Molecular Characterization of Human Atypical Sorbitol-Fermenting Enteropathogenic *Escherichia coli* O157 Reveals High Diversity

Annelene Kossow,[a] Wenlan Zhang,[a] Martina Bielaszewska,[a] Sophie Rhode,[a] Kevin Hansen,[a] Angelika Fruth,[b] Christian Rüter,[c] Helge Karch,[a] Alexander Mellmann[a]

Institute of Hygiene and National Consulting Laboratory for Hemolytic Uremic Syndrome (HUS), University Hospital Muenster, Münster, Germany[a]; National Reference Centre for Salmonella and Other Bacterial Enterics, Robert Koch-Institute, Wernigerode, Germany[b]; Institute of Infectiology, Center for Molecular Biology of Inflammation (ZMBE), University of Münster, Münster, Germany[c]

Alongside the well-characterized enterohemorrhagic *Escherichia coli* (EHEC) O157:H7, serogroup O157 comprises sorbitol-fermenting typical and atypical enteropathogenic *E. coli* (EPEC/aEPEC) strains that carry the intimin-encoding gene *eae* but not Shiga toxin-encoding genes (*stx*). Since little is known about these pathogens, we characterized 30 clinical isolates from patients with hemolytic uremic syndrome (HUS) or uncomplicated diarrhea with respect to their flagellin gene (*fliC*) type and multilocus sequence type (MLST). Moreover, we applied whole-genome sequencing (WGS) to determine the phylogenetic relationship with other *eae*-positive EHEC serotypes and the composition of the *rfb*O157 region. *fliC* typing resulted in five *fliC* types (H7, H16, H34, H39, and H45). Isolates of each *fliC* type shared a unique ST. In comparison to the 42 HUS-associated *E. coli* (HUSEC) strains, only the *stx*-negative isolates with *fliC*H7 shared their ST with EHEC O157:H7/H⁻ strains. With the exception of one O157:H⁻ *fliC*H16 isolate, HUS was exclusively associated with *fliC*H7. WGS corroborated the separation of the *fliC*H7 isolates, which were closely related to the EHEC O157:H7/H⁻ isolates, and the diverse group of isolates exhibiting different *fliC* types, indicating independent evolution of the different serotypes. This was also supported by the heterogeneity within the *rfb*O157 region that exhibited extensive recombinations. The genotypic subtypes and distribution of clinical symptoms suggested that the *stx*-negative O157 strains with *fliC*H7 were originally EHEC strains that lost *stx*. The remaining isolates form a distinct and diverse group of atypical EPEC isolates that do not possess the full spectrum of virulence genes, underlining the importance of identifying the H antigen for clinical risk assessment.

A longside the classical non-sorbitol-fermenting and sorbitol-fermenting (SF) Shiga toxin (Stx)-producing *Escherichia coli* O157:H7/H⁻ strains (nonmotile), which cause severe human diseases ranging from diarrhea to hemolytic uremic syndrome (HUS) and are termed enterohemorrhagic *E. coli* (EHEC) (1), the *E. coli* serogroup O157 also comprises members of other distinct pathogroups, including enterotoxigenic *E. coli* (ETEC) and classical and atypical enteropathogenic *E. coli* (EPEC) (2–7). EPEC and atypical EPEC O157 are both characterized by the presence of the *eae* gene, which is a part of the locus of enterocyte effacement, but atypical EPEC O157 strains lack the *bfpA* gene (5). The *eae* gene is also present in EHEC O157:H7/H⁻ (8); EPEC O157 strains lack, however, the Stx characteristic of EHEC O157 (4, 5). Whereas EHEC O157 strains are the major cause of HUS in children (9–11), EPEC O157 strains mostly cause diarrhea (2, 5). Occasionally, EPEC O157 strains were isolated from patients with HUS (12, 13). Belonging, like the EHEC O157 strains, to the multilocus sequence typing (MLST) clonal complex (CC) 11, they were assumed to be EHEC strains that lost the Stx-encoding gene (EHEC-LST) (12).

Little is known about the characteristics of other atypical SF EPEC O157 strains. We therefore characterized 30 clinical isolates of SF *eae*-positive, *stx*-negative O157 with respect both to their genomic background and to the composition of the *rfb*O157 region, which is responsible for the O157 antigen phenotype.

MATERIALS AND METHODS

The strains were obtained from stool samples of patients with HUS (*n* = 16) or uncomplicated (nonbloody) diarrhea (*n* = 14). A subset of strains was described in earlier studies (5, 12–15). The strains were epidemiologically unrelated. Serotypes and the presence of *rfb*EO157 were deter-

mined and *fliC* subtyping was performed as previously described (16–18). Sorbitol fermentation was detected on sorbitol MacConkey agar (SMAC), production of EHEC hemolysin on enterohemolysin agar (Sifin, Berlin, Germany), and β-D-glucuronidase activity on nutrient agar with 4-methylumbelliferyl-β-D-glucuronide (MUG) (Becton Dickinson, Heidelberg, Germany). Published PCR protocols were applied to detect genes encoding cytolethal distending toxin V (*cdtV*), EHEC *hlyA*, *eae* and its subtypes, and *sfpA* encoding the major subunit of Sfp fimbriae and to analyze occupation of the *stx*-encoding bacteriophage integration site *yecE* (19–25).

For MLST, the internal fragments of seven housekeeping genes were Sanger sequenced (26) and sequence types (STs) were assigned according to the *E. coli* MLST website (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli). Related STs were grouped into CCs in accordance with the MLST website and, for comparison, data from the HUS-associated enterohemorrhagic *E. coli* collection (HUSEC collection) (27) (http://www.ehec.org) were

TABLE 1 Genotypic and phenotypic characteristics of 30 sorbitol-fermenting *stx*-negative, *eae*-positive *E. coli* O157 human isolates

Serotype[a]	Total no. of patients (patients with HUS)	*fliC* type by RFLP	*eae* type	No. of strains with characteristic of:			EHEC Hly phenotype	β-Glucuronidase production
				EHEC *hlyA*	*sfpA*	*cdtV*		
O157:H⁻	23 (15)	H7	γ	18	23	22	2	23
O157:H⁻/Hnt	3 (1)	H16	ε	0	0	0	0	3
O157:H34	1 (0)	H34	ε	0	0	0	0	1
O157:H39	2 (0)	H39	κ	0	0	0	0	0
O157:H45	1 (0)	H45	α	0	0	0	0	1

[a] H⁻, nonmotile; Hnt, not typeable with available antisera.

used. The minimum spanning tree was generated using SeqSphere⁺ software, version 1 (Ridom GmbH, Münster, Germany).

For whole-genome shotgun sequencing of selected O157 strains representative of the different *fliC* types, sequencing libraries were prepared using Nextera XT chemistry (Illumina, Inc., San Diego, CA, USA) for either a 100-bp or a 250-bp paired-end sequencing run on a HiScanSQ or MiSeq sequencer in accordance with the manufacturer's recommendations (Illumina). After quality trimming using the default parameters of CLC Genomics Workbench software, version 6 (CLC bio, Arhus, Denmark), the sequencing reads were assembled using the CLC Genomics Workbench *de novo* assembler (CLC bio). Gene sequences for subsequent analyses were extracted from contigs using SeqSphere⁺ software, version 2.0 beta (Ridom GmbH). For the gene-by-gene core genome analysis as described previously (28), we included all genes present in all strains analyzed (see Table S1 in the supplemental material) and displayed them in a minimum spanning tree using SeqSphere⁺. For comparison, a total of 40 published genome sequences of the major *eae*-positive EHEC serogroups (O26 strain 11368 [GenBank accession no. NC_013361], O103 strain 13353 [NC_013353], O111 strain 13364 [NC_013364], and O157 strain Sakai [NC_002695]), of other published *eae*-positive (*stx*-negative and *stx*-positive) O157 genome sequences (29–32), of other intestinal pathogenic *E. coli* strains (EAEC strains 55989 [NC_011748] and 042 [NC_017626], EPEC strain E2348/69 [NC_011601], ETEC strain E24377A [NC_009801], EIEC strain 53638 [AAKB02000000], and AIEC strain LF82 [CU651637]), and of the commensal *E. coli* strain MG1655 (NC_000913) were included.

For characterization of the *rfb*O157 region, we queried for all 208 genes of the *rfb*O157 region (genes with locus tags ECs2800 to ECs3023) (32) using EHEC O157:H7 strain Sakai (GenBank accession no. NC_002695) as reference. Only genes present in all strains were further analyzed. After concatenation, a multiple alignment was created using MAFFT, version 7 (33), and all variants and putative recombinatory regions were determined using Gubbins (34). Finally, the comparison of the concatenated *rfb*O157 region was visualized using GenVision software, version 12 (DNAStar Inc., Madison, WI, USA).

Nucleotide sequence accession number. All generated raw reads were submitted to the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under study accession no. PRJEB9310.

RESULTS

Phenotypic and genotypic characteristics of *stx*-negative SF *E. coli* O157 strains. Among 30 *stx*-negative SF *E. coli* O157 strains isolated from patients, five different serotypes were identified (Table 1). The *fliC* restriction fragment length polymorphism (RFLP) analysis demonstrated that 23 of the O157:H⁻ (nonmotile) strains had *fliC*H7 and one additional O157:H⁻ strain and two O157:Hnt (nontypeable) strains possessed *fliC*H16; in strains of serotypes O157:H34, O157:H39, and O157:H45, the *fliC* types corresponded to H antigens determined by conventional serotyping (Table 1). All strains contained the *eae* gene, the type of which is shown in Table 1. All strains lacked the *bfpA* gene and thereby

fulfilled the definition of atypical EPEC. Moreover, *yecE*, the common *stx*-encoding bacteriophage integration site in O157, was empty in all strains. All 23 O157:H⁻$_{fliCH7}$ strains possessed *sfpA*, which encodes the major subunit of the SfpA fimbriae typically found in SF EHEC O157:H⁻ strains (20). Moreover, all but one possessed the *cdtV* cluster encoding cytolethal distending toxin V (Table 1), another hallmark of SF EHEC O157:H⁻ strains (21). Eighteen (78%) O157:H⁻$_{fliCH7}$ strains possessed EHEC *hlyA* encoding EHEC hemolysin, but only two expressed the enterohemolytic phenotype (Table 1). The strains of all serotypes except for O157:H39 produced β-D-glucuronidase (Table 1).

MLST. MLST of strains with the five different *fliC* types (H7, H16, H34, H39, and H45) resulted in five different STs: isolates of each *fliC* type shared a unique ST (Fig. 1). Whereas the isolates of *fliC* types H16 and H34 (ST10 and ST752), were closely related based on their allelic profiles, sharing six of the seven housekeeping genes and therefore belonging to the same CC (CC10), there was no close relationship among the other *stx*-negative O157 isolates sharing at maximum one of the seven genes analyzed (Fig. 1). Based on the allelic profiles of the isolates investigated, Fig. 1 further displays the phylogenetic relationships of the different O157 isolates to the 42 HUSEC strains that represent the phylogenetic breadth of HUS-associated EHEC strains (27) in a minimum spanning tree. Whereas the *stx*-negative isolates with *fliC*H7 shared their ST with *stx*-positive O157:H7/H⁻ strains, namely, ST11 (CC11), and the *stx*-negative isolates of serotypes O157:H16 and O157:H34 shared the same CC (CC10) with two HUSEC strains (HUSEC001: ST43, serotype O111:H10; and HUSEC002: ST330, serotype Ont:H⁻), the remaining *stx*-negative O157 isolates of serotypes O157:H39 and O157:H45 were only distantly related to any HUSEC strain (Fig. 1).

Whole-genome sequencing and characterization of *rfb*O157 region. To elucidate in depth the phylogenetic relationship of the five different O157 serotypes, we performed whole-genome sequencing (WGS) of one representative strain for each *fliC* type and subsequently compared the coding regions. Moreover, we included 33 *eae*-positive EHEC and EPEC isolates as well as 6 *E. coli* strains from other intestinal pathotypes and a commensal *E. coli* strain. Among the strains (*n* = 45) investigated, 2,415 genes were present in all (see Table S1 in the supplemental material). Analysis of the clonal relationship based on the allelic profiles derived from the 2,415 gene sequences per strain showed a high diversity both within the O157 strains and in comparison to other prototypical *E. coli* genome sequences of different pathotypes (Fig. 2). Specifically, the O157 isolates of the different serotypes differed by at least 392 alleles in a pairwise comparison (isolates of serotype O157:H16 [8350/98] and O157:H34 [06-8499]). The isolate 3196/
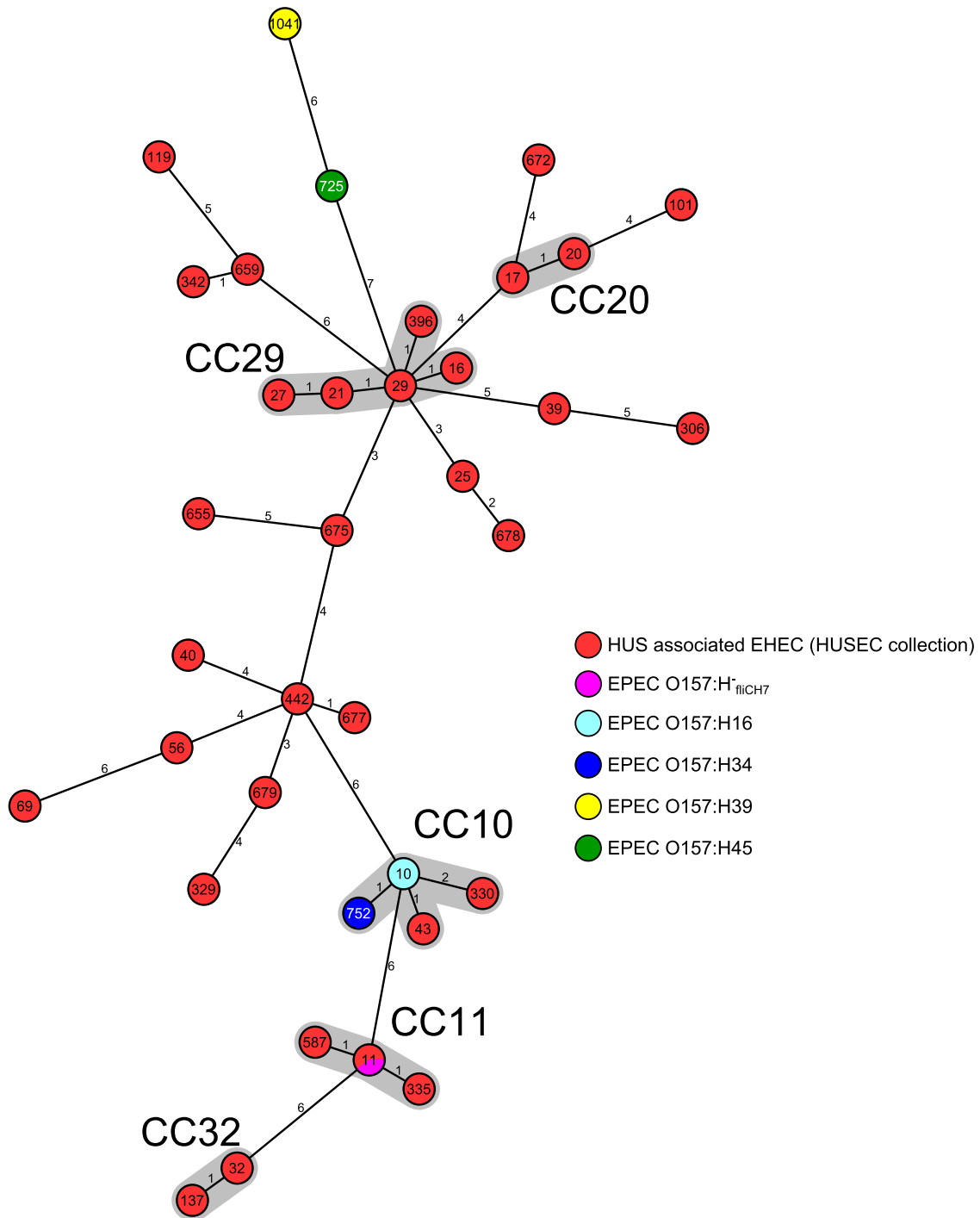
FIG 1 Phylogenetic relationships of *stx*-negative, *eae*-positive *E. coli* O157 to the HUS-associated enterohemorrhagic *E. coli* collection (HUSEC collection). Minimum spanning tree based on MLST allelic profiles portraying the clonal relationships of *stx*-negative, *eae*-positive non-H7 *E. coli* O157 strains to the HUS-associated enterohemorrhagic *E. coli* collection (HUSEC collection) (27) (http://www.ehec.org). Each circle represents a given ST. The numbers on the connecting lines illustrate the number of differing alleles, and the different groups of strains are distinguished by the colors of the circles. If assigned at the MLST database (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli), clonal complexes (CC) are given and shaded in gray.

99, which was O157:H⁻$_{fliCH7}$, clustered together with all other SFO157:H⁻$_{fliCH7}$ strains, which is in accordance with the presence of *fliCH7* in this EPEC strain. Moreover, this clustering was independent of the presence or absence of *stx*. All other O157:H7/ H⁻$_{fliCH7}$ strains also clustered together, depending on their ability to ferment sorbitol, which corroborates the previous finding of an early branching of SF and non-SF O157:H7/H⁻$_{fliCH7}$ strains (32). The three strains of O157:H16 also clustered together (Fig. 2). A similar clustering of the same serotype was also detected among the O157:H45 isolates; in addition to the three isolates of serotype
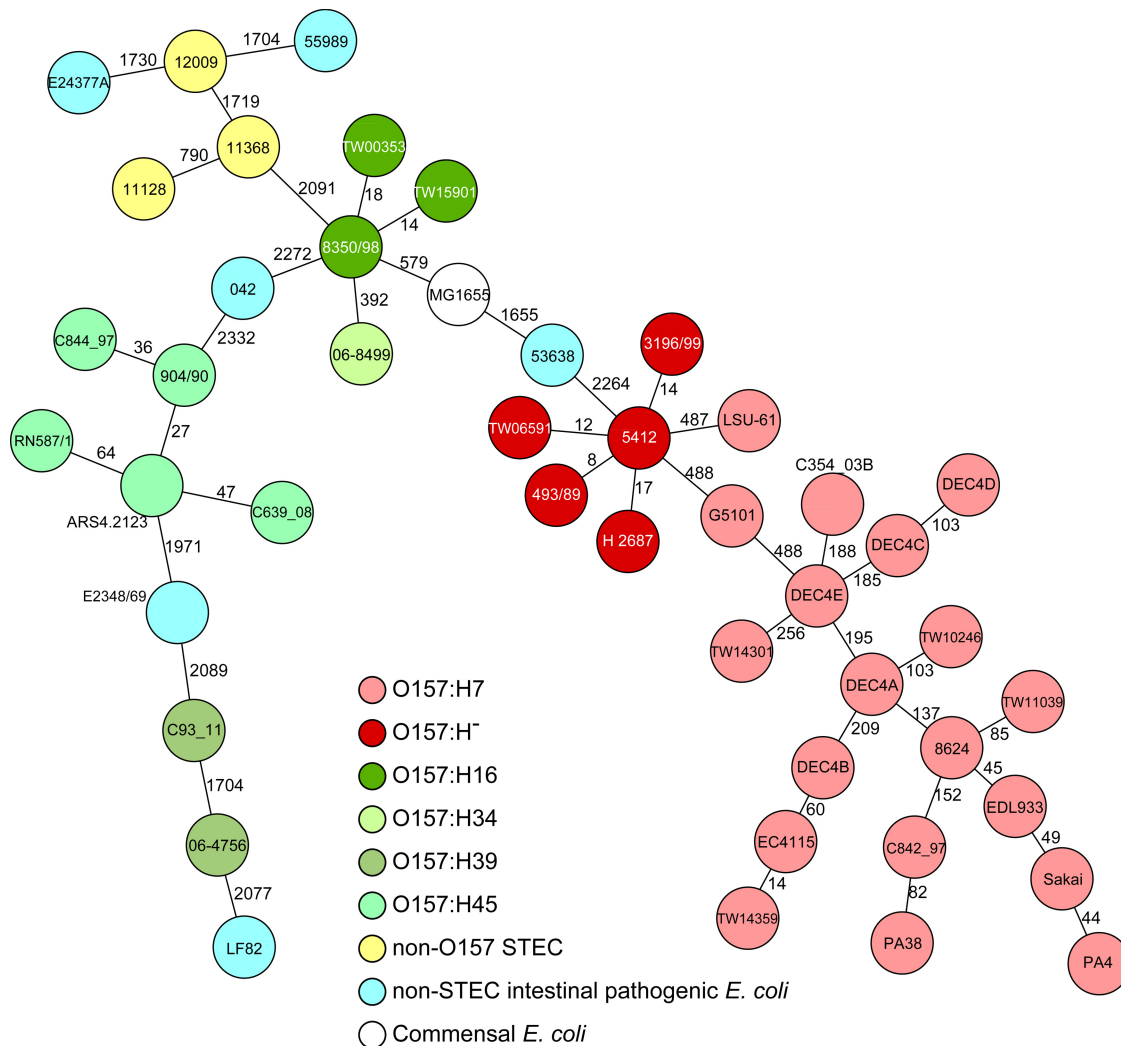
FIG 2 Whole-genome relationship of *stx*-negative, *eae*-positive *E. coli* O157 strains in comparison to other intestinal pathogenic *E. coli* pathotypes. The minimum-spanning tree is based on genome-wide allelic profiles (all 2,415 genes that were present in all 45 strains were included). Each circle represents a given allelic profile based on 2,415 target genes. The numbers on the connecting lines illustrate the number of differing alleles, and the different pathotypes and serotypes, including a commensal *E. coli* strain, are distinguished by the colors of the circles.

O157:H45, two further isolates (RN587/1, serotyped as O157:H8, and ARS4.2123, serotyped as O157:H16 [29]) clustered closely, suggesting a wrong H antigen type. Indeed, the *fliC* allele of RN587/1 was 100% identical to those of the other three isolates and exhibited the H45 genotype, and the *fliC* allele of ARS4.2123 only showed a 6-bp deletion (starting at position 518). Interestingly, the isolates of serotype O157:H39 (06-4756 and C93_11) were neighbors of each other in the minimum spanning tree, but differed in 1,704 of the 2,415 genes tested (Fig. 2). For the remaining serotype (O157:H34), no additional *eae*-positive SF O157 isolate was available.

Interestingly, there is no other obvious link to any other serotype of the same pathotype, indicating independent polyphyletic evolutionary origins of the different O157 strains. To further test the hypothesis of independent evolution of the five different O157 serotypes, we compared the *rfb*O157 regions of the different serotypes with the *rfb*O157 region of O157:H7 reference strain Sakai. Of the 208 genes of the

*rfb*O157 region of O157:H7 strain Sakai, 100 genes (103,182 bp) were present in all five strains (see Table S2 in the supplemental material). After concatenation of the gene sequences and multiple alignment, we determined the single nucleotide polymorphisms (SNPs) and recombinatory sites. With the exception of serotype O157:H$^-$$_{fliCH7}$, which exhibited only 35 SNPs in comparison to those in the *rfb*O157 region of O157:H7 strain Sakai, all other strains showed 1,945 to 3,061 SNPs (Fig. 3). Moreover, the *rfb*O157 region is characterized by several areas of recombination (blue and red blocks in the outer circle of Fig. 3), further corroborating the independent evolution of the O157 determinant. Finally, we performed a BLAST search (http://blast.ncbi.nlm.nih.gov) with all contig sequences that harbored at least one gene of the *rfb*O157 region to determine whether these genes were all located on the chromosome or on plasmids, which might foster the exchange of genomic information. Here, all contigs contained chromosomal sequences surrounding the genes of the *rfb*O157 region.
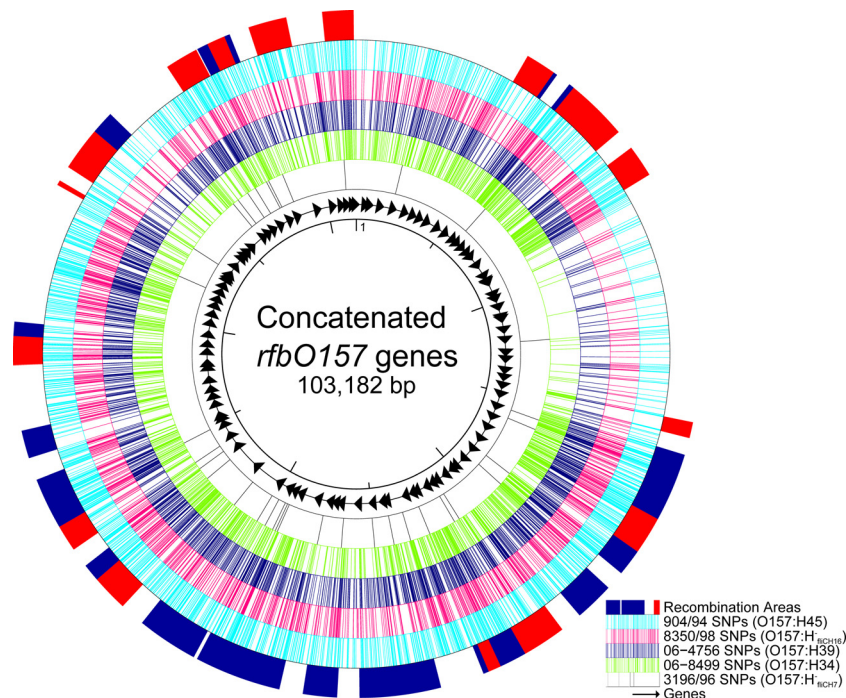
FIG 3 Prediction of recombinations within the *rfb*O157 gene cluster. The diagram illustrates the variation and the predicted areas of recombination within the *rfb*O157 gene cluster (100 genes that were present in all strains investigated were included in a concatenated manner with, in total, 103,182-bp length; the list of loci is given in Table S2 in the supplemental material). From inner to outer circle, the nucleotide position, the genes (arrows), the single nucleotide polymorphisms of the five *stx*-negative, *eae*-positive *E. coli* O157 in comparison to *E. coli* O157:H7 strain Sakai (GenBank accession no. NC_002695) and the predicted areas of recombination using Gubbins (34) are shown. Blue blocks of predicted recombinations are unique to a single isolate, while red blocks are shared by multiple isolates.

DISCUSSION

SF *E. coli* O157 strains without *stx* genes that carry *eae* alleles are categorized as atypical EPEC strains and were shown by WGS in this study to be genetically diverse with the exception of SF O157 strains possessing *fliC*H7 and *sfpA* encoding Sfp fimbriae, which belong to the O157:H7 CC11. All of the other *eae*-positive *E. coli* O157 strains with *fliC*H16, -H34, -H39, or -H45 are distinct from the O157:H7 complex; these clones are also distinct from each other. This finding is in agreement with the genetic diversity reported for other atypical EPEC strains (14, 35). In our previous study, we showed that atypical EPEC strains consist, in general, of two major groups (12, 14). The first group includes strains of a limited spectrum of serotypes that are also frequent among EHEC strains, share non-Stx virulence determinants with such EHEC strains, and cluster into the same MLST CC, demonstrating their close relatedness to EHEC strains. These atypical EPEC serotypes include O26:H11/H⁻, O103:H2/H⁻, O111:H8/H⁻, O121:H19, O145:H28/H⁻, and O157:H7/H⁻ and are presumably EHEC-LST, i.e., EHEC strains that lost their *stx* genes either during infection or during laboratory passages (12, 14). We showed in previous studies that both acquisition and loss of *stx*-encoding phages is possible and actually happens. Friedrich et al. showed that the loss of *stx*-encoding phages in O157 is not a rare phenomenon but happens in more than 10% of the cases (13). Later, we demonstrated that even acquisition of *stx*-encoding phages is possible (36). To do so, we used SF O157 isolates that lost *stx* from patients and showed that these isolates were able to take up *stx*-encoding phages. The typical insertion site of such phages in SF O157, *yecE*,

was also occupied in *stx*-positive strains and empty in *stx*-negative strains. We therefore tested all strains of this study, and indeed they all had an empty *stx* phage insertion site *yecE*, further confirming our hypothesis. A recent study, in which *stx*-negative atypical EPEC isolates of serotype O157:H7/H⁻ were, based on whole-genome data, closely related to *stx*-positive isolates of the same serotype also corroborates our findings (37). The second group of atypical EPEC comprises strains that belong to a broad spectrum of other serotypes or are nontypeable with the presently available antisera and usually do not possess the full spectrum of putative virulence characteristics present in strains of the first group. From our data, SF O157 strains with *fliC*H7 belong to the first group and those with *fliC*H16, -H34, -H39, and -H45 belong to the second group. Our data suggest that at least the O157 antigen evolved via recombination (Fig. 3); however, from our chromosomal sequencing data we could neither determine any directionality of evolution nor determine whether any further subgrouping is present among these serotypes. The bipartition into SF O157 strains with *fliC*H7 and all other O157 strains is also supported by the fact that 15 of 23 SF O157 strains with *fliC*H7 originated from patients with HUS. Among the strains with the other *fliC* types, only one with *fliC*H16 was isolated from a patient with HUS. In another study, an O157:H16 strain isolated from cattle was reported to possess both *stx*₁ and *stx*₂ genes (38). In our laboratory, no Stx-producing *E. coli* O157:H16 strains were isolated during the past 30 years. Therefore, it is not clear whether this HUS-associated *E. coli* O157:H16 strain lost *stx* prior to laboratory diagnostics or whether it was a coinfection with an *stx*-

harboring strain that was the cause of the HUS but was not isolated from the patient's stool sample. From a clinical perspective, our data show the need to determine in *eae*-positive *E. coli* O157 isolates the H antigen to elucidate whether they are putative EHEC-LST strains that had, at least prior to the loss of *stx*, the potential to cause HUS. Whether O157 isolates with *fliC* types other than H7 and H16 can be rated as only low pathogenic is difficult to say from the data presented here since the numbers of isolates of these serotypes were quite low; it is clear, however, that they were very unlikely to have originated from formerly *stx*-positive O157 strains. Several other studies corroborate our findings since the vast majority of O157 isolates with *fliC*H34, -H39, and -H45 were isolated either from animals or food only or from humans without HUS (3, 39–46). Recently, Sanjar et al. performed a similar analysis, focusing on the phylogenomic classification of O157:non-H7 isolates (31). Interestingly, they also determined that the bipartition of O157 into one group likely originating from EHEC O157:H7/H⁻ strains and another very diverse group consisting of *stx*-negative O157:non-H7 isolates. Our findings that the *rfb*O157 region is characterized by several areas of recombination (Fig. 3), which was also shown for EHEC O157:H7/H⁻ isolates (32) and for *E. coli* isolates in general (47), are in line with their findings of a nonmonophyletic origin of these isolates (31).

Our study has some limitations that warrant further comments. First, the number of samples per serotype is low. Nevertheless, our data show at least a tendency that is in line with the literature and should motivate further studies. Second, we restricted our analysis to serogroup O157. This serogroup, however, contains serotype O157:H7, which is the most prevalent serotype associated with severe human disease worldwide. Indeed, a similar situation is likely in other serogroups such as O145, in which different H antigens were present (48).

In summary, our results demonstrated the high diversity among the clinical isolates of the *E. coli* O157 serogroup and the necessity to determine the H antigen for a clinical risk assessment. Since phenotypic determination of the H antigen can be in vain due to nonmotility, the unavailability of antisera, or unspecific cross-reactions, molecular methods, e.g., *fliC* RFLP (18) or sequence-based approaches (49), should be applied.

ACKNOWLEDGMENTS

FUNDING INFORMATION

REFERENCES

1. **Levine MM, Xu JG, Kaper JB, Lior H, Prado V, Tall B, Nataro J, Karch H, Wachsmuth K.** 1987. A DNA probe to identify enterohemorrhagic *Escherichia coli* of O157:H7 and other serotypes that cause hemorrhagic colitis and hemolytic uremic syndrome. J Infect Dis **156:**175–182. http://dx.doi.org/10.1093/infdis/156.1.175.

2. **Trabulsi LR, Keller R, Tardelli Gomes TA.** 2002. Typical and atypical enteropathogenic *Escherichia coli*. Emerg Infect Dis **8:**508–513. http://dx.doi.org/10.3201/eid0805.010385.

3. **Feng PC, Keys C, Lacher D, Monday SR, Shelton D, Rozand C, Rivas M, Whittam T.** 2010. Prevalence, characterization and clonal analysis of *Escherichia coli* O157:non-H7 serotypes that carry *eae* alleles. FEMS Microbiol Lett **308:**62–67. http://dx.doi.org/10.1111/j.1574-6968.2010.01990.x.

4. **Makino S, Asakura H, Shirahata T, Ikeda T, Takeshi K, Arai K, Nagasawa M, Abe T, Sadamoto T.** 1999. Molecular epidemiological study of a mass outbreak caused by enteropathogenic *Escherichia coli* O157:H45. Microbiol Immunol **43:**381–384. http://dx.doi.org/10.1111/j.1348-0421.1999.tb02419.x.

5. **Schmidt H, Rüssmann H, Karch H.** 1993. Virulence determinants in nontoxinogenic *Escherichia coli* O157 strains that cause infantile diarrhea. Infect Immun **61:**4894–4898.

6. **Shelton DR, Karns JS, Higgins JA, Van Kessel JA, Perdue ML, Belt KT, Russell-Anelli J, Debroy C.** 2006. Impact of microbial diversity on rapid detection of enterohemorrhagic *Escherichia coli* in surface waters. FEMS Microbiol Lett **261:**95–101. http://dx.doi.org/10.1111/j.1574-6968.2006.00334.x.

7. **Wilson RA, Francis DH.** 1986. Fimbriae and enterotoxins associated with *Escherichia coli* serogroups isolated from pigs with colibacillosis. Am J Vet Res **47:**213–217.

8. **Donnenberg MS, Tzipori S, McKee ML, O'Brien AD, Alroy J, Kaper JB.** 1993. The role of the *eae* gene of enterohemorrhagic *Escherichia coli* in intimate attachment in vitro and in a porcine model. J Clin Invest **92:**1418–1424. http://dx.doi.org/10.1172/JCI116718.

9. **Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB.** 2013. Recent advances in understanding enteric pathogenic *Escherichia coli*. Clin Microbiol Rev **26:**822–880. http://dx.doi.org/10.1128/CMR.00022-13.

10. **Karch H, Tarr PI, Bielaszewska M.** 2005. Enterohaemorrhagic *Escherichia coli* in human medicine. Int J Med Microbiol **295:**405–418. http://dx.doi.org/10.1016/j.ijmm.2005.06.009.

11. **Nataro JP, Kaper JB.** 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev **11:**142–201.

12. **Bielaszewska M, Köck R, Friedrich AW, von Eiff C, Zimmerhackl LB, Karch H, Mellmann A.** 2007. Shiga toxin-mediated hemolytic uremic syndrome: time to change the diagnostic paradigm? PLoS One **2:**e1024. http://dx.doi.org/10.1371/journal.pone.0001024.

13. **Friedrich AW, Zhang W, Bielaszewska M, Mellmann A, Köck R, Fruth A, Tschäpe H, Karch H.** 2007. Prevalence, virulence profiles, and clinical significance of Shiga toxin-negative variants of enterohemorrhagic *Escherichia coli* O157 infection in humans. Clin Infect Dis **45:**39–45. http://dx.doi.org/10.1086/518573.

14. **Bielaszewska M, Middendorf B, Kock R, Friedrich AW, Fruth A, Karch H, Schmidt MA, Mellmann A.** 2008. Shiga toxin-negative attaching and effacing *Escherichia coli*: distinct clinical associations with bacterial phylogeny and virulence traits and inferred in-host pathogen evolution. Clin Infect Dis **47:**208–217. http://dx.doi.org/10.1086/589245.

15. **Schmidt H, Scheef J, Huppertz HI, Frosch M, Karch H.** 1999. *Escherichia coli* O157:H7 and O157:H⁻ strains that do not produce Shiga toxin: phenotypic and genetic characterization of isolates associated with diarrhea and hemolytic-uremic syndrome. J Clin Microbiol **37:**3491–3496.

16. **Nagano I, Kunishima M, Itoh Y, Wu Z, Takahashi Y.** 1998. Detection of verotoxin-producing *Escherichia coli* O157:H7 by multiplex polymerase chain reaction. Microbiol Immunol **42:**371–376. http://dx.doi.org/10.1111/j.1348-0421.1998.tb02297.x.

17. **Prager R, Strutz U, Fruth A, Tschäpe H.** 2003. Subtyping of pathogenic *Escherichia coli* strains using flagellar (H)-antigens: serotyping versus *fliC* polymorphisms. Int J Med Microbiol **292:**477–486. http://dx.doi.org/10.1078/1438-4221-00226.

18. **Zhang W, Mellmann A, Sonntag A-K, Wieler L, Bielaszewska M, Tschäpe H, Karch H, Friedrich AW.** 2007. Structural and functional differences between disease-associated genes of enterohaemorrhagic *Escherichia coli* O111. Int J Med Microbiol **297:**17–26. http://dx.doi.org/10.1016/j.ijmm.2006.10.004.

19. **Brunder W, Khan AS, Hacker J, Karch H.** 2001. Novel type of fimbriae encoded by the large plasmid of sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H⁻. Infect Immun **69:**4447–4457. http://dx.doi.org/10.1128/IAI.69.7.4447-4457.2001.

20. **Friedrich AW, Nierhoff KV, Bielaszewska M, Mellmann A, Karch H.** 2004. Phylogeny, clinical associations, and diagnostic utility of the pilin subunit gene (*sfpA*) of sorbitol-fermenting, enterohemorrhagic *Esche-*

richia coli O157:H⁻. J Clin Microbiol **42**:4697–4701. http://dx.doi.org/10.1128/JCM.42.10.4697-4701.2004.

21. Janka A, Bielaszewska M, Dobrindt U, Greune L, Schmidt MA, Karch H. 2003. Cytolethal distending toxin gene cluster in enterohemorrhagic *Escherichia coli* O157:H⁻ and O157:H7: characterization and evolutionary considerations. Infect Immun **71**:3634–3638. http://dx.doi.org/10.1128/IAI.71.6.3634-3638.2003.

22. Schmidt H, Beutin L, Karch H. 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. Infect Immun **63**:1055–1061.

23. Zhang WL, Kohler B, Oswald E, Beutin L, Karch H, Morabito S, Caprioli A, Suerbaum S, Schmidt H. 2002. Genetic diversity of intimin genes of attaching and effacing *Escherichia coli* strains. J Clin Microbiol **40**:4486–4492. http://dx.doi.org/10.1128/JCM.40.12.4486-4492.2002.

24. Bielaszewska M, Prager R, Zhang W, Friedrich AW, Mellmann A, Tschape H, Karch H. 2006. Chromosomal dynamism in progeny of outbreak-related sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:NM. Appl Environ Microbiol **72**:1900–1909. http://dx.doi.org/10.1128/AEM.72.3.1900-1909.2006.

25. De Greve H, Qizhi C, Deboeck F, Hernalsteens JP. 2002. The Shiga-toxin VT2-encoding bacteriophage varphi297 integrates at a distinct position in the *Escherichia coli* genome. Biochim Biophys Acta **1579**:196–202. http://dx.doi.org/10.1016/S0167-4781(02)00539-0.

26. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol **60**:1136–1151. http://dx.doi.org/10.1111/j.1365-2958.2006.05172.x.

27. Mellmann A, Bielaszewska M, Köck R, Friedrich AW, Fruth A, Middendorf B, Harmsen D, Schmidt MA, Karch H. 2008. Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic *Escherichia coli*. Emerg Infect Dis **14**:1287–1290. http://dx.doi.org/10.3201/eid1408.071082.

28. Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, McLaughlin SF, Henkhaus JK, Leopold B, Bielaszewska M, Prager R, Brzoska PM, Moore RL, Guenther S, Rothberg JM, Karch H. 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One **6**:e22751. http://dx.doi.org/10.1371/journal.pone.0022751.

29. Hazen TH, Sahl JW, Fraser CM, Donnenberg MS, Scheutz F, Rasko DA. 2013. Refining the pathovar paradigm via phylogenomics of the attaching and effacing *Escherichia coli*. Proc Natl Acad Sci U S A **110**:12810–12815. http://dx.doi.org/10.1073/pnas.1306836110.

30. Rump LV, Strain EA, Cao G, Allard MW, Fischer M, Brown EW, Gonzalez-Escalona N. 2011. Draft genome sequences of six *Escherichia coli* isolates from the stepwise model of emergence of *Escherichia coli* O157:H7. J Bacteriol **193**:2058–2059. http://dx.doi.org/10.1128/JB.00118-11.

31. Sanjar F, Rusconi B, Hazen TH, Koenig SS, Mammel MK, Feng PC, Rasko DA, Eppinger M. 2015. Characterization of the pathogenome and phylogenomic classification of enteropathogenic *Escherichia coli* of the O157:non-H7 serotypes. Pathog Dis **73**:ftv033. http://dx.doi.org/10.1093/femspd/ftv033.

32. Leopold SR, Magrini V, Holt NJ, Shaikh N, Mardis ER, Cagno J, Ogura Y, Iguchi A, Hayashi T, Mellmann A, Karch H, Besser TE, Sawyer SA, Whittam TS, Tarr PI. 2009. A precise reconstruction of the emergence and constrained radiations of *Escherichia coli* O157 portrayed by backbone concatenomic analysis. Proc Natl Acad Sci U S A **106**:8713–8718. http://dx.doi.org/10.1073/pnas.0812949106.

33. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30**:772–780. http://dx.doi.org/10.1093/molbev/mst010.

34. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res **43**:e15. http://dx.doi.org/10.1093/nar/gku1196.

35. Bando SY, Andrade FB, Guth BE, Elias WP, Moreira-Filho CA, Pestana de

Castro AF. 2009. Atypical enteropathogenic *Escherichia coli* genomic background allows the acquisition of non-EPEC virulence factors. FEMS Microbiol Lett **299**:22–30. http://dx.doi.org/10.1111/j.1574-6968.2009.01735.x.

36. Mellmann A, Lu S, Karch H, Xu JG, Harmsen D, Schmidt MA, Bielaszewska M. 2008. Recycling of shiga toxin 2 genes in sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:NM. Appl Environ Microbiol **74**:67–72. http://dx.doi.org/10.1128/AEM.01906-07.

37. Ferdous M, Zhou K, Mellmann A, Morabito S, Croughs PD, de Boer RF, Kooistra-Smid AM, Rossen JW, Friedrich AW. 2015. Is Shiga toxin-negative *Escherichia coli* O157:H7 enteropathogenic or enterohemorrhagic *Escherichia coli*? Comprehensive molecular analysis using whole-genome sequencing. J Clin Microbiol **53**:3530–3538. http://dx.doi.org/10.1128/JCM.01899-15.

38. Ennis C, McDowell D, Bolton DJ. 2012. The prevalence, distribution and characterization of Shiga toxin-producing *Escherichia coli* (STEC) serotypes and virulotypes from a cluster of bovine farms. J Appl Microbiol **113**:1238–1248. http://dx.doi.org/10.1111/j.1365-2672.2012.05421.x.

39. Afset JE, Anderssen E, Bruant G, Harel J, Wieler L, Bergh K. 2008. Phylogenetic backgrounds and virulence profiles of atypical enteropathogenic *Escherichia coli* strains from a case-control study using multilocus sequence typing and DNA microarray analysis. J Clin Microbiol **46**:2280–2290. http://dx.doi.org/10.1128/JCM.01752-07.

40. Bentancor A, Vilte DA, Rumi MV, Carbonari CC, Chinen I, Larzabal M, Cataldi A, Mercado EC. 2010. Characterization of non-Shiga toxin-producing *Escherichia coli* O157 strains isolated from dogs. Rev Argent Microbiol **42**:46–48.

41. Blanco M, Schumacher S, Tasara T, Zweifel C, Blanco JE, Dahbi G, Blanco J, Stephan R. 2005. Serotypes, intimin variants and other virulence factors of eae positive *Escherichia coli* strains isolated from healthy cattle in Switzerland. Identification of a new intimin variant gene (*eae*-η2). BMC Microbiol **5**:23.

42. Feng PC, Keys C, Lacher DW, Beutin L, Bentancor A, Heuvelink A, Afset JE, Rumi V, Monday S. 2012. Clonal relations of atypical enteropathogenic *Escherichia coli* O157:H16 strains isolated from various sources from several countries. FEMS Microbiol Lett **337**:126–131. http://dx.doi.org/10.1111/1574-6968.12017.

43. Hazen TH, Sahl JW, Fraser CM, Donnenberg MS, Scheutz F, Rasko DA. 2013. Draft genome sequences of three O157 enteropathogenic *Escherichia coli* isolates. Genome Announc **1**:e00516–00513. http://dx.doi.org/10.1128/genomeA.00516-13.

44. Mora A, Blanco M, Yamamoto D, Dahbi G, Blanco JE, Lopez C, Alonso MP, Vieira MA, Hernandes RT, Abe CM, Piazza RM, Lacher DW, Elias WP, Gomes TA, Blanco J. 2009. HeLa-cell adherence patterns and actin aggregation of enteropathogenic *Escherichia coli* (EPEC) and Shiga-toxin-producing *E. coli* (STEC) strains carrying different *eae* and *tir* alleles. Int J Microbiol **12**:243–251.

45. Stephan R, Borel N, Zweifel C, Blanco M, Blanco JE. 2004. First isolation and further characterization of enteropathogenic *Escherichia coli* (EPEC) O157:H45 strains from cattle. BMC Microbiol **4**:10. http://dx.doi.org/10.1186/1471-2180-4-10.

46. Willshaw GA, Scotland SM, Smith HR, Cheasty T, Thomas A, Rowe B. 1994. Hybridization of strains of *Escherichia coli* O157 with probes derived from the *eaeA* gene of enteropathogenic *E. coli* and the *eaeA* homolog from a Vero cytotoxin-producing strain of *E. coli* O157. J Clin Microbiol **32**:897–902.

47. Milkman R, Jaeger E, McBride RD. 2003. Molecular evolution of the *Escherichia coli* chromosome. VI. Two regions of high effective recombination. Genetics **163**:475–483.

48. Sonntag A-K, Prager R, Bielaszewska M, Zhang W, Fruth A, Tschäpe H, Karch H. 2004. Phenotypic and genotypic analyses of enterohemorrhagic *Escherichia coli* O145 strains from patients in Germany. J Clin Microbiol **42**:954–962. http://dx.doi.org/10.1128/JCM.42.3.954-962.2004.

49. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol **53**:2410–2426. http://dx.doi.org/10.1128/JCM.00008-15.