

University of Groningen

## Shiga toxin-producing Escherichia coli (STEC) from Humans in the Netherlands

Ferdous, Mithila

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Ferdous, M. (2017). *Shiga toxin-producing Escherichia coli (STEC) from Humans in the Netherlands: Novel diagnostic approach, molecular characterization and phylogenetic background*. University of Groningen.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# CHAPTER 4

## Comprehensive Characterization of *Escherichia coli* O104:H4 Isolated from Patients in the Netherlands

Mithila Ferdous<sup>1\*</sup>, Kai Zhou<sup>1\*</sup>, Richard F. de Boer<sup>2</sup>, Alexander W. Friedrich<sup>1</sup>, Anna M.D. Kooistra-Smid<sup>1,2</sup>, John W.A. Rossen<sup>1</sup>

<sup>1</sup> Department of Medical Microbiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

<sup>2</sup> Certe - Laboratory for Infectious Diseases, Groningen, the Netherlands

\*These authors equally contributed to this work

### Key words

Shiga toxin-producing *E. coli* –STEC, Enterohemorrhagic *E. coli* –EHEC, Hemolytic uremic syndrome, Outbreak, Antimicrobial resistance, Whole genome sequencing, Phylogenetic analysis

Front. Microbiol (2015) 6:1348.

### ABSTRACT

In 2011, a Shiga toxin-producing Enteroaggregative *Escherichia coli* (EAEC Stx2a+) O104:H4 strain caused a serious outbreak of acute gastroenteritis and hemolytic-uremic syndrome (HUS) in Germany. In 2013, *E. coli* O104:H4 isolates were obtained from a patient with HUS and her friend showing only gastrointestinal complaints. The antimicrobial resistance and virulence profiles of these isolates together with three EAEC Stx2a+ O104:H4 isolates from 2011 were determined and compared. Whole-genome sequencing (WGS) was performed for detailed characterization and to determine genetic relationship of the isolates. Four additional genomes of EAEC Stx2a+ O104:H4 isolates of 2009 and 2011 available on NCBI were included in the virulence and phylogenetic analysis. All *E. coli* O104:H4 isolates tested were positive for *stx2a*, *aatA* and *terD* but were negative for *escV*. All, except one 2011 isolate, were positive for *aggR* and were therefore considered EAEC. The EAEC Stx2a+ O104:H4 isolates of 2013 belonged to sequence type (ST) ST678 as the 2011 isolates and showed slightly different resistance and virulence patterns compared to the 2011 isolates. Core-genome phylogenetic analysis showed that the isolates of 2013 formed a separate cluster from the isolates of 2011 and 2009 by 27 and 20 different alleles, respectively. In addition, only a one-allele difference was found between the isolate of the HUS-patient and that of her friend. Our study shows that EAEC Stx2a+ O104:H4 strains highly similar to the 2011 outbreak clone in their core genome are still circulating necessitating proper surveillance to prevent further outbreaks with these potentially pathogenic strains. In addition, WGS not only provided a detailed characterization of the isolates but its high discriminatory power also enabled us to discriminate the 2013 isolates from the isolates of 2009 and 2011 expediting the use of WGS in public health services to rapidly apply proper infection control strategies.

### INTRODUCTION

Shiga toxin-producing *Escherichia coli* (STEC) are a pathogenic group of *Escherichia coli* (*E. coli*) producing the potent cytotoxin Shiga-toxin (Stx) similar to the one produced by *Shigella dysenteriae* serotype 1 (1). STEC may cause a broad spectrum of illness, ranging from diarrhea to the potentially fatal hemolytic uremic syndrome (HUS) (2). A subset of STEC can cause bloody diarrhea in humans and they are known as enterohemorrhagic *E. coli* (EHEC) while a subset of EHEC can cause HUS and are known as hemolytic uremic syndrome-associated *E. coli* HUSEC (3). In addition to Shiga toxin-converting bacteriophages, STEC may contain several other mobile genetic elements encoding virulence factors as pathogenicity islands (PAI), and a large, approximately 90 kb plasmid (pO157) (4). Stx production is common to all HUS-associated *E. coli* isolates regardless of their serotype. When the toxin enters the blood stream it binds to receptors on endothelial cells abundantly present in kidneys and brain, leading to neurological sequel and/or to microvascular disease that may result in HUS (5). In 2011, a large outbreak was reported in Germany caused by an Enteroaggregative *E. coli* (EAEC) O104:H4 strain lysogenized with the Stx2a bacteriophage and thereby becoming an EAEC/STEC hybrid strain (6, 7). Besides the *stx2a* gene, this unusual strain had virulence properties of EAEC including plasmid pAA carrying the aggregative adherence fimbriae (AAF) variant I encoded by the *aggA* gene whose expression is regulated by the *aggR* gene. In addition, it contained a protein-coat secretion system (Aat), dispersin (Aap), a putative type VI secretion system (Aai), and a rare combination of serine protease autotransporters of Enterobacteriaceae (SPATEs) genes, i.e., *sepA*, *sigA* and *pic* (8, 9). It also contained a *terD* gene (tellurite resistance gene as a marker for the *ter* cluster) and a plasmid-borne extended spectrum beta-lactamase (ESBL) gene *bla*<sub>CTX-M-15</sub> resulting in resistance to several antibiotics (7, 8). However, the outbreak clone did not possess the *escV* gene encoding the predicted outer membrane protein and a marker for the locus of enterocyte effacement (LEE) PAI (9, 10).

In clinical microbiology, Whole Genome Sequencing (WGS) has already shown its value in outbreak investigations and epidemiological typing due to its high-resolution discriminatory power and detailed virulence profiling, thereby becoming more and more important in routine diagnostics (11-13). In this study, we characterized EAEC Stx2a+ O104:H4 strains isolated from a HUS patient and her friend who travelled together to Turkey in 2013 prior to diagnosing the patient with HUS. For this, a WGS approach in parallel with routine phenotypic and genotypic laboratory methods was used. Analyses were performed to get more insight into the antibiotic resistance and virulence profiles of the isolates and to reveal their genetic relationship with the 2011 German outbreak EAEC Stx2a+ O104:H4 isolates.

### MATERIALS AND METHODS

#### *E. coli* isolates used in this study

In July 2013, four *E. coli* isolates were obtained from a HUS patient (isolate 338) and her friend (isolates 381-1, 381-3 and 381-4). They were compared to three EAEC Stx2a+ O104:H4 strains named 7N, 8G and LB227103 which were a kind gift of Dr. Alexander Mellmann (Institute of Hygiene, University of Muenster, Muenster, Germany) and were isolated during the 2011 German outbreak period from stool samples of patients, submitted to the National Consulting Laboratory for Hemolytic Uremic Syndrome in Münster, Germany, between May 23 and June 2, 2011 (6, 14). In addition, publically available genomes of five previously reported strains (TY-2482, 2011C-3494, 2009EL-2050, 2009EL-2071 and 55989) were included in virulence and phylogenetic analyses. Detailed information on the isolates used in this study is shown in Table 1.

#### Diagnostic procedures

Fecal samples from the HUS patient (patient 338) and her friend (patient 381) were collected for diagnostic purposes at Certe Laboratory for Infectious Diseases as described previously (18). Shortly, fecal samples were screened for the presence of the virulence genes *stx1*, *stx2* and *escV* by real-time PCR (qPCR) and *stx*-positive samples were subsequently enriched in Brilliant green bile (BGB) broth. After DNA isolation from the enriched broth, qPCR was performed for detection of the EAEC pAA plasmid encoding genes *aggR* and *aatA* and, O-serogroup encoding genes (O26, O91, O103, O104, O111, O121, O145 and O157). In parallel, part of the BGB broth was plated on CHROMagar STEC and Sorbitol MacConkey agar plates (Mediaproducs BV, Groningen, the Netherlands). To obtain a pure isolate from the selective agar plates, more than 30 colonies of each patient were analysed for the presence of *stx* genes by qPCR and for ESBL production by using CHROM ID ESBL agar (bioMérieux, Marcy l'Etoile, France). In addition, sorbitol fermentation and tellurite resistance of the isolates were determined by observing their ability to grow on CT-SMAC plates (Sorbitol MacConkey agar with Cefixime and Tellurite). In addition, conventional PCRs for *stx*-subtypes 2a, 2b, 2c, 2d, 2e, 2f and 2g were performed to determine the subtype of the *stx2* gene as described previously (19). Finally, a multiplex PCR targeting typical molecular features of the 2011 outbreak strain, i.e. the presence of the *wzy*<sub>O104</sub> (O antigen polymerase as a marker for the *E. coli* O104 gene cluster), *fliCH4* (encoding the H4 specific flagellin), *stx2*, and *terD* genes was performed (6, 20).

### Antimicrobial Resistance

Antibiotic resistance patterns of the four isolates of 2013 (338, 381-1, 381-3 and 381-4) and three isolates of 2011 (7N, 8G and LB227103) were analyzed using VITEK2 (bioMérieux, Marcy l'Etoile, France) and E-Test (bioMérieux, Marcy l'Etoile, France) following EUCAST guidelines. Presence of ESBL, carbapenemases and AmpCs genes was confirmed using the Check-MDR CT103 assay according to the manufacturer's protocol (Check-Points BV Wageningen, the Netherlands).

### DNA isolation

For repetitive sequence based PCR fingerprinting and whole genome sequencing, genomic DNA was extracted by the UltraClean® microbial DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, US) according to the manufacturer's protocol. For other molecular assays DNA was extracted using a DNeasy blood and tissue kit (QIAGEN, GmbH, Hilden, Germany) following the manufacturer's protocol.

### Multilocus sequence typing (MLST)

MLST was performed using a 3130xl Genetic Analyzer (Applied Biosystems) to sequence internal fragments of seven housekeeping genes (*adhA*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) according to the protocol described in the *E. coli* MLST databases ([http://mlst.warwick.ac.uk/mlst/dbs/Ecoli/documents/primersColi\\_html](http://mlst.warwick.ac.uk/mlst/dbs/Ecoli/documents/primersColi_html)). The alleles and sequence types (ST) were assigned in accordance with the same database.

### Virulence properties

To assess the presence of known *E. coli* virulence genes, a DNA microarray analysis was performed using the *E. coli* genotyping combined assay kit according to the manufacturer's protocol (Clondiag, Alere Technologies, GmbH, Jena, Germany) (21).

### Repetitive sequence based PCR (rep-PCR)

Isolated DNA was amplified using the DiversiLab *Escherichia* kit for repetitive sequence based PCR fingerprinting following the manufacturer's instructions (bioMérieux, Marcy-l'Etoile, France). The rep-PCR products were detected and analysed using lab-on-a-chip microfluidics technology on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Further analysis was performed with the web-based DiversiLab software (version 3.4) using the Pearson correlation coefficient to calculate pairwise

## Chapter 4

---

similarities among all samples. In this study, we used 95% similarity thresholds for the data analysis (22).

### Whole genome sequencing and analysis

A DNA library was prepared using the Nextera XT kit (Illumina, San Diego, CA, US) according to the manufacturer's instructions and then run on a Miseq (Illumina) for generating paired-end 250-bp reads. *De novo* assembly was performed using default settings of CLC Genomics Workbench v6.0.5 (CLC bio A/S, Denmark) as described previously (23). The MLST sequence type was also confirmed from the whole genome sequence by uploading the assembled genomes to the CGE MLST server (version 1.7) (24); the configuration for MLST was set to *E. coli* scheme 1. The resistome was generated by uploading assembled genomes to the CGE Resfinder server 2.0 (25). The default settings for Resfinder were used except for the threshold of ID that was set to 85%. The serogroup of the isolates was identified using CGE SerotypeFinder 1.0 (26). The virulence genes were identified by mapping reads to a pseudomolecule generated by concatenating a set of *E. coli* virulence genes described before (6, 27). In addition to the isolates of 2013 and 2011, the genomes of two isolates of 2009 (2009EL-2050, 2009EL-2071) were included for comparison of the virulence genes. The genome sequence of 2011C-3493 (GenBank accession no. NC\_018658.1) was used as reference for extracting open reading frames (ORFs) from contigs of each strain by SeqSphere v1.0 (Ridom GmbH, Münster, Germany). Only the ORFs without premature stop codons and ambiguous nucleotides were extracted, and the sequences of those ORFs shared by all samples analyzed here were defined as the core genome for subsequent phylogenetic analysis (28). According to the sequence identity of each ORF, a numerical allele type was assigned by SeqSphere. SeqSphere used the allelic profile formed by the combination of all alleles in each genome for constructing the minimum-spanning tree. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number JRJF01000000 (*E. coli* 338), JRKD01000000 (*E. coli* 381-1), JRLM01000000 (*E. coli* 381-3), JRLD01000000 (*E. coli* 381-4), JRKE01000000 (*E. coli* 7N), JRLN00000000 (*E. coli* 8G), JRKF01000000 (*E. coli* LB227103). The version described in this paper is the first version.

## RESULTS

### Demographic and clinical characteristics of patients

Strain 338 was isolated from the fecal sample of a 22 year-old female patient (338) admitted to the Intensive Care Unit of the University Medical Center Groningen (UMCG) in July 2013, after returning

## Comparing *Escherichia coli* O104:H4 isolates

from a holiday in Turkey. She presented with abdominal pain, bloody diarrhea, vomiting and eventually developed HUS. Patient 381, a 24-year-old female travelling together with patient 338, showed symptoms of diarrhea, abdominal cramps and abdominal pain but was not hospitalized.

**Table 1. Characteristics of isolates analyzed in this study.**

Isolate ID	Date of isolation	Country of Isolation	Clinical manifestations	<i>stx</i> subtype	Serotype	ESBL	MLST	Source
338	July 2013	Netherlands	HUS	2a	O104:H4	negative	ST-678	this study
381-1	July 2013	Netherlands	Diarrhea	2a	O104:H4	positive	ST-678	this study
381-3	July 2013	Netherlands	Diarrhea	NA <sup>a</sup>	O126:H2	positive	ST-10	this study
381-4	July 2013	Netherlands	Diarrhea	2a	O104:H4	negative	ST-678	this study
7N	2011	Germany	Unknown	2a	O104:H4	positive	ST-678	this study
8G	2011	Germany	Unknown	2a	O104:H4	positive	ST-678	this study
LB227103	2011	Germany	HUS	2a	O104:H4	positive	ST-678	(6, 14)
TY-2482	2011	Germany	HUS	2a	O104:H4	positive	ST-678	(12)
2011C-3493	2011	U.S.A.	HUS	2a	O104:H4	positive	ST-678	(15)
2009EL-2050	2009	Republic of Georgia	Bloody diarrhea	2a	O104:H4	negative	ST-678	(15)
2009EL-2071	2009	Republic of Georgia	Bloody diarrhea	2a	O104:H4	negative	ST-678	(15)
55989	1995	Central African Republic	Diarrhea	NA <sup>a</sup>	O104:H4	negative	ST-678	(16, 17)

<sup>a</sup> NA = Not applicable as these isolates are *stx* negative

### Primary screening results

Direct screening of feces by qPCR revealed the presence of *stx1*, *stx2* and *escV* genes (patient 338) and *stx2* gene (patient 381). Additional molecular screening of DNA isolated from enriched BGB broth for other virulence genes and O-serogroups resulted in detection of *aggR*, *aatA* and *wzy*<sub>O104</sub> in both patients. Subsequently, as a result of colony screening for the presence of the *stx* gene and/or for ESBL production one *stx2*-positive non ESBL-producing isolate (isolate 338) was obtained from patient 338, whereas two different *stx2*-positive isolates were recovered from patient 381: an ESBL producing isolate (381-1) and a non ESBL producing isolate (381-4). In addition, one *stx*-negative but ESBL-producing isolate (381-3) was obtained from this patient. All *stx*-positive isolates contained the *aatA* and *terD* but not the *escV* gene, similar to the 2011 German outbreak isolates. No *stx1*-positive



## Chapter 4

---

isolates were obtained. The *aggR* gene was detected in the 2013 and two of the three 2011 isolates (7N and 8G) but unexpectedly not in isolate LB227103. All *stx*-positive isolates were subtyped as *stx2a* and shared the same serotype (O104:H4) as the 2011 German outbreak isolates, whereas the *stx* negative ESBL producing isolate was serotype O126:H2. All isolates were sorbitol fermenting and tellurite resistant. MLST analysis showed that all 2013 EAEC Stx2a+ O104:H4 isolates belonged to ST678 as the 2011 outbreak strains, whereas the *stx*-negative isolate (381-3) was assigned to ST10. The screening results are summarized in Table 1.

### Phenotypic and genotypic resistance profiles

Isolates 338 and 381-4 (both ESBL negative) were resistant to ampicillin and trimethoprim. Isolate 381-1 (ESBL positive) was resistant to ampicillin, cefotaxime, ceftazidime, cefepime, cefuroxime, and trimethoprim and this resistance pattern was almost identical to that of the *stx*-negative/ESBL positive isolate 381-3 from the same patient except that the latter was susceptible to trimethoprim. The resistance pattern of isolate 381-1 differed from that of three of the 2011 outbreak isolates which were resistant to amoxicillin-clavulanic acid, tetracycline and trimethoprim/sulfamethoxazole (Table 2).

In agreement with the results of the phenotypic resistance pattern, the resistome of the three 2013 EAEC Stx2a+ O104:H4 isolates differed from that of the 2011 isolates: the *strA*, *strB* (streptomycin resistance), *sul2* (sulfonamide resistance) and *tetA* (tetracycline resistance) genes were only found in the 2011 outbreak isolates (Table 2). A *bla*<sub>CTX-M-15</sub> gene was detected in isolate 381-1 (*stx*-positive) and 381-3 (*stx*-negative) as well as in the 2011 isolates. In addition, WGS showed the presence of the *dfrA7* gene in all isolates except 381-3 resulting in the latter's trimethoprim-sensitive phenotype. A truncated *sul1* gene resulting from an insertion of *IS26* was found in all EAEC Stx2a+ O104:H4 isolates.

### Virulence Profile

DNA microarray results showed that the virulence profile of the 2013 isolates was similar to that of the EAEC Stx2a+ O104:H4 isolates of 2011 and 2009 used in this study. The only difference was the presence of the microcin H47 biosynthesis gene *mchB* and *mchC*, which were identified in the 2009 and 2011 strains, but were absent in the EAEC Stx2a+ O104:H4 isolates of 2013.

The results obtained by WGS were identical to that of the DNA microarray and confirmed the absence of the *aggR* gene in the outbreak strain LB227103. WGS revealed the presence of several additional virulence genes including *aap*, *aaiC*, *aggA*, *lpfA*<sub>O113</sub>, *lpfA*<sub>O26</sub>, *fyuA*, *irp2* and *set1* (not detected by PCR or microarray) in our isolates (Table 3).

## Comparing *Escherichia coli* O104:H4 isolates

**Table 2. Phenotypic antibiotic resistance patterns and presence of corresponding antibiotic resistance genes among the isolates.**

Name of Antibiotics	2011 isolates			2013 isolates			
	7N	8G	LB227103	338	381-1	381-4	381-3
Ampicillin	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TE-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>
Amoxicillin+ Clav acid	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	R/ <i>bla</i> <sub>TEM-1</sub>	S	I	S	S
Cefuroxime	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>
Cefotaxime	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>
Ceftazidime	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>
Cefepime	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>	S/N	R/ <i>bla</i> <sub>CTX-M-15</sub>
Co-trimoxazole	R/ <i>sul1</i> & 2	R/ <i>sul1</i> & 2	R/ <i>sul1</i> & 2	S/ <i>sul1</i> <sup>a</sup>	S/ <i>sul1</i> <sup>a</sup>	S/ <i>sul1</i> <sup>a</sup>	S/N
Tetracycline	R/ <i>tetA</i>	R/ <i>tetA</i>	R/ <i>tetA</i>	S/N	S/N	S/N	S/N
Trimethoprim	R/ <i>dfrA7</i>	R/ <i>dfrA7</i>	R/ <i>dfrA7</i>	R/ <i>dfrA7</i>	R/ <i>dfrA7</i>	R/ <i>dfrA7</i>	S/N
Streptomycin	ND/ <i>strA-B</i>	ND/ <i>strA-B</i>	ND/ <i>strA-B</i>	ND/N	ND/N	ND/N	ND/N

<sup>a</sup> *sul1* gene was deactivated by an *IS26* insertion

(R represents drug resistance, S represents drug sensitivity; I represents intermediate; N represents not found; ND represents not determined)

## Chapter 4

**Table 3. Virulence genes detected by qPCR, microarray and whole genome sequencing (WGS) analysis.**

Gene name	Function	7N	8G	LB22710	338	381-1	381-4	2009EL-2050	2009EL-2071	55989
<i>aap</i> <sup>a</sup>	synthesis of anti-aggregation protein dispersin	+	+	+	+	+	+	+	+	+
<i>aatA</i>	necessary for translocation of dispersin (Aap)	+	+	+	+	+	+	+	+	+
<i>aaiC</i> <sup>a</sup>	aggR-activated island C	+	+	+	+	+	+	+	+	+
<i>aggA</i> <sup>a</sup> (AAF/I)	Pilin subunit of aggregative adherence fimbriae I	+	+	+	+	+	+	+	+	-
<i>aafA</i> <sup>a</sup>	Pilin subunit of aggregative adherence fimbriae II	-	-	-	-	-	-	-	-	-
<i>agg3A</i> <sup>a</sup> (AAF/III)	Pilin subunit of aggregative adherence fimbriae III	-	-	-	-	-	-	-	-	+
<i>agg4A</i> <sup>a</sup>	Pilin subunit of aggregative adherence fimbriae IV	-	-	-	-	-	-	-	-	-
<i>Agg5A</i> <sup>a</sup>	Pilin subunit of aggregative adherence fimbriae V	-	-	-	-	-	-	-	-	-
<i>aggR</i>	Transcriptional regulator AggR	+	+	-	+	+	+	+	+	+
<i>astA</i> (EAST1)	heat-stable enterotoxin 1	-	-	-	-	-	-	-	-	+
<i>bfpA</i>	Bundle-forming pili	-	-	-	-	-	-	-	-	-
<i>cdt (I-V)</i>	Cytolethal distending toxin	-	-	-	-	-	-	-	-	-
<i>eae</i>	Intimin	-	-	-	-	-	-	-	-	-
<i>EHEC-hlyA</i>	EHEC haemolysin	-	-	-	-	-	-	-	-	-
<i>elt</i>	Heat-labile enterotoxin (LT)	-	-	-	-	-	-	-	-	-
<i>espP</i>	Serine protease	-	-	-	-	-	-	-	-	-
<i>est1a</i>	Heat-stable enterotoxin (STIa)	-	-	-	-	-	-	-	-	-
<i>est1b</i>	Heat-stable enterotoxin (STIb)	-	-	-	-	-	-	-	-	-
<i>fyuA</i> <sup>a</sup>	Component of iron uptake system on HPI	+	+	+	+	+	+	+	+	+
<i>ial</i> <sup>a</sup>	Invasive plasmid (plnv)	-	-	-	-	-	-	-	-	-
<i>iha</i>	Iron-regulated gene	+	+	+	+	+	+	+	+	+
<i>irp2</i> <sup>a</sup>	Component of iron uptake system on HPI	+	+	+	+	+	+	+	+	+
<i>lpfA</i> <sub>O113</sub> <sup>a</sup>	Structural subunit of LPF of STEC O113	+	+	+	+	+	+	+	+	+
<i>lpfA</i> <sub>O157-O1141</sub> <sup>a</sup>	Structural subunit of LPF of STEC O157:H7	-	-	-	-	-	-	-	-	-
<i>lpfA</i> <sub>O157-O1154</sub> <sup>a</sup>	Structural subunit of LPF of STEC O157:H7	-	-	-	-	-	-	-	-	-
<i>lpfA</i> <sub>O26</sub> <sup>a</sup>	Structural subunit of long polar fimbriae (LPF) of STEC O26	+	+	+	+	+	+	+	+	+
<i>mchB</i>	synthesis of Microcin	+	+	+	-	-	-	+	+	-
<i>mchC</i>	synthesis of Microcin	+	+	+	-	-	-	+	+	-
<i>mchF</i>	synthesis of Microcin	+	+	+	+	+	+	+	+	-
<i>pet</i>	plasmid-encoded toxin	-	-	-	-	-	-	-	-	-
<i>pic</i>	gene encoding protein for intestinal colonization	+	+	+	+	+	+	+	+	+
<i>saa</i>	STEC autoagglutinating adhesin	-	-	-	-	-	-	-	-	-
<i>sat</i>	Secreted autotransporter toxin	-	-	-	-	-	-	-	-	-
<i>set1</i> <sup>a</sup>	Shigella enterotoxin 1	+	+	+	+	+	+	+	+	+
<i>sepA</i>	Shigella extracellular protease	+	+	+	+	+	+	+	+	-

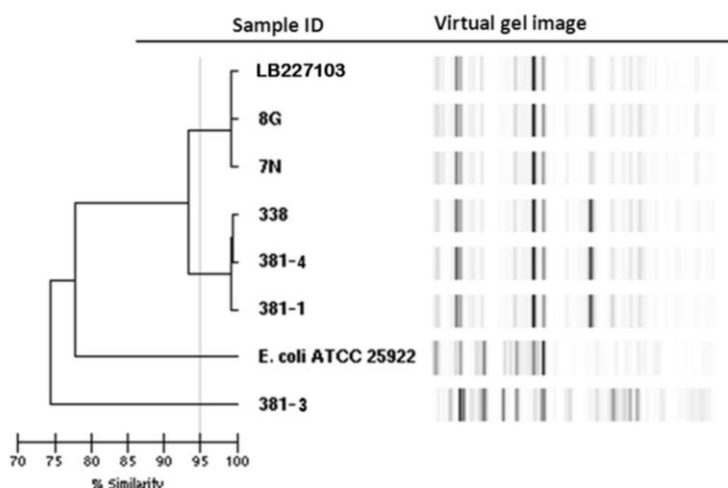
## Comparing *Escherichia coli* O104:H4 isolates

Gene name	Function	7N	8G	LB22710 3	338	381-1	381-4	2009EL- 2050	2009EL- 2071	55989
<i>sfpA</i>	plasmid-borne gene encoding the pilin subunit	-	-	-	-	-	-	-	-	-
<i>sigA</i>	Shigella IgA-like protease homologue	+	+	+	+	+	+	+	+	+
<i>stx1</i>	Shiga toxin 1	-	-	-	-	-	-	-	-	-
<i>stx2</i>	Shiga toxin 2	+	+	+	+	+	+	+	+	-
<i>subA</i>	Subtilase cytotoxin	-	-	-	-	-	-	-	-	-
<i>ter cluster</i>	Tellurite resistance	+	+	+	+	+	+	+	+	-

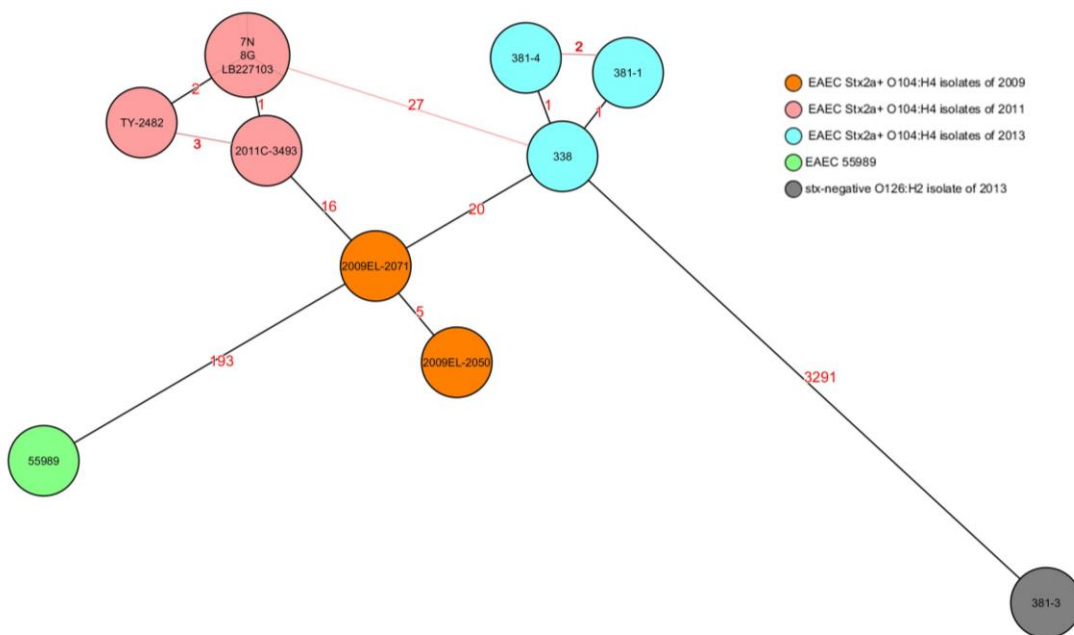
<sup>a</sup>Data derived only from WGS

### Phylogenetic analyses

Rep-PCR analysis was used to investigate the genetic relationship between the different isolates. EAEC Stx2a+ O104:H4 isolates of 2011 and 2013 clustered into two different groups sharing a similarity of less than 95% whereas the two non O104:H4 isolates (*E. coli* ATCC 25922 and *E. coli* O126:H2) separated from the O104:H4 clusters with less than 80% similarity (Figure 1). To type the different isolates with an even higher resolution, a core-genome phylogenetic analysis based on WGS results was performed. In total 3764 ORFs were shared by all isolates analyzed and these were defined as the core genome for further phylogenetic analysis in this study. The minimum-spanning tree shows that the three 2013 EAEC Stx2a+ O104:H4 isolates clustered together with one to two-allele difference among them. One-allele difference of isolate 381-1 from isolates 338 and 381-4 was found in the gene encoding a carboxyterminal protease and was caused by a non-synonymous point mutation (1481T>A) resulting in a premature protein (Leu494Stop) in isolate 381-1. In addition, one-allele difference of isolate 381-4 from isolates 338 and 381-1 was found in a gene encoding a putative cation:proton antiport protein caused by a non-synonymous point mutation (1465A>C) resulting in an amino-acid change in the protein (Ile489Leu) in isolate 381-4. No allele differences were found between the three 2011 outbreak isolates (7N, 8G and LB227103), and they clustered together with two other reported outbreak strains (TY-2482 and 2011C-3493) with two to three-allele difference among them (Figure 2). The 2009 isolates (2009EI-2071 and 2009EI-2050) were separated from each other with a five-allele difference. The cluster of 2013 O104:H4 isolates were separated from the 2009 and 2011 O104:H4 clusters with a minimum of 20 and 27 different alleles, respectively (Figure 2 and Table S1). There were more than 193 different alleles between the typical EAEC strain 55989 (O104:H4 *stx*-negative) and the 2013 EAEC *stx*2a+ O104:H4 isolates, whereas isolate 381-3 (O126:H2, *stx*-negative) was far distinct from any of the O104:H4 strains analyzed, with more than 3291-allele differences (Figure 2).



**Figure 1. Dendrogram of *E. coli* isolates based on rep-PCR results.** Stx2a+ O104:H4 isolates of 2011 and 2013 are separated into two clusters with a similarity of less than 95% whereas the non O104:H4 *E. coli* isolates are clearly separated with a similarity of less than 80%.



**Figure 2. Core-genome phylogenetic analysis of *E. coli* strains.** The minimum spanning tree was generated by SeqSphere and based on allelic profiles comparing 3764 alleles present in all analyzed strains and defined as their core genome. Numbers in the lines indicate the number of allele differences between isolates. Black lines indicate minimum distances and red lines connect isolates with more than minimum distances. Different colors represent isolates of different time periods except isolate 381-3, which was isolated in 2013 but marked in a different color as it is *stx* negative. EAEC 55989 was used as the hypothetical EAEC Stx2a+ O104:H4 progenitor.

### DISCUSSION

In this study, EAEC Stx2a+ O104:H4 isolates obtained from stool samples of a HUS patient and her friend, travelling together to Turkey in 2013, were characterized and compared with the 2011 German EAEC Stx2a+ O104:H4 outbreak isolates and with two isolates from cases of bloody diarrhea that occurred in the Republic of Georgia in 2009. The phylogenetic relationship of these isolates was also compared with EAEC 55989, the hypothetical EAEC Stx2a+ O104:H4 progenitor strain isolated in central Africa in 1995. Besides using established molecular typing methods routinely used in our laboratory, WGS was used. Initially, the three EAEC Stx2a+ O104:H4 isolates obtained in 2013 were suspected to be similar to the 2011 German outbreak strains based on PCR results. The 2013 isolates could not be distinguished from the 2011 outbreak strains using a multiplex screening PCR targeting O104wzy, *fliCH4*, *stx2*, and *terD* genes as characteristic features of the 2011 outbreak strains (6, 20). Our microarray analyses revealed more detailed information but did not allow us to definitely discriminate the 2013 EAEC Stx2a+ O104:H4 isolates from the 2011 ones. Differences found between the 2011 and 2013 EAEC Stx2a+ O104:H4 isolates included two virulence genes (*mchB* and *mchC*) and three drug-resistant genes (*tetA*, *strA-B* and *sul2*). However, these genes are known to be located in multiple genomic islands and could therefore be easily lost or obtained (29, 30). Although rep-PCR separated the 2011 and 2013 EAEC Stx2a+ O104:H4 isolates into two clusters, this technique is not suitable for large-scale/global outbreak typing, as the inter-laboratory reproducibility is not sufficient for this. In addition, also the intra-laboratory reproducibility may vary making it necessary for an accurate typing to include all the isolates that need to be compared within the same run (22). The high resolution of WGS enabled us to discriminate the 2013 isolates from the 2011 German outbreak in the most accurate way.

In this study, a gene-by-gene typing approach also known as whole genome MLST (wgMLST), extended MLST, MLST+ or core genome MLST (cgMLST) was used (31). The EAEC Stx2a+ O104:H4 isolates of 2013 appear to differ from those of 2011 and 2009 by a minimum of 27 and 20 alleles, respectively. This suggests that 2013 isolates were relatively closer related to the 2009 clone of southeast Europe than to the 2011 outbreak clone. Unfortunately, there is no established threshold for the MLST+ approach to address intra- and inter-cluster differences yet. In addition, only the core genome was used for constructing the phylogenetic tree minimizing the chance of including mobile genetic elements (MGEs) in the analysis as there is no consensus on how MGEs should be taken into account in bacterial phylogeny (32, 33). As the estimated mutation rate of *E. coli* is reported to be 1.1 nucleotide per genome per year (34), the 27-allele difference between the 2013 and 2011 isolates suggests that the EAEC Stx2a+ O104:H4 isolates from the different time periods belong to different

## Chapter 4

---

clones. Moreover, only a one-allele difference was found between the isolate of the HUS patient and that of her friend, strongly supporting the hypothesis that either both patients obtained STEC from the same source or that one transmitted STEC to the other.

In addition to high resolution typing, WGS provided us with more detailed molecular characteristics of the isolates. This helped us to explain in more detail the phenotypic results of the isolates. For example, an insertion within the *su1* gene resulting in protein truncation was found in all 2013 EAEC Stx2a+ O104:H4 isolates, which could explain why the isolates were susceptible to trimethoprim/sulfamethoxazole in spite of the presence of the *su1* gene detected by the microarray. On the other hand, a *su2* gene existing in the 2011 but not the 2013 isolates may confer resistance to trimethoprim/sulfamethoxazole in spite of the fact that also the 2011 isolates carried the truncated *su1*.

Notably, one of the 2013 EAEC Stx2a+ O104:H4 isolates (381-1) was *bla*<sub>CTX-M-15</sub> positive as the 2011 outbreak isolates. As isolate 381-3 (*stx*-negative/O126:H2/*bla*<sub>CTX-M-15</sub>) was also obtained from the same host, it is likely that 381-1 acquired this ESBL gene from isolate 381-3 or the other way around. Our study did, however, not address a detailed analysis of the MGEs and plasmids as the aim of the study was to show the value of WGS technique in routine clinical diagnostics to identify and compare potentially pathogenic strains.

Recently, two similar isolates of EAEC Stx2a+ O104:H4 causing bloody diarrhea and HUS were reported in Belgium. Patients of both cases travelled to Tunisia and Turkey respectively, and the infection may have been acquired in those regions (35). This is similar to our and other studies describing EAEC Stx2a+ O104:H4 infections associated with travelling to Turkey, Tunisia, Egypt and North Africa (30, 36).

Our study shows that WGS has the potential to be integrated in current routine laboratories to relatively rapidly obtain reproducible and detailed information in one single step. This facilitates hospitals and public health organizations to make appropriate strategies for infection control and public health measures in real time (37). Surely, some challenges have to be overcome before applying WGS to routine diagnostics, like the availability of user-friendly bioinformatics tools, automation of the workflow and availability of genomic databases (38).

In conclusion, the detailed characterization of EAEC Stx2a+ O104:H4 isolates obtained in 2013 including the high resolution typing approach helped us to distinguish them from the 2011 German outbreak strains. It further supported the general idea that, although the 2011 outbreak is stopped, EAEC Stx2a+ O104:H4 strains highly similar to the 2011 outbreak strain in their core genome still

circulate. They still pose a serious risk for public health being a potential source for causing a new EAEC Stx2a+ O104:H4 outbreak if not monitored carefully.

### ACKNOWLEDGEMENTS

We thank Dr. Alexander Mellmann for providing us with *E. coli* O104:H4 strains of the German outbreak and Brigitte Dijkhuizen for technical assistance.

This study was partly supported by the Interreg IVA-funded projects EurSafety Health-net (III-1-02=73) and SafeGuard (III-2-03=025) and by a University Medical Center Groningen Healthy Ageing Pilots grant.

### REFERENCES

1. **Gyles CL.** 2007. Shiga toxin-producing *Escherichia coli*: an overview. *J Anim Sci* **85**.
2. **Farrokh C, Jordan K, Auvray F, Glass K, Oppegaard H, Raynaud S, Thevenot D, Condron R, De Reu K, Govaris A, Heggum K, Heyndrickx M, Hummerjohann J, Lindsay.** 2013. Review of Shiga-toxin-producing *Escherichia coli* (STEC) and their significance in dairy production. *Int J Food Microbiol* **162**:190–212.
3. **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**:1287-90.
4. **Karch H.** 2001. The role of virulence factors in enterohemorrhagic *Escherichia coli* (EHEC)-associated hemolytic-uremic syndrome. *Semin. Thromb Hemost* **27**:207–213.
5. **Welinder-Olsson C, Kaijser B.** 2005. Enterohemorrhagic *Escherichia coli* (EHEC). *Scand J Infect Dis* **37**:405–416.
6. **Bielaszewska M, Mellmann A, Zhang W, Köck R, Fruth, A, Bauwens A, Peters G, Karch H.** 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: A microbiological study. *Lancet Infect Dis* **11**:671–676.
7. **Brzuszkiewicz E, Thürmer A, Schuldes J, Leimbach A, Liesegang H, Meyer FD, Boelter J, Petersen H, Gottschalk G, Daniel R.** 2011. Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Enter-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch Microbiol* **193**:883–891.
8. **Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, Paxinos EE, Sebra R, Chin CS, Iliopoulos D, Klammer A, Peluso P, Lee L, Kislyuk AO, Bullard J, Kasarskis A, Wang S, Eid J, Rank D, Redman JC, Steyert SR, Frimodt-Møller J, Struve C, Petersen AM, Krogfelt KA, Nataro JP, Schadt EE, Waldor MK.** 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N Engl J Med* **365**:709–717.
9. **Scheutz F, Nielsen EM, Frimodt-Møller J, Boisen N, Morabito S, Tozzoli R, Nataro JP, Caprioli A.** 2011. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro Surveill* **16**.



10. Müller D, Benz I, Liebchen A, Gallitz I, Karch H, Schmidt MA. 2009. Comparative analysis of the locus of enterocyte effacement and its flanking regions. *Infect Immun* **77**:3501–3513.
11. 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.
12. Rohde H, Qin J, Cui Y, Li D, Loman NJ, Hentschke M, Chen W, Pu F, Peng Y, Li J, Xi F, Li S, Li Y, Zhang Z, Yang X, Zhao M, Wang P, Guan Y, Cen Z, Zhao X, Christner M, Kobbe R, Loos S, Oh J, Yang L, Danchin A, Gao GF, Song Y, Li Y, Yang H, Wang J, Xu J, Pallen MJ, Wang J, Aepfelbacher M, Yang R, *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium. 2011. Open-source genomic analysis of Shiga-toxin-producing *E. coli* O104:H4. *N Engl J Med* **365**:718–724.
13. Grad YH, Lipsitch M, Feldgarden M, Arachchi HM, Cerqueira GC, FitzGerald M, Godfrey P, Haas BJ, Murphy CI, Russ C, Sykes S, Walker BJ, Wortman JR, Young S, Zeng Q, Abouelleil A, Bochicchio J, Chauvin S, Desmet T, Gujja S, McCowan C, Montmayeur A, Steelman S, Fridomt-Møller J, Petersen AM, Struve C, Krogfelt KA, Bingen E, Weill FX, Lander ES, Nusbaum C, Birren BW, Hung DT, Hanage WP. 2012. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc Natl Acad Sci* **109**:3065–3070.
14. Pritchard L, Holden NJ, Bielaszewska M, Karch H, Toth IK. 2012. Alignment-free design of highly discriminatory diagnostic primer sets for *Escherichia coli* O104:H4 outbreak strains. *PLoS One* **7**:e34498.
15. Ahmed SA, Awosika J, Baldwin C, Bishop-Lilly K. a, Biswas B, Broomall S, Chain PS, Chertkov O, Chokoshvili O, Coyne S, Davenport K, Detter JC, Dorman W, Erkkila TH, Folster JP, Frey KG, George M, Gleasner C, Henry M, Hill KK, Hubbard K, Insalaco J, Johnson S, Kitzmiller A, Krepps M, Lo CC, Luu T, McNew LA, Minogue T, Munk CA, Osborne B, Patel M, Reitenga KG, Rosenzweig CN, Shea A, Shen X, Strockbine N, Tarr C, Teshima H, van Gieson E, Verratti K, Wolcott M, Xie G, Sozhamannan S, Gibbons HS; Threat Characterization Consortium. 2012. Genomic Comparison of *Escherichia coli* O104:H4 Isolates from 2009 and 2011 Reveals Plasmid, and Prophage Heterogeneity, Including Shiga Toxin Encoding Phage stx2. *PLoS One* **7**.
16. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiapello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Le Bouguéne C, Lescat M, Mangenot S, Martinez-Jéhanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tournet J, Vacherie B, Vallenet D, Médigue C, Rocha EP, Denamur E. 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**.
17. Miko A, Delannoy S, Fach P, Strockbine N. a, Lindstedt BA, Mariani-Kurkdjian P, Reetz J, Beutin L. 2013. Genotypes and virulence characteristics of Shiga toxin-producing *Escherichia coli* O104 strains from different origins and sources. *Int J Med Microbiol* **303**:410–421.
18. de Boer RF, Ferdous M, Ott A, Scheper HR, Wisselink GJ, Heck ME, Rossen JW, Kooistra-Smid AM. 2015. Assessing the public health risk of Shiga toxin-producing *Escherichia coli* by use of a rapid diagnostic screening algorithm. *J Clin Microbiol* **53**:1588–1598.
19. Scheutz F, Teel LD, Beutin L, Pierard D, Buvens G, Karch H, Mellmann A, Caprioli A, Tozzoli R, Morabito S, Strockbine NA, Melton-Celsa AR, Sanchez M, Persson S, O'Brien AD. 2012. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* **50**:2951–2963.
20. Zhang W, Bielaszewska M, Bauwens A, Fruth A, Mellmann A, Karch H. 2012. Real-time multiplex PCR for detecting Shiga toxin 2-producing *Escherichia coli* O104:H4 in human stools. *J Clin Microbiol* **50**:1752–1754.

## Comparing *Escherichia coli* O104:H4 isolates

---

21. **Geue L, Schares S, Mintel B, Conraths FJ, Müller E, Ehricht R.** 2010. Rapid microarray-based genotyping of enterohemorrhagic *Escherichia coli* serotype O156:H25/H-/Hnt isolates from cattle and clonal relationship analysis. *Appl Environ Microbiol* **76**:5510–5519.
22. **Voets GM, Leverstein-Van Hall Ma, Kolbe-Busch S, Van Der Zanden A, Church D, Kaase M, Grisold A, Upton M, Cloutman-Green E, Cantón R, Friedrich AW, Fluit AC; DiversiLab Study Group.** 2013. International multicenter evaluation of the diversilab bacterial typing system for *Escherichia coli* and *Klebsiella* spp. *J Clin Microbiol* **51**:3944–3949.
23. **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.
24. **Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Ponten T, Ussery DW, Aarestrup FM, Lund O.** 2012. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* **50**:1355-1361.
25. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV.** 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* **67**:2640-2644.
26. **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.
27. **Lima IF, Boisen N, Quetz Jda S, Havt A, de Carvalho EB, Soares AM, Lima NL, Mota RM, Nataro JP, Guerrant RL, Lima AA.** 2013. Prevalence of enteroaggregative *Escherichia coli* and its virulence-related genes in a case-control study among children from north-eastern Brazil. *J Med Microbiol* **62**:683-93.
28. **Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A.** 2014. Bacterial whole-genome sequencing revisited: Portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J Clin Microbiol* **52**:2365–2370.
29. **Guy L, Jernberg C, Arvén Norling J, Ivarsson S., Hedenström I, Melefors Ö, Liljedahl U, Engstrand L, Andersson SG.** 2013. Adaptive Mutations and Replacements of Virulence Traits in the *Escherichia coli* O104:H4 Outbreak Population. *PLoS One* **8**: 1–13.
30. **Grad YH, Godfrey P, Cerquiera GC, Mariani-Kurkdjian P, Gouali M, Bingen E, Shea TP, Haas BJ, Griggs A, Young S, Zeng Q, Lipsitch M, Waldor MK, Weill FX, Wortman JR, Hanage WP.** 2013. Comparative genomics of recent Shiga toxin-producing *Escherichia coli* O104:H4: short-term evolution of an emerging pathogen. *MBio* **4**, 1–10.
31. **Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford Sa, Jolley Ka, McCarthy ND.** 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol* **11**:728–36.
32. **Daubin V, Gouy M, Perrière G.** 2002. A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome Res* **12**:1080-90.
33. **Frost LS, Lepplae R, Summers AO, Toussaint A.** 2005. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**:722-32.
34. **Reeves PR, Liu B, Zhou Z, Li D, Guo D, Ren Y, Clabots C, Lan R, Johnson JR, Wang L.** 2011. Rates of mutation and host transmission for an *Escherichia coli* clone over 3 years. *PLoS One* **6**.
35. **De Rauw K, Vincken S, Garabedian L, Levtchenko E, Hubloue I, Verhaegen J, Craeghs J, Glupczynski Y, Mossong J, Piérard D.** 2014. Enteroaggregative Shiga toxin-producing *Escherichia coli* of serotype O104:H4 in Belgium and Luxembourg. *New Microbes New Infect* **2**:138-43.

## Chapter 4

---

36. **Ecdc/Efsa** 2011. STEC/VTEC in humans, food and animals in the EU/EEA. ECDC/EFSA joint Technical report **23**.
37. **Reuter S, Ellington MJ, Cartwright EJP, Köser CU, Török, ME, Gouliouris T, Harris SR, Brown NM, Holden MT, Quail M, Parkhill J, Smith GP, Bentley SD, Peacock SJ.** 2013. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med* **173**:1397–404.
38. **Sherry NL, Porter JL, Seemann T, Watkins A, Stinear TP, Howden BP.** 2013. Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory. *J. Clin.Microbiol* **51**,1396–1401.