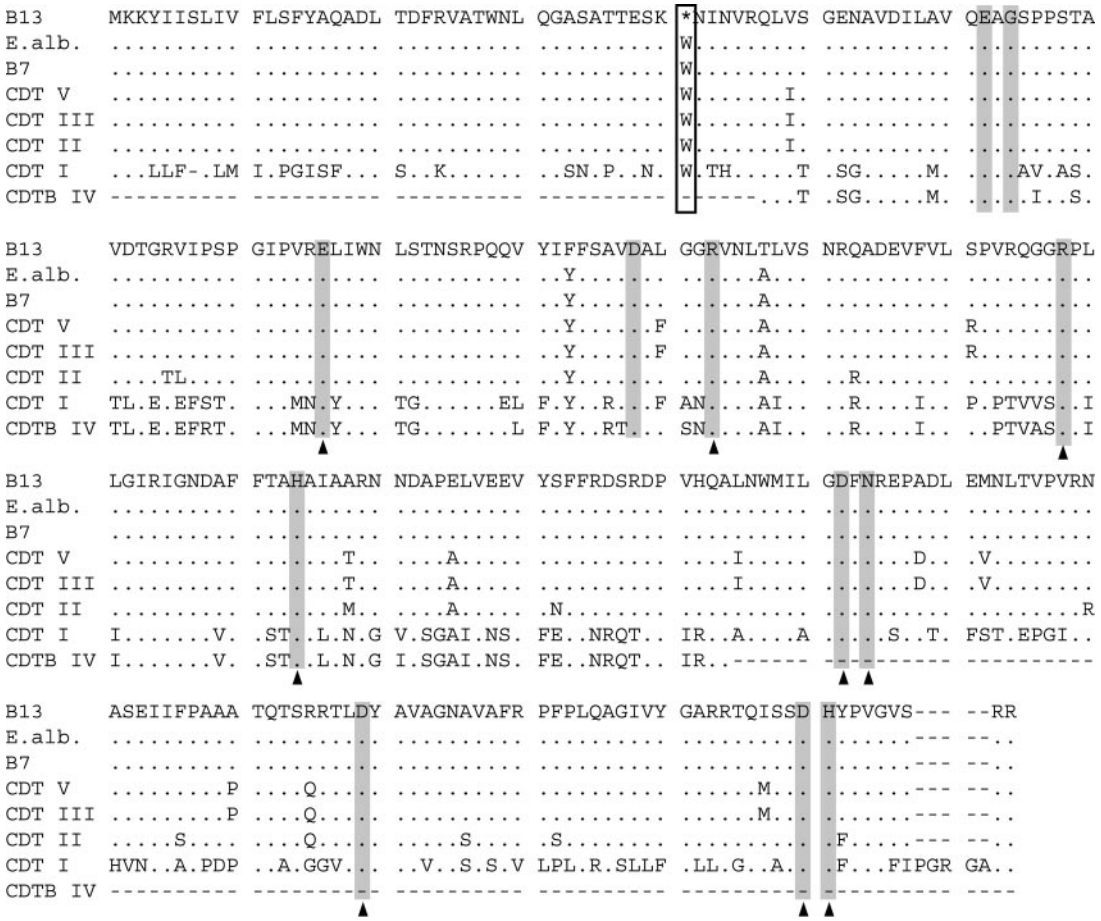


FIG. 3. Multiple sequence alignment of CdtBs from representatives of the five described *E. coli* alleles, CDT I (44), CDT II (37), CDT III (36), CDT IV (49), and CDT V (4, 22), and representatives of three distinct alleles of the *E. albertii* lineage (B13, E.alb., and B7). At amino acid position 41, shown here in a box, *Shigella* B13 strains are predicted to translate a stop codon rather than a tryptophan codon due to a single G-to-A transition at the second codon position. Conserved residues thought to be important for enzymatic activity (11, 35) are highlighted in gray. Arrowheads indicate residues that were mutated and shown to completely or partially abolish the cell cycle arrest activity of CDT (11, 35).

ers clpX_28F and clpX_411R were designed to amplify a region of *clpX* in all *E. coli* and *Shigella* strains as well as the *E. albertii* lineage. A total of 134 *E. coli* and *Shigella* strains representing all major pathotypes as well as all identified members of the *E. albertii* lineage were analyzed by PCR (data not shown). The multiplex PCR was tested and validated with all strains of the *E. albertii* lineage as well as representatives of *E. coli*, *Shigella*, and *Salmonella*.

DISCUSSION

Following their discovery in 1991, isolates of *eae*-positive *H. alvei* were described as distinct from *eae*-negative *H. alvei* both phenotypically and genotypically (2, 3, 19–21, 43). These unusual bacteria were unlike other diarrheagenic pathogens in that they did not produce heat-labile and heat-stable enterotoxins or Shiga toxins and were not enteroinvasive. However, it was shown that they were able to cause attaching-effacing lesions accompanied by diarrhea in rabbits, as well as the typical pedestal formation and actin polymerization in vitro (2, 3, 19). Twelve years after they were first discovered, the *eae*-positive *H. alvei* isolates were elevated to a new species of pathogen within the *Enterobacteriaceae* (18).

It has been recognized for nearly 50 years that *Shigella* B13 is highly divergent from other shigellae and has been classified incorrectly (5, 14, 31, 40). For example, in a DNA relatedness study, *Shigella* B13 was the only known group within the *E. coli-Shigella* complex that could be separated with certainty (5). Later studies based on sequence analysis of multiple housekeeping genes confirmed that *Shigella* B13 is highly divergent from *E. coli* and other *Shigella* serotypes (40). The key characteristic of typical *Shigella* strains is invasiveness due to the presence of a large invasion plasmid, pINV (16). Although *Shigella* B13 strains were originally considered to be *Shigella*, there has been no evidence that they are invasive, nor do they carry the known virulence genes present on pINV. Our analyses also failed to PCR amplify pINV genes in any of the *Shigella* B13 strains or *E. albertii* strains examined by multilocus virulence gene profiling.

We conclude, based on evolutionary analyses of multiple housekeeping genes, the 16S rRNA gene, and putative virulence genes, that *E. albertii* and *Shigella* B13 strains are monophyletic and represent a distinct lineage of enteric pathogens that separated from an ancient *Escherichia* ancestor after *E. coli* and *S. enterica* had diverged from a common ancestor.

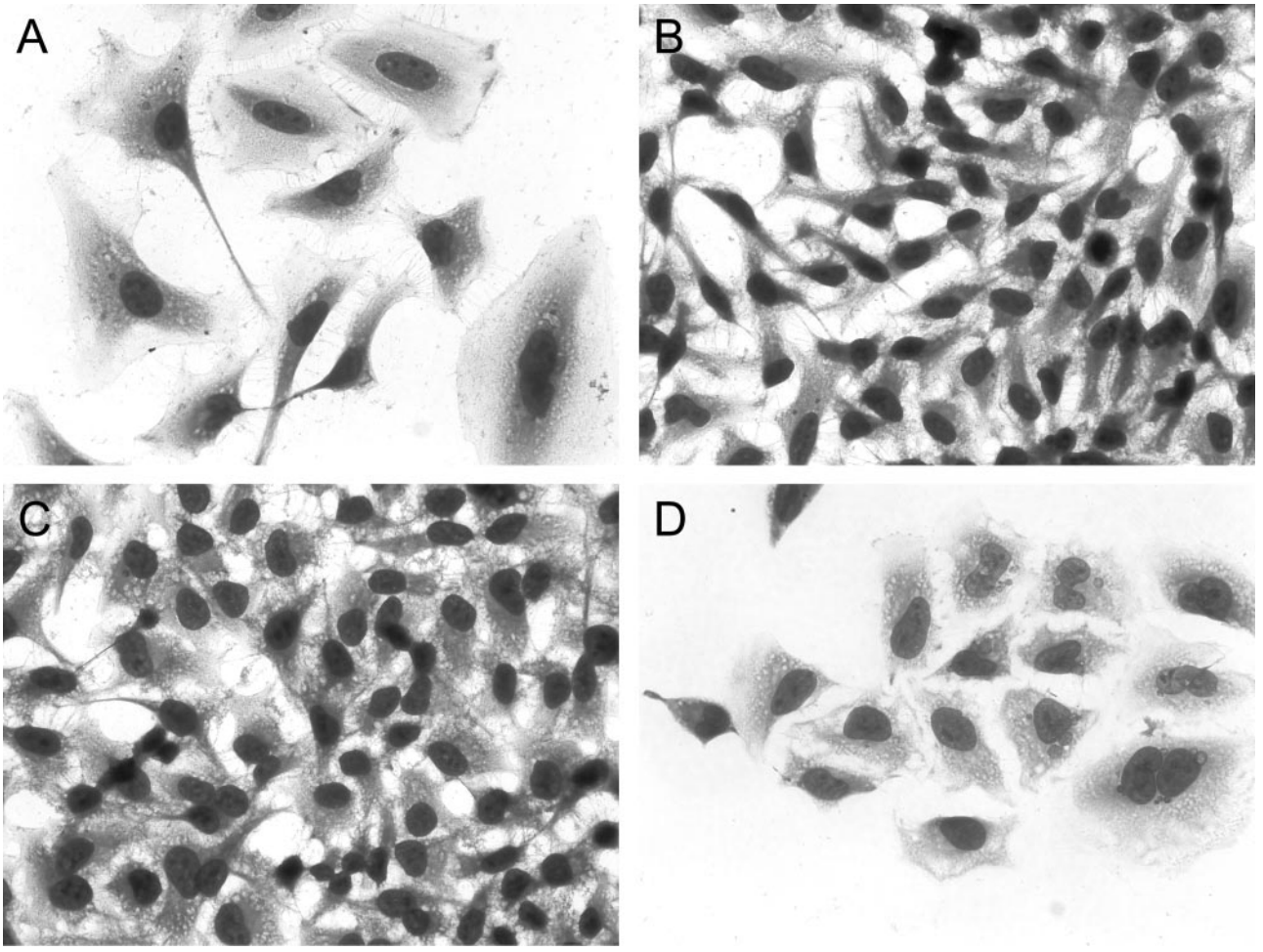


FIG. 4. Cytopathic effects of bacterial cell extracts on HeLa cell monolayers. Cytoplasmic distension and nuclear abnormalities were apparent 72 h after HeLa cells were exposed to *E. albertii* and *Shigella* B7-related strains (A), whereas *Shigella* B13 strains did not appear to produce active toxin (B). Included for comparison are CDT-negative *E. coli* K-12 (C) and CDT-positive *E. coli* 493/89 (D). Magnification, ×40.

On the basis of the rate of synonymous substitutions for *E. coli* and *S. enterica* (4.7×10^{-9} per site per year) (42), the lineage shared a common ancestor with *E. coli* ~28 million years ago. The molecular evolutionary analysis supports the proposal that certain strains, such as nontypeable strain 3097-02 and *Shigella* B7-related strain K-1, should also be included in the new species, *E. albertii*.

Although most *Shigella* B13 strains belong to the *E. albertii* lineage, certain *E. coli*-*Shigella*-like strains (atypical *Shigella* B13 strains) also express the *Shigella* B13 antigen, suggesting there has been either horizontal transfer of genes responsible for this phenotype or convergence to a similar antigenic phenotype. Atypical *Shigella* B13 strains 3556-77 and 3557-77 were also detected as intermediate between *E. coli* and *Shigella* B13 strains by DNA relatedness, with levels similar to those found between *E. coli* and *Shigella* B13 strain but lower than those found within other *Shigella* B13 strains (5). The chemical and genetic organization of the O antigen of *Shigella* B13 was recently described; based on nucleotide sequence data, the authors concluded that the *Shigella* B13 *rfb* gene cluster was recently assembled and has also undergone recent evolutionary change (15). The inclusion of nontypeable and *Shigella* B7-related strains in the *E. albertii* lineage and the evidence for change in the genes responsible for antigenic variations suggest that the lineage may be very diverse, but that diversity has probably been seriously underestimated because of sampling bias.

To explore putative virulence factors present in the *E. albertii* lineage, we described two genes that *E. albertii* strains are known to express, the intimin (*eae*) (3) and CDT (*cdt*) (38) genes, and screened the strains for the presence of other known *E. coli* and *Shigella* virulence factors. Phylogenetic analyses of virulence genes revealed three subclusters within the *E. albertii* lineage, consistent with phylogenetic analyses of multiple housekeeping genes, and supported the conclusion that these strains belong to an evolutionary branch of enteric pathogens that have further diverged into three distinct groups.

The intimin gene (*eae*), which was found to be present in all members of the *E. albertii* lineage by virulence gene profiling, was sequenced in representatives of the *E. albertii* subclusters. Phylogenetic analyses indicated that the sequences of *Shigella* B13 strains, which are identical, are similar to the sequence of the iota2 allele of pathogenic *E. coli*. The sequences of strains 19982 (*E. albertii*) and K-1 (*Shigella* B7 related) are similar to the sequence of the iota2 allele in the central conserved region

TABLE 4. Presence or absence of *E. coli* and *Shigella* virulence genes in the *E. albertii-Shigella* B13 lineage as detected by multilocus virulence gene profiling[a]

Strain	Lineage	Presence (+) or absence (−) of the indicated gene in:											
		E. coli								*Shigella* (chromosomal)			
		O43			LEE (*eae*)	c/p (*astA*)	pO157 (*toxB*)	Plasmid					
		ureA	*terC*	*iha*				*ehx*	*pet*	*sigA*	*pic*	*shuA*	set_{1a}
9194	*E. albertii*	+	+	−	+	−	−	−	−	−	−	−	−
10457	*E. albertii*	−	−	−	+	−	+	−	−	−	−	−	−
10790	*E. albertii*	+	+	−	+	−	+	−	−	−	−	−	+
12502	*E. albertii*	−	−	−	+	−	+	−	−	−	−	−	−
19982	*E. albertii*	+	−	−	+	−	+	−	−	−	−	−	−
2045-51	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	−
2046-51	*Shigella* B13	+	−	−	+	−	−	−	−	−	−	−	−
1610-55	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	−
603-73	*Shigella* B13	+	−	+	+	−	−	−	−	−	−	−	−
3552-77	*Shigella* B13	+	−	+	+	−	−	−	−	−	−	−	−
3553-77	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	−
3554-77	*Shigella* B13	+	−	+	+	−	−	−	−	−	−	−	−
3555-77	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	−
3103-99	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	+
3097-02	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	+
C-425	*Shigella* B13	−	−	+	+	−	−	−	−	−	−	−	+
616	*Shigella* B13	+	+	+	+	−	−	−	−	−	−	−	−
K-694	*Shigella* B13	−	−	+	+	−	−	−	−	−	+	−	+
K-1	*Shigella* B7	+	−	−	+	−	−	−	−	−	−	−	−
5216-70	Atypical *Shigella* B13	+	−	+	−	+	−	−	−	+	−	−	−
3556-77	Atypical *Shigella* B13	−	−	−	+	+	+	−	−	−	−	+	−
3557-77	Atypical *Shigella* B13	−	−	+	+	+	−	−	−	+	−	+	−
3052-94	Atypical *Shigella* B13	−	−	−	+	−	−	+	+	−	−	+	−
3053-94	Atypical *Shigella* B13	−	−	+	+	+	+	+	−	−	−	−	−

[a] Abbreviations: O43, O island 43 *E. coli* O157:H7 EDL-933; LEE, locus of enterocyte effacement; c/p, duplicate copies exist, one chromosomal and one carried on a plasmid. All strains were found to be negative by PCR for *espP*, *espC*, *sepA*, *sat*, stx_1, stx_2, *saa*, *spaP*, *invG*, *shuA*, *cdtA*, irp_2, *fyuA*, *senA*, *estA*, *bfpA*, and *elt*.

but have different alleles most similar to the alpha and zeta alleles of *E. coli* in the Int280 region.

CDT activity does not appear to be required for the virulence of *E. albertii* and its relatives. In two *Shigella* B13 strains (3553-77 and 3555-77), originally isolated from human patients, the *cdtB* gene was not amplified by PCR with degenerate primers; the remainder of the *Shigella* B13 strains contained a putative stop codon and lacked CDT activity. The silencing of CDT expression suggests that it does not have a required role in the virulence of these organisms. The short evolutionary distance between the active *E. albertii cdt* alleles and the inactive *Shigella* B13 *cdt* alleles suggests that the inactivation was a recent evolutionary event.

Both virulence genes examined have a lower GC content and a lower codon adaptation index, as measured by the rate of amelioration of GC content at the third codon positions, than do housekeeping genes (Table 3). The low codon adaptation index suggests recent acquisition of the virulence genes by lateral gene transfer. In particular, *cdt* sequences have a very low codon adaptation index. Levels of polymorphism within the *E. albertii* lineage are similar between the virulence and housekeeping genes, with the exception of the 3′ Int280 region of *eae*. The sequence diversity of this extracellular region in *E. coli* strains is thought to be under positive selection (33), and the sequence diversity in the *E. albertii* lineage supports the hypothesis that in the *E. albertii* lineage, as in *E. coli*, positive selection drives sequence divergence in this region.

To facilitate the economical screening of isolates, a diagnostic multiplex PCR was designed to identify bacteria of the *E. albertii* lineage based on sequence variations in housekeeping genes. This method of classification is independent of biochemical and antigenic phenotypes and therefore may provide a useful starting point for discovering genetic and antigenic diversity within the lineage. Infection by members of the *E. albertii* lineage may contribute to unexplained diarrhea, and molecular markers for the rapid detection of the new lineage will be crucial to further elucidating the role of these novel organisms in enteric disease.

ACKNOWLEDGMENTS

We thank Lindsey Ouellette for technical assistance and Kaisar Talukder, International Center for Diarrheal Disease Research, for kindly supplying several strains used in this study.

This project was funded in part with an American Society for Microbiology undergraduate research fellowship to K.E.H. and in part with federal funds from the NIAID, NIH, and DHHS under NIH research contract N01-AI-30058.

REFERENCES

1. **Adu-Bobie, J., G. Frankel, C. Bain, A. G. Goncalves, L. R. Trabulsi, G. Douce, S. Knutton, and G. Dougan.** 1998. Detection of intimins alpha, beta, gamma, and delta, four intimin derivatives expressed by attaching and effacing microbial pathogens. J. Clin. Microbiol. **36:**662–668.
2. **Albert, M. J., K. Alam, M. Islam, J. Montanaro, A. S. Rahaman, K. Haider, M. A. Hossain, A. K. Kibriya, and S. Tzipori.** 1991. *Hafnia alvei*, a probable cause of diarrhea in humans. Infect. Immun. **59:**1507–1513.
3. **Albert, M. J., S. M. Faruque, M. Ansaruzzaman, M. M. Islam, K. Haider, K. Alam, I. Kabir, and R. Robins-Browne.** 1992. Sharing of virulence-associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. J. Med. Microbiol. **37:**310–314.
4. **Bielaszewska, M., M. Fell, L. Greune, R. Prager, A. Fruth, H. Tschape, M. A. Schmidt, and H. Karch.** 2004. Characterization of cytolethal distending toxin

genes and expression in Shiga toxin-producing *Escherichia coli* strains of non-O157 serogroups. Infect. Immun. **72:**1812–1816.
5. **Brenner, D. J., A. G. Steigerwalt, H. G. Wathen, R. J. Gross, and B. Rowe.** 1982. Confirmation of aerogenic strains of *Shigella boydii* 13 and further study of *Shigella* serotypes by DNA relatedness. J. Clin. Microbiol. **16:**432–436.
6. **Bumbaugh, A. C., K. E. Hyma, D. W. Lacher, T. M. Large, L. M. Ouellette, C. L. Tarr, N. A. Strockbine, K. A. Talukder, D. A. Sack, and T. S. Whittam.** Unpublished data.
7. **Chien, C. C., N. S. Taylor, Z. Ge, D. B. Schauer, V. B. Young, and J. G. Fox.** 2000. Identification of *cdtB* homologues and cytolethal distending toxin activity in enterohepatic *Helicobacter* spp. J. Med. Microbiol. **49:**525–534.
8. **Cope, L. D., S. Lumbley, J. L. Latimer, J. Klesney-Tait, M. K. Stevens, L. S. Johnson, M. Purven, R. S. Munson, Jr., T. Lagergard, J. D. Radolf, and E. J. Hansen.** 1997. A diffusible cytotoxin of *Haemophilus ducreyi*. Proc. Natl. Acad. Sci. USA **94:**4056–4061.
9. **Cortes-Bratti, X., T. Frisan, and M. Thelestam.** 2001. The cytolethal distending toxins induce DNA damage and cell cycle arrest. Toxicon **39:**1729–1736.
10. **Deng, K., and E. J. Hansen.** 2003. A CdtA-CdtC complex can block killing of HeLa cells by *Haemophilus ducreyi* cytolethal distending toxin. Infect. Immun. **71:**6633–6640.
11. **De Rycke, J., and E. Oswald.** 2001. Cytolethal distending toxin (CDT): a bacterial weapon to control host cell proliferation? FEMS Microbiol Lett. **203:**141–148.
12. **Elwell, C. A., and L. A. Dreyfus.** 2000. DNase I homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest. Mol. Microbiol. **37:**952–963.
13. **Eserstam, R., T. P. Rajaguru, P. E. Jansson, A. Weintraub, and M. J. Albert.** 2002. The structure of the O-chain of the lipopolysaccharide of a prototypal diarrheagenic strain of *Hafnia alvei* that has characteristics of a new species under the genus *Escherichia*. Eur. J. Biochem. **269:**3289–3295.
14. **Ewing, W. H., R. W. Reavis, and B. R. Davis.** 1958. Provisional *Shigella* serotypes. Can. J. Microbiol. **4:**89–107.
15. **Feng, L., S. N. Senchenkova, J. Yang, A. S. Shashkov, J. Tao, H. Guo, G. Zhao, Y. A. Knirel, P. Reeves, and L. Wang.** 2004. Structural and genetic characterization of the *Shigella boydii* type 13 O antigen. J. Bacteriol. **186:**383–392.
16. **Hale, T. L.** 1991. Genetic basis of virulence in *Shigella* species. Microbiol. Rev. **55:**206–224.
17. **Hassane, D. C., R. B. Lee, and C. L. Pickett.** 2003. *Campylobacter jejuni* cytolethal distending toxin promotes DNA repair responses in normal human cells. Infect. Immun. **71:**541–545.
18. **Huys, G., M. Cnockaert, J. M. Janda, and J. Swings.** 2003. *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. Int. J. Syst. Evol. Microbiol. **53:**807–810.
19. **Ismaili, A., B. Bourke, J. C. de Azavedo, S. Ratnam, M. A. Karmali, and P. M. Sherman.** 1996. Heterogeneity in phenotypic and genotypic characteristics among strains of *Hafnia alvei*. J. Clin. Microbiol. **34:**2973–2979.
20. **Janda, J. M., S. L. Abbott, and M. J. Albert.** 1999. Prototypal diarrheagenic strains of *Hafnia alvei* are actually members of the genus *Escherichia*. J. Clin. Microbiol. **37:**2399–2401.
21. **Janda, J. M., S. L. Abbott, S. Khashe, and W. Probert.** 2002. Phenotypic and genotypic properties of the genus *Hafnia*. J. Med. Microbiol. **51:**575–580.
22. **Janka, A., M. Bielaszewska, U. Dobrindt, L. Greune, M. A. Schmidt, and H. Karch.** 2003. Cytolethal distending toxin gene cluster in enterohemorrhagic *Escherichia coli* O157:H$^-$ and O157:H7: characterization and evolutionary considerations. Infect. Immun. **71:**3634–3638.
23. **Johnson, W. M., and H. Lior.** 1988. A new heat-labile cytolethal distending toxin (CLDT) produced by *Campylobacter* spp. Microb. Pathog. **4:**115–126.
24. **Johnson, W. M., and H. Lior.** 1988. A new heat-labile cytolethal distending toxin (CLDT) produced by *Escherichia coli* isolates from clinical material. Microb. Pathog. **4:**103–113.
25. **Jores, J., K. Zehmke, J. Eichberg, L. Rumer, and L. H. Wieler.** 2003. Description of a novel intimin variant (type zeta) in the bovine O84:NM verotoxin-producing *Escherichia coli* strain 537/89 and the diagnostic value of intimin typing. Exp. Biol. Med. (Maywood) **228:**370–376.
26. **Lan, R., and P. Reeves.** 2002. *Escherichia coli* in disguise: molecular origins of *Shigella*. Microbes Infect. **4:**1125.
27. **Lara-Tejero, M., and J. E. Galan.** 2000. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. Science **290:**354–357.
28. **Lara-Tejero, M., and J. E. Galan.** 2001. CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. Infect. Immun. **69:**4358–4365.
29. **Lee, R. B., D. C. Hassane, D. L. Cottle, and C. L. Pickett.** 2003. Interactions of *Campylobacter jejuni* cytolethal distending toxin subunits CdtA and CdtC with HeLa cells. Infect. Immun. **71:**4883–4890.
30. **Luo, Y., E. A. Frey, R. A. Pfuetzner, A. L. Creagh, D. G. Knoechel, C. A. Haynes, B. B. Finlay, and N. C. Strynadka.** 2000. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. Nature **405:**1073–1077.
31. **Manolov, D. G.** 1959. A new type of the genus *Shigella*—"*Shigella* 13." J. Hyg. Epidemiol. Microbiol. Immunol. **3:**184–190.
32. **Mao, X., and J. M. DiRienzo.** 2002. Functional studies of the recombinant subunits of a cytolethal distending holotoxin. Cell. Microbiol. **4:**245–255.
33. **McGraw, E. A., J. Li, R. K. Selander, and T. S. Whittam.** 1999. Molecular evolution and mosaic structure of alpha, beta, and gamma intimins of pathogenic *Escherichia coli*. Mol. Biol. Evol. **16:**12–22.
34. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
35. **Nesic, D., Y. Hsu, and C. E. Stebbins.** 2004. Assembly and function of a bacterial genotoxin. Nature **429:**429–433.
36. **Peres, S. Y., O. Marches, F. Daigle, J. P. Nougayrede, F. Herault, C. Tasca, J. De Rycke, and E. Oswald.** 1997. A new cytolethal distending toxin (CDT) from *Escherichia coli* producing CNF2 blocks HeLa cell division in G2/M phase. Mol. Microbiol. **24:**1095–1107.
37. **Pickett, C. L., D. L. Cottle, E. C. Pesci, and G. Bikah.** 1994. Cloning, sequencing, and expression of the *Escherichia coli* cytolethal distending toxin genes. Infect. Immun. **62:**1046–1051.
38. **Pickett, C. L., R. B. Lee, A. Eyigor, B. Elitzur, E. M. Fox, and N. A. Strockbine.** 2004. Patterns of variations in *Escherichia coli* strains that produce cytolethal distending toxin. Infect. Immun. **72:**684–690.
39. **Pickett, C. L., E. C. Pesci, D. L. Cottle, G. Russell, A. N. Erdem, and H. Zeytin.** 1996. Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* sp. *cdtB* gene. Infect. Immun. **64:**2070–2078.
40. **Pupo, G. M., R. Lan, and P. R. Reeves.** 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proc. Natl. Acad. Sci. USA **97:**10567–10572.
41. **Ramachandran, V., K. Brett, M. A. Hornitzky, M. Dowton, K. A. Bettelheim, M. J. Walker, and S. P. Djordjevic.** 2003. Distribution of intimin subtypes among *Escherichia coli* isolates from ruminant and human sources. J. Clin. Microbiol. **41:**5022–5032.
42. **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.
43. **Ridell, J., A. Siitonen, L. Paulin, O. Lindroos, H. Korkeala, and M. J. Albert.** 1995. Characterization of *Hafnia alvei* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. J. Clin. Microbiol. **33:**2372–2376.
44. **Scott, D. A., and J. B. Kaper.** 1994. Cloning and sequencing of the genes encoding *Escherichia coli* cytolethal distending toxin. Infect. Immun. **62:**244–251.
45. **Suau, A., R. Bonnet, M. Sutren, J. J. Godon, G. R. Gibson, M. D. Collins, and J. Dore.** 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. **65:**4799–4807.
46. **Sugai, M., T. Kawamoto, S. Y. Peres, Y. Ueno, H. Komatsuzawa, T. Fujiwara, H. Kurihara, H. Suginaka, and E. Oswald.** 1998. The cell cycle-specific growth-inhibitory factor produced by *Actinobacillus actinomycetemcomitans* is a cytolethal distending toxin. Infect. Immun. **66:**5008–5019.
47. **Tarr, C. L., T. M. Large, C. L. Moeller, D. W. Lacher, P. I. Tarr, D. W. Acheson, and T. S. Whittam.** 2002. Molecular characterization of a serotype O121:H19 clone, a distinct Shiga toxin-producing clone of pathogenic *Escherichia coli*. Infect. Immun. **70:**6853–6859.
48. **Tarr, C. L., and T. S. Whittam.** 2002. Molecular evolution of the intimin gene in O111 clones of pathogenic *Escherichia coli*. J. Bacteriol. **184:**479–487.
49. **Toth, I., F. Herault, L. Beutin, and E. Oswald.** 2003. Production of cytolethal distending toxins by pathogenic *Escherichia coli* strains isolated from human and animal sources: establishment of the existence of a new *cdt* variant (type IV). J. Clin. Microbiol. **41:**4285–4291.
50. **Young, V. B., C. C. Chien, K. A. Knox, N. S. Taylor, D. B. Schauer, and J. G. Fox.** 2000. Cytolethal distending toxin in avian and human isolates of *Helicobacter pullorum*. J. Infect. Dis. **182:**620–623.
51. **Young, V. B., K. A. Knox, J. S. Pratt, J. S. Cortez, L. S. Mansfield, A. B. Rogers, J. G. Fox, and D. B. Schauer.** 2004. In vitro and in vivo characterization of *Helicobacter hepaticus* cytolethal distending toxin mutants. Infect. Immun. **72:**2521–2527.
52. **Young, V. B., K. A. Knox, and D. B. Schauer.** 2000. Cytolethal distending toxin sequence and activity in the enterohepatic pathogen *Helicobacter hepaticus*. Infect. Immun. **68:**184–191.
53. **Yu, J., and J. B. Kaper.** 1992. Cloning and characterization of the *eae* gene of enterohaemorrhagic *Escherichia coli* O157:H7. Mol. Microbiol. **6:**411–417.
54. **Zhang, W. L., B. Kohler, E. Oswald, L. Beutin, H. Karch, S. Morabito, A. Caprioli, S. Suerbaum, and H. Schmidt.** 2002. Genetic diversity of intimin genes of attaching and effacing *Escherichia coli* strains. J. Clin. Microbiol. **40:**4486–4492.