International Journal of Medical Microbiology 316 (2024) 151628

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

# Enteroaggregative *Escherichia coli*: Frequent, yet underdiagnosed pathotype among *E. coli* O111 strains isolated from children with gastrointestinal disorders in the Czech Republic

Klára Schlosserová [a,b,1], Ondřej Daniel [a,b,1], Klára Labská [a,1,2], Vladislav Jakubů [a,c], Tereza Stárková [a], Jan Bílý [a], Jiří Dresler [d], Christina Lang [e], Angelika Fruth [e], Antje Flieger [e], Helena Žemličková [a,c], Martina Bielaszewska [a,b,*], Monika Havlíčková [a]

[a] Centre for Epidemiology and Microbiology, National Institute of Public Health, Šrobárova 48, Prague 100 00, Czech Republic
[b] 2nd Faculty of Medicine, Charles University, Prague, V Úvalu 84, Prague 150 06, Czech Republic
[c] 3rd Faculty of Medicine, Charles University, Prague, Ruská 87, Prague 100 00, Czech Republic
[d] Central Military Medical Institute, Military University Hospital, U Vojenské nemocnice 1200, Prague 160 01, Czech Republic
[e] Division of Enteropathogenic Bacteria and Legionella and National Reference Centre for Salmonella and other Bacterial Enteric Pathogens, Robert Koch Institute, Burgstrasse 37, Wernigerode 38855, Germany

ARTICLE INFO

*Keywords:*
Enteroaggregative *Escherichia coli* O111
Pediatric diarrhea
Czech Republic
Surveillance
Whole genome sequencing
$PAA_{O111}$ plasmid

ABSTRACT

Enteroaggregative *Escherichia coli* (EAEC) strains including those of serogroup O111 are important causes of diarrhea in children. In the Czech Republic, no information is available on the etiological role of EAEC in pediatric diarrhea due to the lack of their targeted surveillance. To fill this gap, we determined the proportion of EAEC among *E. coli* O111 isolates from children with gastrointestinal disorders ≤ 2 years of age submitted to the National Reference Laboratory for *E. coli* and Shigella during 2013–2022. EAEC accounted for 177 of 384 (46.1 %) *E. coli* O111 isolates, being the second most frequent *E. coli* O111 pathotype. Most of them (75.7 %) were typical EAEC that carried *aggR*, usually with *aaiC* and *aatA* marker genes; the remaining 24.3 % were atypical EAEC that lacked *aggR* but carried *aaiC* and/or *aatA*. Whole genome sequencing of 11 typical and two atypical EAEC O111 strains demonstrated differences in serotypes, sequence types (ST), virulence gene profiles, and the core genomes between these two groups. Typical EAEC O111:H21/ST40 strains resembled by their virulence profiles including the presence of the aggregative adherence fimbriae V (AAF/V)-encoding cluster to such strains from other countries and clustered with them in the core genome multilocus sequence typing (cgMLST). Atypical EAEC O111:H12/ST10 strains lacked virulence genes of typical EAEC and differed from them in cgMLST. All tested EAEC O111 strains displayed stacked-brick aggregative adherence to human intestinal epithelial cells. The AAF/V-encoding cluster was located on a plasmid of 95,749 bp or 93,286 bp ($pAA_{O111}$) which also carried *aggR*, *aap*, *aar*, *sepA*, and *aat* cluster. EAEC O111 strains were resistant to antibiotics, in particular to aminopenicillins and cephalosporins; 88.3 % produced AmpC β-lactamase, and 4.1 % extended spectrum β-lactamase. We conclude that EAEC are frequent among *E. coli* O111 strains isolated from children with gastrointestinal disorders in the Czech Republic. To reliably assess the etiological role of EAEC in pediatric diarrhea, a serotype-independent, PCR-based pathotype surveillance system needs to be implemented in the future.

## 1. Introduction

*Escherichia coli* strains of serogroup O111 belong to at least three different pathotypes including enteropathogenic *E. coli* (EPEC), Shiga toxin-producing (also designated enterohemorrhagic) *E. coli* (EHEC), and enteroaggregative *E. coli* (EAEC) (Campos et al., 1994). EPEC O111

* Corresponding author at: Centre for Epidemiology and Microbiology, National Institute of Public Health, Šrobárova 48, Prague 100 00, Czech Republic.
*E-mail address:* martina.bielaszewska@szu.cz (M. Bielaszewska).
[1] # These authors contributed equally to this work.
[2] Institute for Hematology and Blood Transfusion, Laboratory of Molecular Microbiology, Prague, Czech Republic

https://doi.org/10.1016/j.ijmm.2024.151628
Received 19 February 2024; Received in revised form 27 May 2024; Accepted 16 June 2024
Available online 22 June 2024

strains are characterized by the presence of the *eae* gene that encodes the major EPEC adhesin intimin and cause diarrhea in children worldwide (Croxen et al., 2013; Nataro and Kaper, 1998). EHEC O111 strains possess, in addition to *eae*, the genes encoding Shiga toxins including $stx_1$, $stx_2$ or both (Karch et al., 2005; Marejková et al., 2013; Mellmann et al., 2008). They cause, besides diarrhea, the life-threatening hemolytic uremic syndrome (HUS) (De Rauw et al., 2018; Karch et al., 2005; Mellmann et al., 2008; Valilis et al., 2018), the most common cause of acute renal failure in children (Karch et al., 2005). EHEC O111 belong, together with EHEC O26, O103, O145, and O157, to the "TOP 5" EHEC serogroups which are most frequently associated with human disease including HUS worldwide (Karch et al., 2005; Marejková et al., 2013; Mellmann et al., 2008; De Rauw et al., 2018; Valilis et al., 2018).

EAEC, an additional pathotype identified among *E. coli* O111 strains, are associated with acute and persistent diarrhea in children mainly in developing but also in industrialized countries, and cause travelers´ diarrhea (Croxen et al., 2013; Hebbelstrup Jensen et al., 2014; Jenkins, 2018). They are a highly heterogeneous group regarding their serotypes and putative virulence factors (Boisen et al., 2012; Boisen et al., 2020; Hebbelstrup Jensen et al., 2014; Jenkins, 2018). Their common characteristic is the aggregative "stacked-brick" adherence pattern to epithelial cells (Nataro et al., 1987) which is mediated by aggregative adherence fimbriae (AAF) presently known in five variants, AAF/I to AAF/V (Boisen et al., 2020; Jønsson et al., 2015). EAEC virulence genes are located on the plasmid of aggregative adherence (pAA) or on the chromosome within a pathogenicity island (Dudley et al., 2006; Nataro et al., 1994). The expression of both chromosomal and plasmid-encoded virulence genes is controlled by the transcriptional activator AggR (Dudley et al., 2006; Nataro et al., 1994). EAEC harboring the *aggR* gene are termed typical EAEC (Dudley et al., 2006; Jenkins, 2018) and many studies have strongly associated them with diarrhea (Cohen et al., 2005; Llorente et al., 2023; Pabst et al., 2003). EAEC lacking *aggR* are termed atypical EAEC (Dudley et al., 2006; Jenkins, 2018) and are considered of uncertain pathogenicity (Boisen et al., 2020), though in some studies they were significantly associated with diarrhea (Joffré and Iñiguez Rojas, 2020; Tobias et al., 2015).

Due to the severity of infections caused by EHEC strains, the surveillance of diarrheagenic *E. coli* in the Czech Republic is focused on this pathotype. EHEC infections are reportable and have been included into the system of epidemiological surveillance (Decree No. 275/2010 Coll., 2010). The surveillance is primarily based on serotyping using antisera against the "TOP 5" EHEC serogroups, which are used in all microbiological laboratories in the Czech Republic for typing *E. coli* strains isolated from stools of children with gastrointestinal disorders $\leq$ 2 years of age. According to the Czech legislation (Decree No. 275/2010 Coll., 2010), the routine microbiological laboratories are obliged to forward all *E. coli* isolates of the "TOP 5" serogroups to the National Reference Laboratory for *E. coli* and Shigella (NRL ECS) at the National Institute of Public Health, Prague, for confirmation and analysis of virulence genes. There, the isolates are tested, besides *stx* genes, also for marker virulence genes of other *E. coli* pathotypes including EPEC, EAEC, enterotoxigenic *E. coli* (ETEC), and enteroinvasive *E. coli* (EIEC) (Ileninová et al., 2022). Infections with EHEC, EPEC, ETEC, and EIEC are subsequently reported to the national Information System of Infectious Diseases (ISIN) and cases of EHEC also to The European Surveillance System (TESSy).

In contrast to other European countries where EAEC have been identified as causes of diarrhea in children (Ellis et al., 2020; Hebbelstrup Jensen et al., 2016; Huppertz et al., 1997; Llorente et al., 2023; Pabst et al., 2003; Scavia et al., 2008; Tré-Hardy et al., 2023), the etiological role of this pathotype in pediatric diarrhea in the Czech Republic remains unknown. In a pilot study performed by the NRL ECS in 2020 and 2021, we identified the EAEC pathotype in 14 % and 6 %, respectively, of *E. coli* O111 isolates from children with diarrhea obtained through the national surveillance system, but in none of isolates of the other "TOP 5" serogroups (O26, O103, O145, O157) (Ileninová et al., 2022). This prompted us to determine in the present study the proportion of EAEC among *E. coli* O111 isolates obtained through the national surveillance system during a ten-year period, 2013–2022, and to evaluate the contribution of different *E. coli* O111 pathotypes, in particular EAEC, to pediatric diarrhea. We determined the presence of marker virulence genes in EAEC O111 isolates, their adherence to human intestinal epithelial cells, and their susceptibility to antibiotics. Moreover, we performed the whole genome sequencing (WGS) of selected isolates to determine their molecular characteristics and genomic relatedness to EAEC O111 strains from other countries.

## 2. Materials and methods

### 2.1. E. coli O111 strains

Altogether, 384 *E. coli* O111 strains isolated from stool samples of children with gastrointestinal disorders $\leq$ 2 years of age between January 2013 and December 2022 were investigated. Three hundred and eighty-two of them were sent to the NRL ECS by routine microbiological laboratories from all over the Czech Republic as a part of the national surveillance of intestinal pathogenic *E. coli*. They originated from epidemiologically unrelated children who had one or more symptoms including diarrhea, abdominal pains or cramps, nausea, and vomiting. Two additional *E. coli* O111 strains were isolated from stools of patients with HUS in the NRL ECS using a protocol described previously (Marejková et al., 2013). The patients´ ages were between 2 weeks and 24 months (mean, 14.9 months; median, 15.2 months), and the male to female ratio was 1.06. Twelve patients (3.1 %) reported traveling abroad (Tunisia, Egypt, Turkey) within 2 weeks before the disease onset. Case-related data including sex, clinical symptoms, and travel history were obtained from treating physicians and/or from the national electronic reporting system ISIN.

### 2.2. Analyses of E. coli O111 strains in NRL ECS

Strains were confirmed as *E. coli* using the Microflex LT MALDI-TOF mass spectrometer (Bruker Daltonics). The O111 serogroup was determined by slide agglutination with anti-*E. coli* O111 antiserum (Mast Diagnostics) and confirmed by PCR targeting the $wbdl_{O111}$ gene, a component of the O111 lipopolysaccharide biosynthetic cluster (Perelle et al., 2007). H antigens were determined by PCR with primers targeting the flagellin subunit-encoding $fliC_{H12}$ and $fliC_{H21}$ genes (Banjo et al., 2018). All *E. coli* O111 strains were PCR tested for pathotype-specific marker virulence genes including $stx_1$, $stx_2$ (EHEC) (Scheutz et al., 2012), *eae* (EPEC) (Schmidt et al., 1994), *aggR*, *aatA*, *aaiC* (EAEC) (Boisen et al., 2008; Hebbelstrup Jensen et al., 2017; Schmidt et al., 1995), *elt*, *estIA*, *estIB* (ETEC) (Müller et al., 2007), and *ipaH* (EIEC) (Li et al., 2009). PCR primers are listed in Table 1.

### 2.3. Cell adherence assay

Human intestinal epithelial cells HCT-8 (ATCC CCL-244) were grown in RPMI 1640 medium (Thermo Fisher Scientific) with 10 % fetal bovine serum (Sigma-Aldrich), 2 mM L-glutamine, and 1 mM sodium pyruvate (both Thermo Fisher Scientific). For the adherence assay, $1\times10^5$ cells/well were seeded into 24-well plates (Corning) containing circular coverslips and grown to 80 % confluency. Monolayers were washed with phosphate-buffered saline (PBS), inoculated with $1\times10^7$ colony forming units of EAEC overnight cultures in Luria-Bertani broth, and incubated for 3 h (37 °C, 5 % $CO_2$). Cells were washed eight times with PBS, fixed (70 % methanol, 5 min), and stained (10 % Giemsa solution, 30 min) (all Sigma-Aldrich). After washing with deionised water, coverslips were mounted on glass slides and bacterial adherence patterns were examined under a light microscope (Olympus BH-2, Olympus). EAEC reference strain 042 (O44:H18) (Boisen et al., 2012) was used as a positive control and *E. coli* O111 strain 17–436 from our collection, which lacked marker genes of all pathotypes, as a negative control.

**Table 1**
PCRs used for analyses of *E. coli* O111 strains.

| Primer designation | Primer sequence (5- 3) | Target gene[a,b] | Amplicon size (bp) | Reference |
|---|---|---|---|---|
| O111-f | CGAGGCAACACATTATATAGTGCTTT | $wbdl_{O111}$ | 146 | Perelle et al., 2007 |
| O111-r | TTTTTGAATAGTTATGAACATCTTGTTTAGC | | | |
| H1/H12-f | ATGCGCTGACTGCATCAAAG | $fliC_{H1/H12}$ | 774 | Banjo et al., 2018 |
| H1/H12-r | CCTTGCCGTTGTTAGCATCG | | | |
| H21-f | GCAACTAAGCTTGCAGTGGC | $fliC_{H21}$ | 304 | Banjo et al., 2018 |
| H21-r | TCTTGGCAGCGTTCAGATCA | | | |
| Stx1-det-F1 | GTACGGGGATGCAGATAAATCGC | $stx_1$ | 209 | Scheutz et al., 2012 |
| Stx1-det-R1 | AGCAGTCATTACATAAGAACGYCCACT | | | |
| stx1a-F1 | CCTTTCCAGGTACAACAGCGGTT | $stx_{1a}$ | 478 | Scheutz et al., 2012 |
| stx1a-R2 | GGAAACTCATCAGATGCCATTCTGG | | | |
| F4 | GGCACTGTCTGAAACTGCTCCTGT | $stx_2$ | 625–627 | Scheutz et al., 2012 |
| R1 | ATTAAACTGCACTTCAGCAAATCC | | | |
| F4-f | CGCTGTCTGAGGCATCTCCGCT | | | |
| R1-e/f | TAAACTTCACCTGGGCAAAGCC | | | |
| stx2a-F2 | GCGATACTGRGBACTGTGGCC | $stx_{2a}$ | 347–349 | Scheutz et al., 2012 |
| stx2a-R2 | GCCACCTTCACTGTGAATGTG | | | |
| stx2a-R3 | CCGKCAACCTTCACTGTAAATGTG | | | |
| SK1 | CCCGAATTCGGCACAAGCATAAGC | *eae* | 863 | Schmidt et al., 1994 |
| SK2 | CCCGGATCCGTCTCGCCAGTATTCG | | | |
| aggR-f | GTATGAAATTAAAACAAACATCG | *aggR* | 800 | Boisen et al., 2008 |
| aggR-r | GTTTATTGGCTTTTAAAATAGTC | | | |
| aaiC-f | ATTGTCCTCAGGCATTTCAC | *aaiC* | 215 | Boisen et al., 2008 |
| aaiC-r | ACGACAACCCCTGATAAACAA | | | |
| Eagg-1 | CTGGCGAAAGACTGTATCAT | *aatA* | 630 | Schmidt et al., 1995 |
| Eagg-2 | CAATGTATAGAAATCCGCTGTT | | | |
| MP2-LT-F | GAACAGGAGGTTTCTGCGTTAGGTG | *elt* | 655 | Müller et al., 2007 |
| MP2-LT-R | CTTTCAATGGCTTTTTTTTGGGAGTC | | | |
| MP4-STIa-F | CCTCTTTTAGYCAGACARCTGAATCASTTG | *estIA* | 157 | Müller et al., 2007 |
| MP4-STIa-R | CAGGCAGGATTACAACAAAGTTCACAG | | | |
| MP2-STI-F | TGTCTTTTTCACCTTTCGCTC | *estIB* | 171 | Müller et al., 2007 |
| MP2-STI-R | CGGTACAAGCAGGATTACAACAC | | | |
| IpaH-F | TGACCGCCTTTCCGATA | *ipaH* | 584 | Li et al., 2009 |
| IpaH-R | TTCTCCAGCATCTCATA | | | |

[a] The genes encode the following proteins: $wbdl_{O111}$, O111 lipopolysaccharide; *fliC*, flagellin subunit; Stx, Shiga toxin; *eae*, intimin; *aggR*, transcriptional activator AggR; *aaiC*, protein of type VI secretion system; *aatA*, dispersin transporter protein; *elt*, heat-labile enterotoxin; *estIA*, *estIB*, heat-stable enterotoxins; *ipaH*, ipaH invazin.
[b] Positive control strains were as described previously including EDL933 ($stx_1$, $stx_2$, *eae*), 164/82 (*elt*, *estIA)*, 117/86 (*estIb)* (Müller et al., 2007), 042 (*aggR, aatA, aaiC*) (Boisen et al., 2012) or from the strain collection of the NRL ECS, Prague: 12/481 (*ipaH*), 10/070 ($wbdl_{O111}$), 16/938 ($fliC_{H1/H12}$), 10/250 ($fliC_{H21}$); *E. coli* K-12 strain C600 was a negative control.

### 2.4. Antimicrobial susceptibility testing

Susceptibility of EAEC O111 strains to 17 antimicrobials was tested using the broth microdilution method according to the guidelines of the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (The European Committee on Antimicrobial Susceptibility Testing, 2021). Minimum inhibitory concentrations (MICs) were determined according to the EUCAST breakpoints (The European Committee on Antimicrobial Susceptibility Testing, 2021). For confirmation of the extended-spectrum β-lactamase (ESBL) phenotype, the combination disc test according to the EUCAST guidelines was used (EUCAST, 2017). The test was performed with ceftazidime and ceftazidime/clavulanic acid, and cefotaxime and cefotaxime/clavulanic acid discs, respectively (Bio-Rad). The result was considered positive if the diameter of the inhibition zone was $\geq$ 5 mm larger with clavulanic acid than without clavulanic acid. For confirmation of the AmpC β-lactamase phenotype, the combination disc test with ceftazidime and cefotaxime discs was performed on Mueller-Hinton agar without and with oxacillin (Hrabák, 2007; EUCAST, 2017). The result was considered positive if the diameter of the inhibition zone was $\geq$ 5 mm larger on Mueller-Hinton agar with oxacillin than without oxacillin.

### 2.5. DNA extraction, Illumina sequencing and data analysis

EAEC O111 strains were grown on Columbia blood agar (Oxoid) for 20 h. Bacterial DNA was isolated from a single colony using the QIAamp DNA Mini Kit (Qiagen) following the manufacturer's protocol. The quality and concentration of purified DNA were assessed with Nanodrop and Qubit, respectively (both Thermofisher). The libraries for sequencing on a MiSeq sequencer (Illumina) were prepared following the Laboratory standard operation procedure for Pulsnet Nextera XT Library preparation (CDC, 2020). The raw reads were first checked for contamination with the Kraken 2 tool (v2.1.2) (Wood et al., 2019). The fastp (v0.23.1) tool was used for reads quality control and filtering (Chen et al., 2018). Serotypes, sequence types (STs), and virulence genes were identified by uploading the reads to SerotypeFinder 2.0 (https://cge.food.dtu.dk/services/SerotypeFinder/), MLST 2.0 (https://cge.food.dtu.dk/services/MLST/), and VirulenceFinder 2.0 (https://cge.food.dtu.dk/services/VirulenceFinder/), respectively, available on the Center for Genomic Epidemiology website (https://www.genomicepidemiology.org/). The threshold of sequence identity was set to 85 % for SerotypeFinder 2.0, and VirulenceFinder 2.0. The seven loci (*adk*, *gyrB*, *fumC*, *icd*, *mdh*, *purA*, *recA*) scheme (Wirth et al., 2006) was used for MLST 2.0.

### 2.6. Core genome multilocus sequence typing (cgMLST)

Details of 406 *E. coli* O111 strains analyzed by cgMLST are shown in Table S1. Illumina raw reads of Czech EAEC O111 strains were uploaded to the EnteroBase database (https://enterobase.warwick.ac.uk/). Data analysis was performed with the EnteroBase Backend Pipeline for cgMLST (Zhou et al., 2020). The minimum spanning tree was generated using the MSTreeV2 algorithm in the Enterobase with the GrapeTree tool (Zhou et al., 2018). Sequences of 13 Czech EAEC O111 strains were compared with sequences of 90 EAEC O111 strains from other countries available in the EnteroBase including the United Kingdom (n = 55), Spain (n = 21), Belgium (n = 5), France (n = 1), Germany (n = 1), the Netherlands (n = 1), the United States (n = 1), and Japan (n = 3); in two

strains the country of origin was unknown (Table S1). Since there is no option to identify EAEC in the EnteroBase based on marker virulence genes, such sequences were preliminarily selected as those where no pathotype was given; subsequently they were downloaded from the EnteroBase and analyzed by the AMRFinderPlus tool (v3.12.8) (Feldgarden et al., 2021). Sequences which contained *aggR* (with or without *aaiC*) were considered as typical EAEC and those which lacked *aggR* but contained *aaiC* as atypical EAEC (Table S1). In addition, the 103 EAEC O111 sequences were compared with sequences of four EAEC/EHEC O111:H21/ST40 hybrids available in the EnteroBase (Table S1) and with sequences of 299 human EHEC O111 isolates from eight countries and belonging to three different serotypes and STs including O111:H8/ST16 (n = 290), O111:H2/ST17 (n = 5), and O111:H11/ST21 (n = 4) (Table S1).

### 2.7. Nanopore sequencing and data analysis

The DNA samples used for Nanopore sequencing were the same as those for Illumina sequencing. The libraries for sequencing on GridION device (Oxford Nanopore Technologies) were prepared using Ligation sequencing gDNA - Native Barcoding Kit 24 V14 (SQK-NBD114.24) protocol (Oxford Nanopore Technologies, 2022). The basecalling was made using Dorado basecaller (v7.3.9) in super accurate singleton mode. The Nanoplot (v1.42.0) was used for read quality control (De Coster and Rademakers, 2023). The long-sequencing reads (in fastq format) were checked for contamination with the KmerFinder tool (v3.0.2) (Clausen et al., 2018; Hasman et al., 2014; Larsen et al., 2014). The worst reads were removed until 500 Mbp remain with Filtlong (v1.42.0) (Wick, 2017). The Flye assembler (v2.9.3) was used for *de novo* assembly of long reads (Kolmogorov et al. 2019). Flye assemblies and filtered long reads were used to create consensus sequence using Medaka tool (v1.11.3) with default model option (Oxford Nanopore Technologies, 2017). The Medaka consensus sequences were polished using filtered Illumina reads (see Section 2.5) using polypolish (v0.6.0) (Wick and Holt, 2022). Final hybrid assemblies were annotated using light database (v5.1) in Bakta tool (v1.9.3) (Schwengers et al., 2021) and visualized interactively via Proksee (Grant et al., 2023). Comparison of plasmid sequences was performed using the CGView server (Stothard and Wishart, 2005). For plasmid annotation in strain 19–683 mobileOG database tool (Brown et al., 2022) and CARD database tool in Proksee (Alcock et al.,. 2023) were used.

### 2.8. Data deposition

Illumina raw reads of 13 EAEC O111 strains sequenced in this study, and complete sequences of pAA plasmids from EAEC strains 19–683 and 21–582, and of pAR plasmid from strain 19–683 were deposited to the European Nucleotide Archive (ENA) (http://www.ebi.ac.uk/ena/) under study accession PRJEB72449. Illumina raw reads of EAEC O111 strains are also available in the EnteroBase database (https://enterobase.warwick.ac.uk/).

## 3. Results

### 3.1. E. coli O111 pathotypes identified in the NRL ECS in 2013–2022

PCR analyses demonstrated that 11 (2.9 %) of 384 *E. coli* O111 isolates from children ≤ 2 years with gastrointestinal disorders obtained through the national surveillance in 2013–2022 were EHEC, 183 (47.6 %) were EPEC, and 177 (46.1 %) were EAEC. No isolate was ETEC or EIEC, and 13 isolates (3.4 %) did not possess any of the pathotype-specific genes (Table 2). The annual numbers of *E. coli* O111 isolates and the proportions of different pathotypes in individual years are shown in Fig. 1. The numbers of *E. coli* O111 isolates received by the NRL ECS had an increasing trend between 2013 and 2019 and subsequently decreased (Fig. 1A). In 2014–2019, EAEC and EPEC pathotypes accounted for similar proportions of annual *E. coli* O111 isolates (45 %-58 % and 38 %-48 %, respectively) (Fig. 1B). In contrast, in 2013, EAEC was the predominant pathotype (74 % of isolates), while in 2020–2022, EPEC (69 %-89 % of isolates) prevailed over EAEC (7 %-28 % of isolates) (Fig. 1B).

**Table 2**
Pathotypes of 384 *E. coli* O111 strains analyzed in NRL ECS in 2013–2022.

| Pathotype (marker virulence genes) | Total number (%) of strains | Virulence gene profile | Number (%) of pathotype strains with virulence gene profile |
|---|---|---|---|
| EHEC ($stx_1$, $stx_2$) | 11 (2.9) | $stx_{1a}$ | 6 (54.5) |
| | | $stx_{2a}$ | 2[a] (18.2) |
| | | $stx_{1a}$+$stx_{2a}$ | 3 (27.3) |
| EPEC (*eae*) | 183 (47.6) | *eae* | 183 (100) |
| EAEC (*aggR*, *aatA*, *aaiC*) | 177 (46.1) | *aggR*+*aatA*+*aaiC* | 116 (65.5) |
| | | *aggR*+*aaiC* | 17 (9.6) |
| | | *aggR* | 1 (0.6) |
| | | *aaiC*+*aatA* | 19 (10.7) |
| | | *aaiC* | 19 (10.7) |
| | | *aatA* | 5 (2.8) |
| ETEC (*elt*, *estlA*, *estlB*) | 0 | *elt*, *estlA*, *estlB* | 0 |
| EIEC (*ipaH*) | 0 | *ipaH* | 0 |
| None of tested | 13 (3.4) | None of tested | 13 |

[a] Both strains originated from HUS patients; one strain lost $stx_{2a}$.

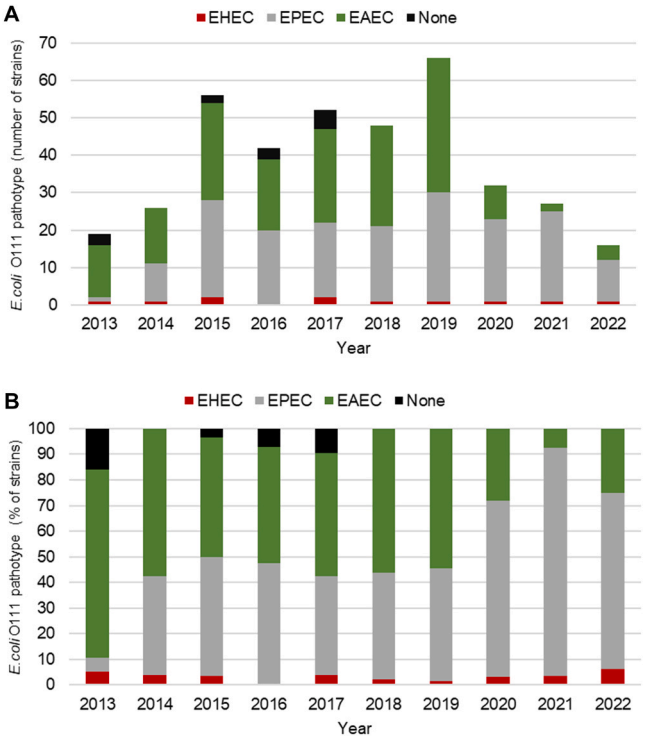


**Fig. 1.** Distribution of *E. coli* O111 pathotypes during 2013–2022. (A) Numbers and (B) percentages of strains of the particular pathotypes. EHEC, enterohemorrhagic *E. coli*; EPEC, enteropathogenic *E. coli*; EAEC, enteroaggregative *E. coli*.

### 3.2. Marker virulence gene profiles of EAEC O111 and their distribution during 2013–2022

To characterize EAEC O111 strains, we first determined their marker virulence gene profiles. The majority of the strains (134 of 177, 75.7 %) were typical EAEC which possessed *aggR*, mostly (in 116 strains) together with *aatA* (encoding dispersin transporter protein) and *aaiC* (encoding type VI secreted protein). The remaining 18 typical EAEC

strains contained *aggR* together with *aaiC* or alone (Table 2). Forty-three of 177 EAEC O111 strains (24.3 %) were atypical EAEC that lacked *aggR* but carried *aaiC* and/or *aatA* (Table 2). Distribution of strains with different virulence gene profiles in individual years of the 2013–2022 period is shown in Fig. 2. Except for 2019 and 2022, typical EAEC prevailed over atypical EAEC, accounting for 55.6 %-100 % of annual isolates (Fig. 2A, B). In comparison, in 2019 and 2022, atypical EAEC were more frequent being responsible for 55.6 % and 75 %, respectively, of EAEC O111 strains (Fig. 2A, B).

### 3.3. Characterization of EAEC O111 strains by whole genome sequencing

Thirteen EAEC O111 strains including 11 typical and two atypical were subjected to Illumina WGS. The strains were randomly selected from the 2018–2021 period when broad spectra of marker virulence gene profiles were observed (Fig. 2) and the introduction of WGS into the NRL ECS allowed sequencing of freshly isolated strains. H antigens, STs, and putative virulence genes obtained through WGS analysis are listed in Table 3. All typical EAEC strains belonged to the serotype O111:H21 and ST40, whereas atypical EAEC strains belonged to the serotype O111:H12 and ST10. The typical and atypical EAEC O111 strains also differed in their virulence gene profiles. Typical EAEC strains possessed, in addition to *aggR*, the *aaiC*, *aap*, *aar*, and *anr* genes encoding the secreted protein, dispersin, AggR-activated regulator, and AraC negative regulator, respectively. They also contained the gene cluster (*agg5A*, *agg3B*, *agg3C*, *agg3D*) encoding the aggregative adherence fimbriae AAF/V (Jønsson et al., 2015), *astA* encoding the heat-stable enterotoxin EAST-I, *shiaA*, a component of the *Shigella flexneri* SHI-2 pathogenicity island, and the *espI*, *pic*, *sat*, and *sepA* genes encoding the serine protease autotransporters of *Enterobacteriaceae* (SPATEs) (Boisen et al., 2009).

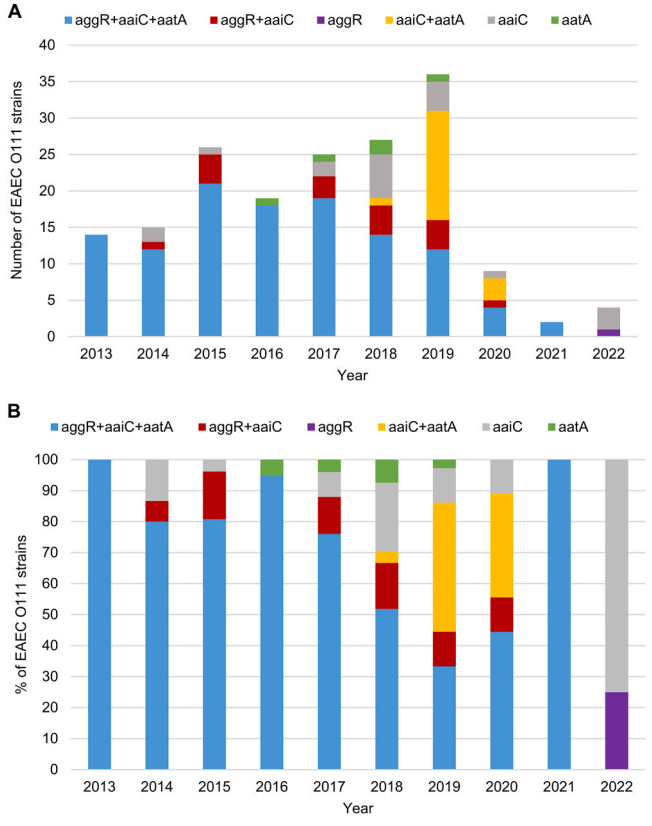


**Fig. 2.** Distribution of EAEC O111 strains with different marker virulence gene profiles determined by PCR during 2013–2022. (A) Numbers and (B) percentages of isolates. Typical EAEC are defined by the presence of *aggR*, either alone or in combination with *aatA* and/or *aaiC*. Atypical EAEC lack *aggR* and carry *aatA* and/or *aaiC*.

Besides the AAF/V-encoding gene cluster, all typical EAEC O111:H21/ST40 strains also contained genes encoding other putative adhesins such as the afimbrial adhesin AfaD, curli fimbriae, intimin-like adhesin FdeC, type I fimbriae, Iha (iron-regulated gene A homologue adhesin), long polar fimbriae, and the YHD fimbrial cluster (*yehA*, *yehB*, *yehC*, *yehD*) (Table 3). They also carried several genes which frequently occur in extraintestinal pathogenic *E. coli* such as *iss*, *iut*, *iucC*, *fyuA* and *irp2,* as well as the *terC* gene (Table 3), a part of the tellurite resistance-encoding cluster which also occurs in EHEC O111 (Karch et al., 2005; Marejková et al., 2013).

Two atypical EAEC O111:H12/ST10 strains partially differed from each other in their virulence gene contents, but both lacked, besides *aggR*, also *aatA*, *aap*, *aar*, *sepA*, and the AAF/V-encoding cluster (Table 3). They also lacked the gene clusters encoding AAF/I-AAF/IV, as well as the *afp* cluster encoding aggregate-forming pili and *cseA* encoding ETEC colonization factor CS22 that have recently been found in EAEC strains lacking all known AAF variants (Boisen et al., 2020; Lang et al., 2018; Llorente et al., 2023). On the other hand, they shared with typical EAEC O111 isolates the genes encoding other adhesins (*csgA*, *fimH*, *iha*, and the *yehA*, *yehB*, *yehC*, *yehD* cluster), and *astA*, *pic, fyuA*, *irp2*, and *terC* genes (Table 3). In contrast to typical EAEC O111, both atypical isolates contained the genes *mchB*, *mchC*, and *mchF* encoding the antibacterial peptide antibiotic microcin H47 and its secretion system (Rodríguez et al., 1999) (Table 3).

### 3.4. Localization of AAF/V-encoding gene cluster on pAA plasmid

To find out if the AAF/V-encoding cluster of EAEC O111:H21/ST40 strains is located on a pAA plasmid, we performed Nanopore sequencing of four randomly selected Illumina-sequenced strains (19−683, 19−806, 20−243, 21−582) and used hybrid assembly approach to reconstruct plasmids. pAA plasmids were identified in all strains. Visualization of pAA from strain 19–683 ($pAA_{O111-19-683}$) (Fig. 3) demonstrated that it contained the AAF/V-encoding gene cluster (*agg5A*, *agg3B*, *agg3C*, *agg3D*) and several other EAEC virulence genes including *aggR*, *aap*, *aar*, *aatPABCD* cluster, and *sepA*. The $pAA_{O111-19-683}$ was identical in the size (95,749 bp) and the gene content including the presence of EAEC virulence genes to pAA plasmids from strains 19–806 ($pAA_{O111-19-806}$) and 20–243 ($pAA_{O111-20-243}$) (Fig. S1). In comparison, pAA from strain 21–582 ($pAA_{O111-21-582}$) was smaller (93,286 bp) but it also carried all EAEC virulence genes detected in the other pAA plasmids (Fig. S1). It lacked a region that contained genes encoding a transposase and a ClbS/DfsB family four-helix bundle protein in the other pAA plasmids (Fig. S1). All four pAA plasmids belonged to the IncFII/IncFIB incompatibility group.

### 3.5. Genomic relatedness of EAEC O111 strains from the Czech Republic and other countries based on cgMLST

To determine the genomic relatedness of EAEC O111 strains from the Czech Republic and other countries, the core genome coding regions of 13 Czech strains were compared to each other and to those of 90 EAEC O111 strains from eight other countries around the world available in the EnteroBase database (Table S1). The minimum spanning tree based on allelic profiles derived from the scheme of 2512 *Escherichia/Shigella* core genome genes in the EnteroBase demonstrated that there were two distinct clusters among the Czech EAEC O111 strains (Fig. 4). The first included 11 typical EAEC O111:H21/ST40 strains which differed from each other by a maximum of 39 alleles. The second included two atypical EAEC O111:H12/ST10 strains which differed from each other by 48 alleles and from the EAEC O111:H21/ST40 cluster by 2164 alleles (Fig. 4).

Comparison of Czech EAEC O111 strains with 90 typical and atypical EAEC O111 strains from eight other countries is shown in Fig. 5. The 11 Czech typical EAEC O111:H21/ST40 strains clustered together with EAEC O111:H21/ST40 strains from European countries with a

**Table 3**
Virulence genes of EAEC O111 strains identified by whole genome sequencing.

| Gene[a] | Protein | EAEC O111 strain (number and sequence type; ST) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 19–725 ST40 | 20–243 ST40 | 19–698 ST40 | 19–325 ST40 | 19–422 ST40 | 19–683 ST40 | 19–615 ST40 | 19–806 ST40 | 19–815 ST40 | 20–019 ST40 | 21–582 ST40 | 18–510 ST10 | 19–457 ST10 |
| *fliC* | Flagellin subunit | H21 | H21 | H21 | H21 | H21 | H21 | H21 | H21 | H21 | H21 | H21 | H12 | H12 |
| *aaiC* | Type VI secretion protein | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *aap* | Dispersin | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *aar* | AggR-activated regulator | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *aatA* | Dispersin transporter protein | - | + | - | - | - | + | + | + | + | + | + | - | - |
| *afaD* | Afimbrial adhesin | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *agg5A* | AAF/V major subunit | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *agg3B* | AAF/III minor adhesin | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *agg3C* | Usher; AAF/III assembly unit | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *agg3D* | Chaperone; AAF/III assembly unit | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *aggR* | AraC transcriptional activator | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *anr* | AraC negative regulator | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *astA* | Heat-stable enterotoxin EAST-I | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *capU* | Hexosyltransferase homologue | - | - | - | - | - | - | - | - | - | - | - | - | + |
| *cia* | Colicin Ia | - | - | - | - | - | + | - | - | + | - | - | - | - |
| *cib* | Colicin Ib | - | - | - | - | - | + | - | - | - | - | - | - | - |
| *csgA* | Curli major subunit CsgA | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *espI*[b] | *E. coli* secreted protease | + | + | + | + | + | + | + | + | + | + | - | - | - |
| *fdeC* | Intimin-like adhesin FdeC | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *fimH* | Type 1 fimbriae | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *fyuA* | Siderophore receptor | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *gad* | Glutamate decarboxylase | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *hha* | Hemolysin expression modulator | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *hlyE* | *E. coli* hemolysin E | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *iha* | Adhesin, siderophore receptor | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *irp2* | High molecular weight protein 2 non-ribosomal peptide synthetase | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *iss* | Increased serum survival | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *iucC* | Aerobactin synthetase | + | + | + | + | + | + | + | + | + | + | + | - | + |
| *iutA* | Ferric aerobactin receptor | + | + | + | + | + | + | + | + | + | + | + | - | + |
| *lpfA* | Long polar fimbriae major subunit | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *mchB* | Microcin H47 part of colicin H | - | - | - | - | - | - | - | - | - | - | - | + | + |
| *mchC* | MchC protein | - | - | - | - | - | - | - | - | - | - | - | + | + |
| *mchF* | ABC transporter protein MchF | - | - | - | - | - | - | - | - | - | - | - | + | + |
| *nlpI* | lipoprotein NlpI precursor | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *ompT* | Outer membrane protease | + | + | + | + | + | + | + | + | + | + | + | - | - |
| ORF3 | Isoprenoid biosynthesis | - | - | + | + | - | + | + | + | + | + | + | - | + |
| ORF4 | Putative isopentenyl-diphosphate delta-isomerase | - | + | + | + | + | + | + | + | + | + | + | - | + |
| *pic*[b] | Protein involved in intestinal colonization | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *sat*[b] | Secreted autotransporter toxin | + | + | + | + | + | + | + | + | + | + | + | - | + |
| *sepA*[b] | Shigella extracellular protease | + | + | + | + | + | + | + | + | + | + | + | - | - |
| *shiA* | Homolog of *Shigella flexneri* SHI-2 pathogenicity island gene *shiA* | + | + | + | + | + | + | + | + | + | + | + | - | + |
| *terC* | Tellurite resistance | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *traT* | Outer membrane protein complement resistance | - | - | - | - | - | + | + | - | + | + | - | + | + |
| *yehA* | Outer membrane lipoprotein, YHD fimbrial cluster | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *yehB* | Usher, YHD fimbrial cluster | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *yehC* | Chaperone, YHD fimbrial cluster | + | + | + | + | + | + | + | + | + | + | + | + | + |
| *yehD* | Major pilin subunit, YHD fimbrial cluster | + | + | + | + | + | + | + | + | + | + | + | + | + |

[a] +, the gene is present; -, the gene is absent.
[b] serine-protease autotransporter of *Enterobacteriaceae* (SPATE).

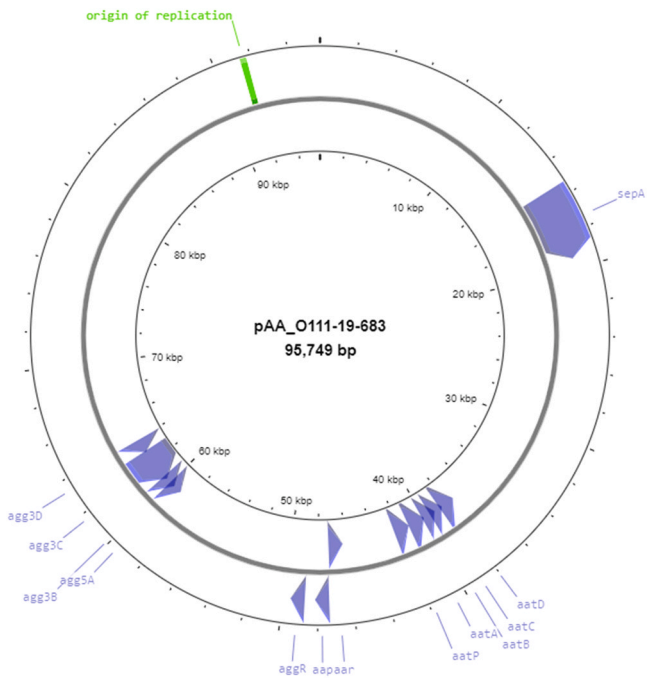


**Fig. 3.** Visualization of pAA plasmid from EAEC O111:H21/ST40 strain 19–683 ($pAA_{O111-19-683}$) harboring the AAF/V-encoding cluster (*agg5A, agg3B, agg3C, agg3D*). Other EAEC virulence genes located on $pAA_{O111-19-683}$ (*aggR, aap, aar, aatPABCD* cluster, and *sepA*) are also shown. The plasmid was visualized using Proksee web server and annotated with Bakta tool.

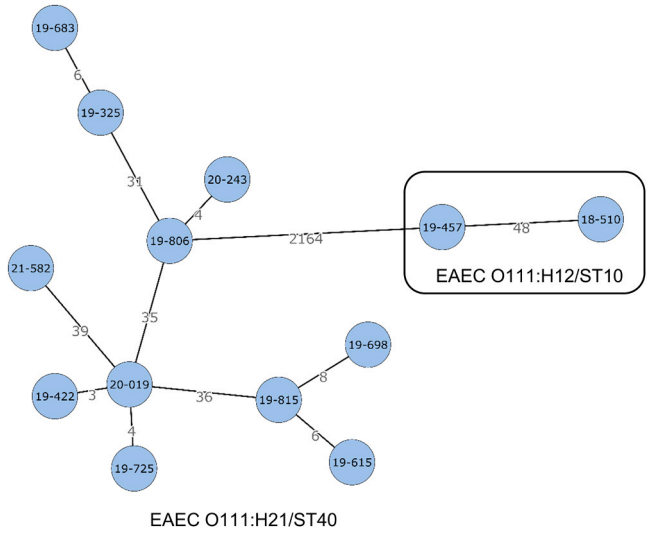


**Fig. 4.** Genomic relatedness of Czech EAEC O111:H21/ST40 and EAEC O111:H12/ST10 strains based on cgMLST analysis. The minimum spanning tree represents visualization of allelic profiles derived from the scheme of 2512 *Escherichia/Shigella* core genome genes in the EnteroBase. Each circle with strain number represents a given allelic profile. The numbers on the connecting lines illustrate the numbers of alleles by which the respective strains differ.

maximum of 33 allelic differences between Czech strain 19–815 and the most closely related Spanish strain 3150–15 (Fig. 5). Strains from the Czech Republic and other European countries (Belgium, Spain, the United Kindgom) tended to form smaller clusters of more closely related strains (Fig. 5). Except for four Belgian strains isolated during a nursery outbreak (Tré-Hardy et al., 2023), no epidemiological links between strains in such clusters were known (Llorente et al., 2023; Ellis et al., 2020; this study). Notably, the only strain from the United States (EC_1488) differed from two most closely related European strains (both from the United Kingdom) by 78 and 135 alleles, respectively. Three strains from Japan were even more distant, differing by 175 alleles from the United States strain but being related to each other (22 and 26 allelic differences). The cgMLST analysis thus indicated a close genomic relatedness of typical EAEC O111:H21/ST40 strains across Europe and a more distant genomic relatedness of such strains within and outside Europe. In contrast, the EAEC O111:H21/ST40 cluster was distant (2170 allelic differences) from the EAEC O111:H12/ST10 cluster that consisted of five strains from the Czech Republic, France, Germany and an unknown country (Fig. 5).

### 3.6. Genomic relatedness of EAEC O111, EHEC O111, and EAEC/EHEC O111 hybrids based on cgMLST

*E. coli* O111 belongs to several serogroups (O23, O59, O86, O104, O181) in which hybrid pathotypes combining virulence traits of EAEC and EHEC have occasionally been reported (Bielaszewska et al., 2011; Dallman et al., 2012; Iyoda et al., 2000; Kislichkina et al., 2022; Lang et al., 2018; Mellmann et al., 2011; Morabito et al., 1998; Prager et al., 2014; Rasko et al., 2011). In the case of a highly virulent EAEC/EHEC O104:H4 hybrid (Bielaszewska et al., 2011; Mellmann et al., 2011) it was hypothesized that it originated from an EAEC progenitor that acquired a Stx2-encoding bacteriophage (Kampmeier et al., 2018; Rasko et al., 2011). To test this hypothesis for EAEC/EHEC O111 hybrids, we compared the core genome sequences of EAEC, EHEC, and EAEC/EHEC O111 strains available in the EnteroBase. As demonstrated in Fig. 6, each of the pathotype/serotype/ST groups including EAEC O111:H21/ST40, EAEC O111:H12/ST10, EHEC O111:H8/ST16 (the most common EHEC O111 group), EHEC O111:H2/ST17, and EHEC O111:H11/ST21 formed a distinct cluster which differed from the neighbouring cluster by 822–2170 alleles. The different EHEC serotypes/STs were unrelated to each other (Fig. 6). In contrast, four EAEC/EHEC O111:H21/ST40 hybrids were closely related to EAEC O111:H21/ST40 strains but unrelated to any of the EHEC O111 serotype/ST groups (Fig. 6). This indicated that similar to *E. coli* O104:H4 hybrid, the EAEC/EHEC O111:H21 hybrids evolved from an EAEC O111:H21 progenitor by acquisition of a $stx_2$-harboring bacteriophage.

### 3.7. Adherence of EAEC O111 strains to human intestinal epithelial cells

The 13 sequenced EAEC O111 strains were tested for their adherence to human intestinal epithelial cells HCT-8. We used these cells instead of the reference HEp-2 line (Nataro et al., 1987) since their intestinal origin better reflects the situation during EAEC infection. Examples of adherence patterns of four EAEC O111:H21/H40 strains that possessed the AAF/V-encoding cluster and two EAEC O111:H12/H10 strains that lacked all presently known AAF variants (Table 3) are shown in Fig. 7. All strains whether or not they possessed AAF/V displayed the stacked-brick aggregative adherence (Fig. 7B-G) which was similar to that of the reference EAEC strain 042 used as a positive control (Fig. 7A). In contrast, only single rare bacteria adhered to HCT-8 cells infected with an *E. coli* O111 strain which lacked virulence genes of all pathotypes and served as a negative control (Fig. 7H).

### 3.8. Antimicrobial susceptibilities of EAEC O111 strains

One hundred and forty-five EAEC O111 strains which represented different marker virulence gene profiles were tested for susceptibilities to ampicillin, ampicillin/sulbactam, piperacillin/tazobactam, cefoxitin, cefotaxime, ceftazidime, cefepime, meropenem, trimethoprim, trimethoprim/sulfamethoxazole, ciprofloxacin, ceftolozane/tazobactam, gentamicin, amikacin, tobramycin, colistin, and tigecycline. The proportions of strains that were resistant to each antimicrobial are summarized in Fig. 8. Data for individual strains including their MICs, ESBL and AmpC phenotypes are shown in Table S2. There was a high

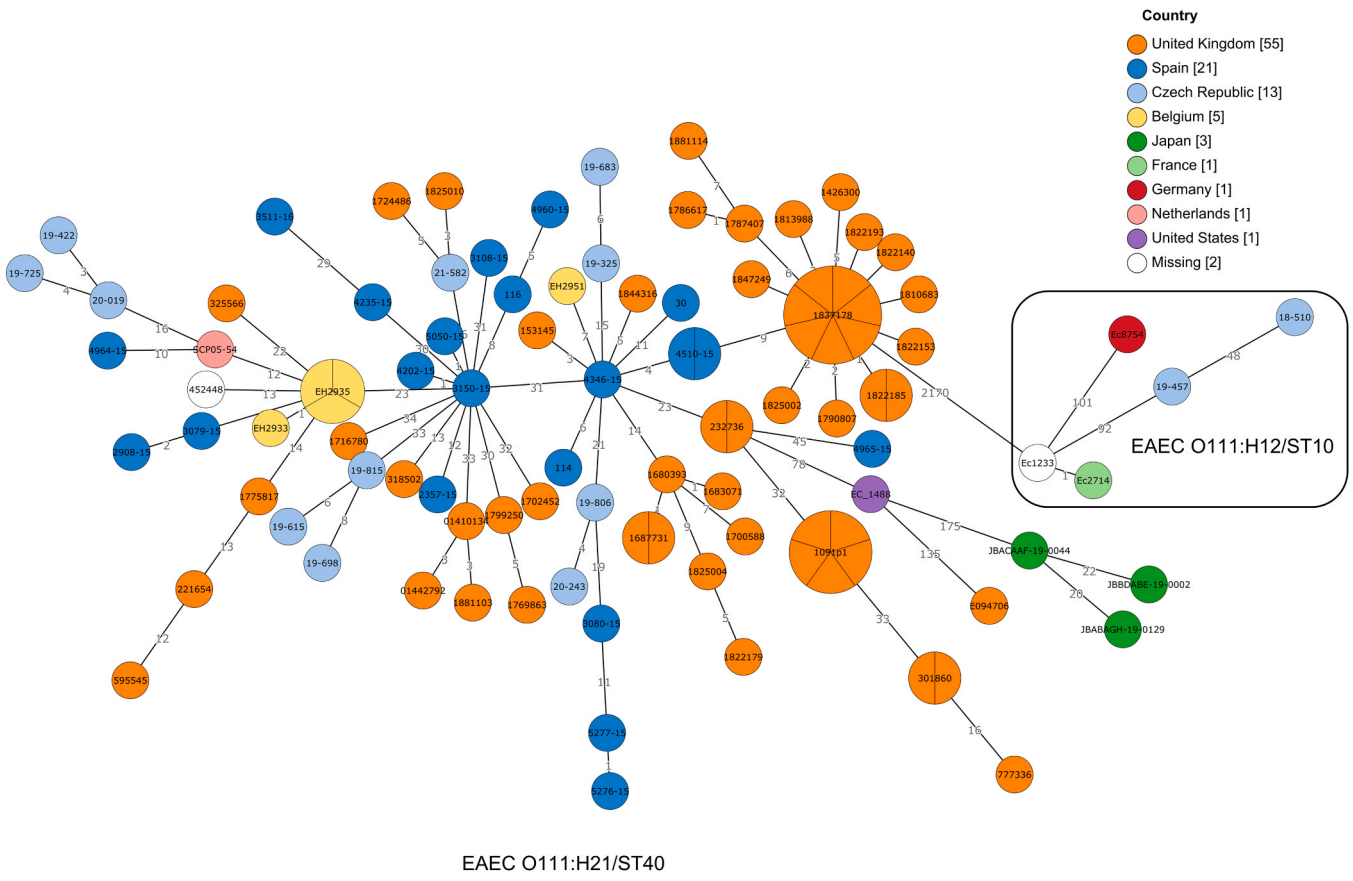


**Fig. 5.** Genomic relatedness of Czech EAEC O111 strains with EAEC O111 strains from other countries based on cgMLST. The minimum spanning tree represents visualization of allelic profiles derived from the scheme of 2512 *Escherichia/Shigella* core genome genes in the EnteroBase. Each circle with strain number represents a given allelic profile. The sizes of circles correspond to the number of strains within the circle. Circles that contain more than one strain are designated by the number of one strain and numbers of all strains are given in parentheses as follows: Circle EH2935 (EH2935, EH2934, EH2932); circle 4510–15 (4510–15, 4522–15); circle 1687731 (1687731, 1687755); circle 232736 (232736, 232726); circle 1837178 (1837178, 1831243, 1822206, 1825005, 1826064, 1825006, 1825003); circle 1091p1 (1091p1, 767p4, 657p1, 1337p1, 1150p1); circle 301860 (301860, 249369); circle 1822185 (1822185, 1822207). The numbers on the connecting lines illustrate the numbers of alleles by which the respective strains differ. Strains from different countries are distinguished by the colors of the circles and their numbers are given in the legend above the tree. Two major clusters including EAEC O111:H21/ST40 and EAEC O111:H12/ST10 are indicated. The Czech EAEC O111 strains were sequenced in this study, sequences of the other strains originated in the EnteroBase. The strains are listed in Table S1.

level of resistance, in particular to aminopenicillins and cephalosporins (Fig. 8). One hundred and thirty-nine (95.9 %) strains were resistant to ampicillin, 137 (94.5 %) strains to ampicillin/sulbactam, and 109 (75.2 %) strains to piperacillin/tazobactam. Most of strains (130, 89.7 %) were also resistant to the second-generation cephalosporin cefoxitin, and 16 (11.0 %) and 89 (61.4 %) strains to the third-generation cephalosporins cefotaxime and ceftazidime, respectively. Six (4.1 %) strains were, in addition, resistant to the fourth-generation cephalosporin cefepime and displayed the ESBL phenotype (Fig. 8, Table S2). All strains that were resistant to the second- and/or third-generation cephalosporins (128 of 145, 88.3 %) expressed the AmpC phenotype (Fig. 8). Notably, most of AmpC β-lactamase producers (100 of 128, 78.1 %) were typical EAEC, whereas most of ESBL producers (5 of 6, 83.3 %) were atypical EAEC. All 145 strains were susceptible to the fifth-generation cephalosporin ceftolozan/tazobactam, to meropenem, and ciprofloxacin (Fig. 8). Only few strains were resistant to trimethoprim/sulfamethoxazole and/or trimethoprim (7 strains, 4.8 %), or to tigecycline (4 strains, 2.8 %). One strain (0.7 %) was resistant to gentamicin and tobramycin, and one strain (0.7 %) to colistin (Fig. 8). Two strains (1.4 %) were multidrug resistant (MDR) (Fig. 8) demonstrating resistance to: i) β-lactams, trimethoprim, and tigecycline (strain 17–549), and ii) β-lactams, trimethoprim, trimethoprim/sulfamethoxazole, and aminoglycosides (strain 15–812) (Table S2). Only 4 strains, all atypical EAEC, were susceptible to all antimicrobials tested (Table S2).

### 3.9. Antimicrobial resistance determinants identified by whole genome sequencing

To determine the molecular basis of antimicrobial resistance, Illumina sequences of 13 EAEC O111 strains were analyzed for antimicrobial resistance determinants. Comparison of the antimicrobial resistance phenotypes and genotypes is shown in Table 4. Sequences of all 11 O111:H21/ST40 strains that expressed the AmpC phenotype displayed the mutation C to T at the position −42 of the *ampC* promoter region (Table 4) which has been reported to be responsible for overexpression of the chromosomal AmpC β-lactamase conferring resistance to the second and third generation cephalosporins (Bogaerts et al., 2010; Päivärinta et al., 2016). None of these strains carried any of the plasmid-mediated AmpC genes including $bla_{CMY}$, $bla_{MOX}$, $bla_{FOX}$ $bla_{DHA}$, $bla_{ACT}$, $bla_{MIR}$, $bla_{CIT}$, $bla_{ACC}$, $bla_{LAT}$, and $bla_{BIL}$ (Pérez-Pérez and Hanson, 2002). This indicated that the AmpC phenotype of the EAEC O111:H21/ST40 strains resulted from overexpression of the chromosomal *ampC* gene due to the mutation in its promoter region rather than from the expression of a plasmid-mediated AmpC. The EAEC O111:H12/ST10 strain 18–510 contained, in accordance with its ESBL phenotype, the $bla_{CTX\text{-}M\text{-}15}$ gene encoding ESBL CTX-M-15 (Table 4). The other EAEC O111:H12/ST10 strain 19–457 that was susceptible to all antimicrobials tested (Table S2) did not contain any antimicrobial resistance determinant (Table 4). Two EAEC O111 strains were also resistant to

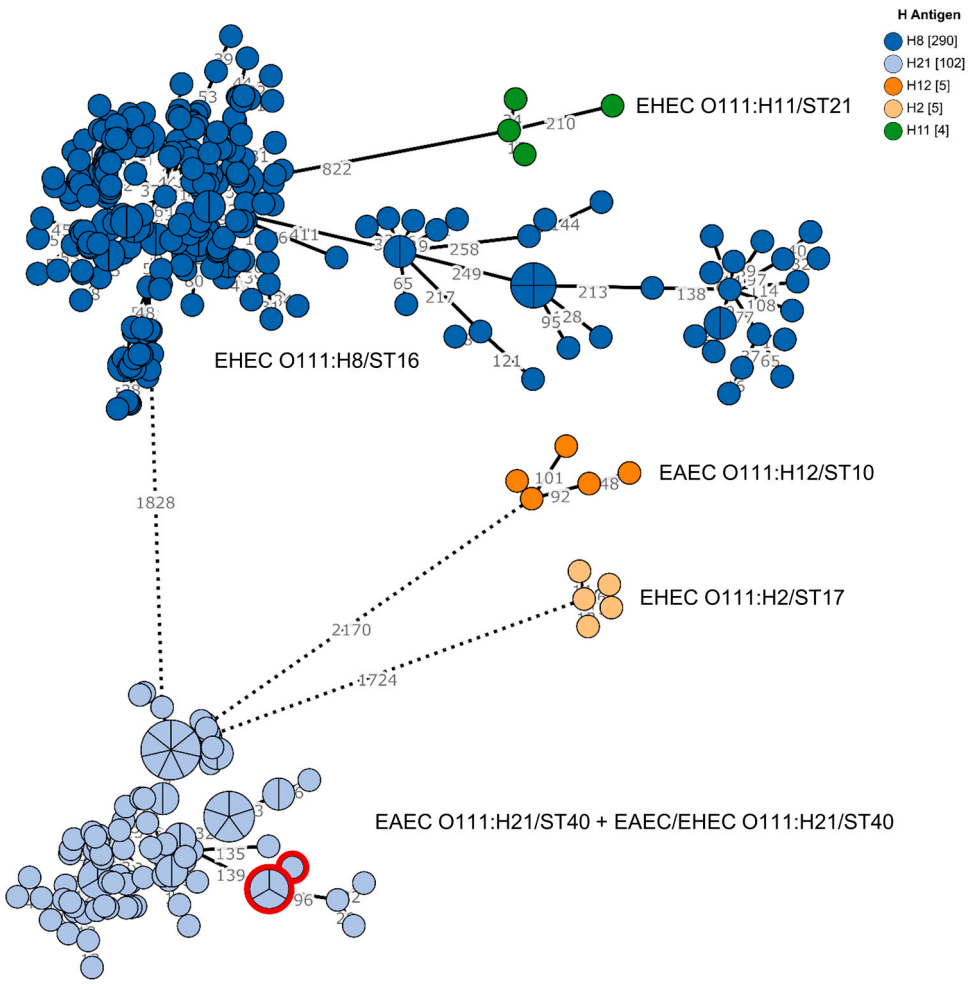


**Fig. 6.** Genomic relatedness of EAEC O111, EHEC O111, and EAEC/EHEC O111 hybrids. The minimum spanning tree is based on allelic profiles derived from the scheme of 2512 *Escherichia/Shigella* core genome genes in the EnteroBase. Each circle represents a given allelic profile. The sizes of the circles correspond to the number of strains within the circle and the circles that contain more than one strain are divided into the respective parts. Clusters of strains of different pathotypes/serotypes/STs are distinguished by the colors of the circles (according to H antigens) and their numbers are given in the legend above the tree. The respective information is also given near to each cluster. Red-bordered circles in the „EAEC O111:H21/ST40 + EAEC/EHEC O111:H21/ST40" cluster designate four EAEC/EHEC hybrids. The numbers on the connecting lines illustrate the numbers of alleles by which the respective clusters and strains differ. The Czech strains were sequenced in this study, sequences of the other strains originated in the EnteroBase. The strains are listed in Table S1.

non-β-lactam antibiotics (Table 4). In strain 19–815, resistance to trimethoprim and trimethoprim/sulfamethoxazole correlated with the presence of the *dfrA14* and *sul2* genes (Table 4). In comparison, strain 19–683 which also carried *sul2* but not *dfrA14* was susceptible to trimethoprim/sulfamethoxazole (Table 4, Table S2) since the *sul2* gene only without *dfrA14* did not confer resistance to this combined drug. The presence of the *aph(3")-lb* and *aph(6)-Id* genes that encode resistance to streptomycin (Shaw et al.,1993) in strain 19–683 (and also in strain 19–815) could not be correlated with the respective phenotype since streptomycin is used in the Czech Republic only for the treatment of tuberculosis and is therefore not part of the set of antibiotics used for susceptibility testing of *Enterobacteriaceae*.

Nanopore sequencing and subsequent hybrid assembly of plasmids from strain 19–683 demonstrated that besides the pAA plasmid (Fig. 3), this strain also harbored a plasmid of 103,622 bp that carried the *sul2*, *aph(3")-lb*, *aph(6)-Id*, and $bla_{\text{TEM-1 C}}$ genes (Fig. S2). We designated this plasmid $pAR_{O111\text{–}19\text{–}683}$. It belonged to the IncB/O/K/Z incompatibility group.

## 4. Discussion

This is the first study that investigated the occurrence of EAEC in the stools of children with diarrhea and other gastrointestinal disorders in the Czech Republic. Using the current surveillance system of diarrheagenic *E. coli*, which is based on serotyping of *E. coli* isolates from children ≤ 2 years in routine microbiological laboratories and subsequent pathotype detection in the NRL ECS, we demonstrated that EAEC accounted for almost half (46.1 %) of all *E. coli* O111 isolates during 2013–2022. Notably, only 3.1 % of the patients from whom the EAEC O111 strains originated reported a recent travel to the countries where EAEC infections are frequent such as Tunisia, Egypt, and Turkey (Jenkins, 2018). This strongly suggests that EAEC O111 are important domestically acquired pathogens responsible for pediatric diarrhea in the Czech Republic, similar to what was reported in Spain (Llorente et al., 2023). Although the focus on a single serogroup may be a limitation of our study, the validity and epidemiological significance of our findings is supported by a recent report from Spain that O111:H21 was the second most frequent EAEC serotype identified in children with diarrhea (Llorente et al., 2023), by the high intrinsic potential of EAEC O111:H21/ST40 strains to cause diarrhea (Ellis et al., 2020), and by the ability of EAEC O111 strains to cause outbreaks (Hebbelstrup Jensen et al., 2016; Tré-Hardy et al., 2023; Yatsuyanagi et al., 2002).

A noteworthy finding in our study is that the current surveillance system of intestinal pathogenic *E. coli* based on serotyping rather than on

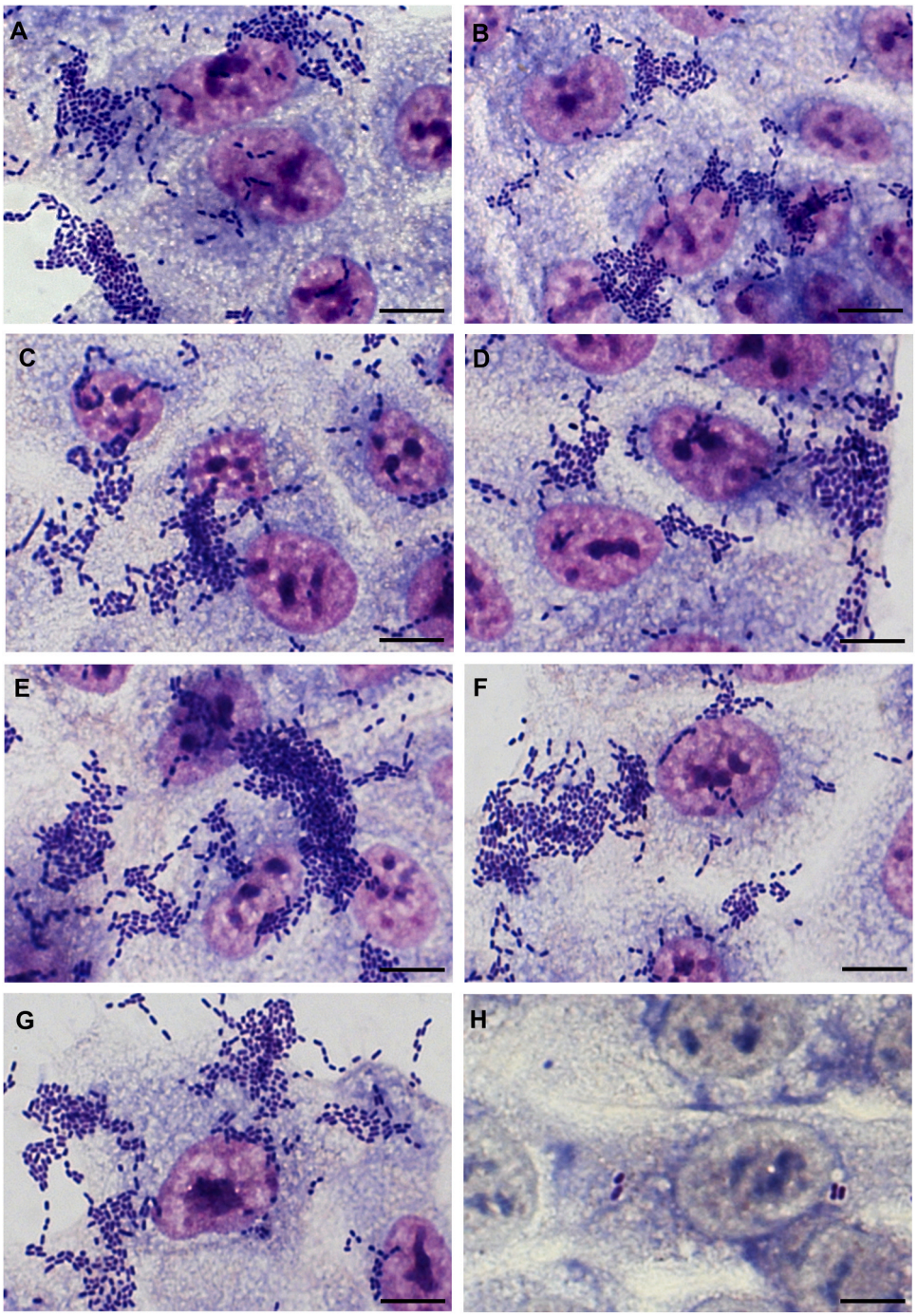


**Fig. 7.** Adherence of EAEC O111 strains to human intestinal epithelial cells HCT-8. (A) EAEC reference strain 042 (positive control). (B-E) Typical EAEC O111:H21/ST40 strains containing AAF/V-encoding gene cluster: (B) 19–725; (C) 20–243; (D) 19–422; (E) 20–019. (F-G) Atypical EAEC O111:H12/ST10 strains lacking AAF/V-encoding gene cluster: (F) 18–510; (G) 19–457. (H) *E. coli* O111 strain 17–436 lacking virulence genes of EAEC and of all other pathotypes (negative control). Scale bars are 10 μm.

PCR-based pathotype detection leads to underdiagnosis of EAEC and, consequently, to underestimation of their etiological role in diarrhea. This is evident from the fact that the EAEC pathotype was identified in only one of the "TOP 5" serogroups targeted by the Czech surveillance system (Ileninová et al., 2022), and, on the other hand, many other EAEC serogroups remain undetected. To improve this situation, implementation of a serotype-independent surveillance based on PCR detection of pathotype-specific marker genes already at the level of routine microbiological laboratories is needed. Such a surveillance concept would make it possible to obtain an objective information about the role of EAEC in the etiology of diarrhea in the Czech Republic which is currently lacking.

The sequenced Czech EAEC O111:H21/ST40 strains resembled in their virulence characteristics and cgMLST such strains from other countries in Europe and beyond (Ellis et al., 2020, Llorente et al., 2023, Tré-Hardy et al., 2023), indicating that they belong to the EAEC O111:H21/ST40 clone which causes diarrhea in children worldwide. Similarly, EAEC O111:H12/ST10 strains that accounted for a minority of Czech EAEC O111 isolates and differed from EAEC O111:H21/H40 strains by the virulence gene profiles (Table 3) and core genomes (Fig. 4)

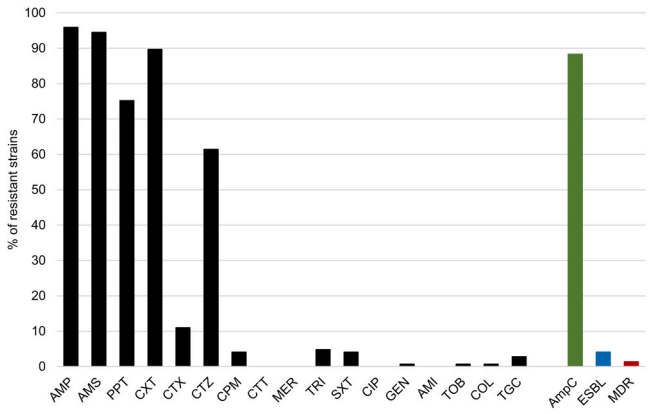


**Fig. 8.** Resistance of EAEC O111 strains to antimicrobials. AMP, ampicillin; AMS, ampicillin/sulbactam; PPT, piperacillin/tazobactam; CXT, cefoxitin; CTX, cefotaxime; CTZ, ceftazidime; CPM, cefepime; CTT, ceftolozane/tazobactam; MER, meropenem; TRI, trimethoprim; SXT, trimethoprim/sulfamethoxazole; CIP, ciprofloxacin; GEN, gentamicin; AMI, amikacin; TOB, tobramycin; COL, colistin. ESBL, extended spectrum β-lactamase; AmpC, AmpC β-lactamase; MDR, multidrug resistant.

**Table 4**
Comparison of antimicrobial resistance phenotypes and genotypes in EAEC O111 strains subjected to whole genome sequencing.

| Strain no. | Resistance phenotype[a] | ESBL[b] | AmpC[b] | Genetic background of resistance (WGS) |
|---|---|---|---|---|
| 19–815 | AMP, AMS, PPT, CXT, TRI, SXT | - | + | $bla_{TEM-1\ C}$, *ampC* promoter mutation (-42 C>T), *dfrA14*, *sul2*, *aph(3")-lb*, *aph(6)-ld* |
| 19–806 | AMP, AMS, PPT, CXT | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–683 | AMP, AMS, PPT, CXT, CTZ | - | + | $bla_{TEM-1\ C}$, *ampC* promoter mutation (-42 C>T), *sul2*, *aph (3")-lb*, *aph(6)-ld* |
| 20–019 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–325 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–422 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–615 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–725 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 19–698 | AMP, AMS, PPT, CXT, CTX, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 20–243 | AMP, AMS, PPT, CXT, CTZ | - | + | *ampC* promoter mutation (-42 C>T) |
| 21–582 | AMP, AMS, PPT, CXT | - | + | *ampC* promoter mutation (-42 C>T) |
| 18–510 | AMP, CXT, CTX, CTZ, CPM | + | - | $bla_{CTX-M-15}$ |
| 19–457 | None | - | - | None |

[a] AMP, ampicillin; AMS, ampicillin/sulbactam; PPT, piperacillin/tazobactam; CXT, cefoxitin; CTX, cefotaxime; CTZ, ceftazidime; CPM, cefepime; TRI, trimethoprim; SXT, trimethoprim/sulfamethoxazole.

[b] ESBL, extended spectrum β-lactamase; AmpC, AmpC β-lactamase; +, the phenotype is present; -, the phenotype is absent.

were also occasionally identified in other countries (Fig. 5). This indicates that the global EAEC O111 population consists of a major O111:H21/ST40 and a minor O111:H12/ST10 groups.

Our observation that AAF/V-encoding cluster which mediates aggregative adherence (Jønsson et al., 2015) is in EAEC O111:H21/ST40 strains located on pAA plasmids is in accordance with reports for other types of AAF fimbriae in other EAEC serotypes (Boisen et al., 2008; Mellmann et al., 2011; Nataro et al., 1994; Rasko et al., 2011). Localization of AAF/V cluster on pAA plasmid was previously reported in an EAEC/EHEC O111:H21 hybrid using Illumina sequencing (Dallman et al., 2012). The hybrid assembly of long (Nanopore) and short (Illumina) reads enabled us to reconstruct, to the best of our knowledge for the first time, complete circularized sequences of $pAA_{O111}$ plasmids and to show that though they may differ by size, they always carried, besides the AAF/V cluster, also other EAEC virulence genes including *aggR*, *aap*, *aar*, *aat* cluster, and *sepA*. Interestingly, EAEC O111:H12/ST10 strains that lacked all presently known types of AAF fimbriae (Table 3) also displayed the typical stacked-brick adherence to human intestinal epithelial cells HCT-8 (Fig. 7F, G). We therefore speculated that the aggregate-forming pili (AFP) that have been shown to confer aggregative adherence to AAF/I-AAF/V-negative EAEC strains (Lang et al., 2018) might be responsible for aggregative adherence of EAEC O111:H12/ST10 strains. However, we did not find any homologies to AFP-encoding genes in Illumina sequences of our EAEC O111:H12/ST10 strains suggesting yet another mechanism of their aggregative adherence. Putative involvement of curli fimbriae which are encoded in both EAEC O111:H12/ST10 strains (Table 3) and have been reported to confer biofilm formation and cell-cell aggregation in *E. coli* and other *Enterobacteriaceae* (Collinson et al., 1992; Hu, 2018) might be considered. The molecular basis of aggregative adherence of AAF-negative and AFP-negative EAEC strains thus needs further investigation.

In addition to their etiological role in diarrhea, another public health importance of EAEC is due to their ability to integrate $stx_2$-harboring bacteriophages into their genomes (Bielaszewska et al., 2011; Dallman et al., 2012; Iyoda et al., 2000; Mellmann et al., 2011). The resulting hybrids combine Stx2 production with aggregative adherence to intestinal epithelial cells (Bielaszewska et al., 2011) and have thus a high virulence potential including the ability to cause HUS and outbreaks. Specifically, EAEC/EHEC O111 hybrid caused a community outbreak of HUS in France (Boudailliez et al., 1997; Morabito et al., 1998) and a household outbreak of HUS and diarrhea in Northern Ireland (Dallman et al., 2012). EAEC/EHEC O104:H4 hybrid caused a devastating outbreak of HUS in Germany (Kampmeier et al., 2018; Karch et al., 2012). The cgMLST clustering of EAEC/EHEC O111:H21/ST40 hybrids with EAEC O111:H21/ST40 strains, and their core genomic distance from EHEC O111 (Fig. 6) indicate that EAEC/EHEC O111 hybrids evolved from EAEC O111 through lysogenic conversion with $stx_2$-harboring bacteriophages rather than from EHEC O111 by acquisition of EAEC virulence features.

The high level of antibiotic resistance in EAEC O111 in our study is in agreement with the resistance of EAEC strains including those of O111 serogroup from other industrialized countries, but the resistance patterns somewhat differ. In most studies, EAEC resistance to ampicillin was, like in our study, high (Do Nascimento et al., 2017; Hebbelstrup Jensen et al., 2018; Nüesch-Inderbinen et al., 2013; Tobias et al., 2015; Tokuda et al., 2010). On the other hand, the high rates of resistance to cephalosporins and of the AmpC production observed in our study (88.3 %) were infrequently reported in other countries (Do Nascimento et al., 2017; Hebbelstrup Jensen et al., 2016). The frequency of the ESBL phenotype among our EAEC isolates (4.1 %) was similar to that reported in EAEC from Spain (2.5 %) (Llorente et al., 2023), but substantially lower than those reported in Switzerland (20 %) (Nüesch-Inderbinen et al., 2013), the United Kingdom (34 %) (Do Nascimento et al., 2017), and Japan (19 %) (Imuta et al., 2016). Also, the rates of resistance to trimethoprim and trimethoprim/sulfamethoxazole in our study (4.8 % and 4.1 %, respectively) were considerably lower than those in the studies from the United Kingdom (48.4 % and 49.7 %, respectively) (Do Nascimento et al., 2017), Denmark (58 % and 53 %, respectively (Hebbelstrup Jensen et al., 2018), Switzerland (80 % and 70 %, respectively) (Nüesch-Inderbinen et al., 2013), and Israel (33.5 %) (Tobias et al., 2015). The same was true for the rates of MDR in EAEC strains in our study (1.4 %) compared to those in other European countries and the United States (23 %-60 %) (Beczkiewicz et al., 2019; Do Nascimento et al., 2017; Hebbelstrup Jensen et al., 2016;

Hebbelstrup Jensen et al., 2018; Llorente et al., 2023; Nüesch-Inderbinen et al., 2013). Another encouraging finding in our study is the susceptibility of all 145 EAEC O111 strains to meropenem and ciprofloxacin. EAEC susceptibility to ciprofloxacin was also reported in Israel (Tobias et al., 2015) and in Danish children with EAEC diarrhea (Hebbelstrup Jensen et al., 2016), whereas 34 % of EAEC isolates from adults in Denmark were ciprofloxacin resistant (Hebbelstrup Jensen et al., 2018). On the other hand, a worrying finding in our study is the resistance of an EAEC O111 strain to colistin, which is the last resort antibiotic to treat infections caused by highly resistant Gram-negative bacteria.

In conclusion, our study demonstrates that the EAEC pathotype is frequent among *E. coli* O111 isolates from children with diarrhea and other gastrointestinal disorders in the Czech Republic. EAEC strains may thus be important, but presently underdiagnosed causes of pediatric diarrhea in this country. Serotype-independent, PCR-based pathotype surveillance needs to be implemented to assess the role of EAEC as intestinal pathogens.

## Ethical approval

The study was approved by the Ethics Committee of the National Institute of Public Health, Prague (No. SZU/01661/2019 from April 17, 2018). Written informed consent was not required from the patients' parents because the stool analyses were performed as part of routine microbiological examination.

## Funding information

This study was supported by the Ministry of Health of the Czech Republic - conceptual development of research organization ('The National Institute of Public Health—NIPH, 75010330'). The whole genome sequencing was performed as part of the project „Genomic surveillance of selected infectious diseases in the Czech Republic" (Grant agreement no. 101113387 - HERA2CZ) co-funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor HADEA can be held responsible for them. The funders had no role in the study design, in the collection, analysis and interpretation of data, in the writing of the article, and in the decision to submit the article for publication.

## CRediT authorship contribution statement

**Antje Flieger:** Writing – review & editing, Supervision. **Helena Žemličková:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization. **Martina Bielaszewska:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Monika Havlíčková:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Jiří Dresler:** Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – review & editing. **Klára Schlosserová:** Writing – review & editing, Writing – original draft, Investigation, Data curation. **Klára Labská:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Ondřej Daniel:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Vladislav Jakubů:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. **Tereza Stárková:** Writing – review & editing, Investigation. **Jan Bílý:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Christina Lang:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Data curation. **Angelika Fruth:** Writing – review & editing, Investigation.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Illumina raw reads of Czech EAEC O111 strains, and complete sequences of plasmids from EAEC strains were deposited to the European Nucleotide Archive (ENA) (study accession PRJEB72449).

## Acknowledgments

The computational resources for WGS analysis were provided through the e-INFRA CZ project (ID: 90254) supported by the Ministry of Education, Youth and Sports of the Czech Republic. We are grateful to Denis Piérard and Florence Crombé (Universitair Ziekenhuis Brussel, Belgium) for providing us with raw reads of the Belgian EAEC O111 strains and Ulrich Dobrindt (Institute for Hygiene, University of Muenster, Germany) for the EAEC referencestrain 042. We thank Júlia Kseničová and Kristián Kozák (NRL for *E. coli* and Shigella, NIPH, Prague), Iveta Vrbová) (NRL for Antibiotics, NIPH, Prague), and Thomas Garn and Tobias Grössl (Robert Koch Institute, Wernigerode) for skilled technical assistance.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ijmm.2024.151628.

## References

Alcock, B.P., Huynh, W., Chalil, R., Smith, K.W., Raphenya, A.R., Wlodarski, M.A., Edalatmand, A., Petkau, A., Syed, S.A., Tsang, K.K., Baker, S.J.C., Dave, M., McCarthy, M.C., Mukiri, K.M., Nasir, J.A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., Kwong, M., Liang, Z.C., Niu, K.C., Shan, P., Yang, J.Y.J., Gray, K.L., Hoad, G.R., Jia, B., Bhando, T., Carfrae, L.A., Farha, M.A., French, S., Gordzevich, R., Rachwalski, K., Tu, M.M., Bordeleau, E., Dooley, D., Griffiths, E., Zubyk, H.L., Brown, E.D., Maguire, F., Beiko, R.G., Hsiao, W.W.L., Brinkman, F.S.L., Van Domselaar, G., McArthur, A.G., 2023. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 51 (D1), D690–D699. https://doi.org/10.1093/nar/gkac920.

Banjo, M., Iguchi, A., Seto, K., Kikuchi, T., Harada, T., Scheutz, F., Iyoda, S., Pathogenic, E., coli Working Group in Japan, 2018. *Escherichia coli* H-genotyping PCR: a complete and practical platform for molecular H typing. J. Clin. Microbiol. 56, e00190-18 https://doi.org/10.1128/JCM.00190-18.

Beczkiewicz, A., Cebelinski, E., Decuir, M., Lappi, V., Wang, X., Smith, K., Boxrud, D., Medus, C., 2019. High relative frequency of enteroaggregative *Escherichia coli* among patients with reportable enteric pathogens, Minnesota, 2016-2017. Clin. Infect. Dis. 69, 473–479. https://doi.org/10.1093/cid/ciy890.

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. https://doi.org/10.1016/S1473-3099(11)70165-7.

Bogaerts, P., Rodriguez-Villalobos, H., Bauraing, C., Deplano, A., Laurent, C., Berhin, C., Struelens, M.J., Glupczynski, Y., 2010. Molecular characterization of AmpC-producing *Escherichia coli* clinical isolates recovered at two Belgian hospitals. Pathol. Biol. (Paris) 58, 78–83. https://doi.org/10.1016/j.patbio.2009.07.035.

Boisen, N., Struve, C., Scheutz, F., Krogfelt, K.A., Nataro, J.P., 2008. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. Infect. Immun. 76, 3281–3292. https://doi.org/10.1128/IAI.01646-07.

Boisen, N., Ruiz-Perez, F., Scheutz, F., Krogfelt, K.A., Nataro, J.P., 2009. Short report: high prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative *Escherichia coli*. Am. J. Trop. Med. Hyg. 80, 294–301. https://doi.org/10.4269/ajtmh.2009.80.294.

Boisen, N., Scheutz, F., Rasko, D.A., Redman, J.C., Persson, S., Simon, J., Kotloff, K.L., Levine, M.M., Sow, S., Tamboura, B., Toure, A., Malle, D., Panchalingam, S., Krogfelt, K.A., Nataro, J.P., 2012. Genomic characterization of enteroaggregative *Escherichia coli* from children in Mali. J. Infect. Dis. 205, 431–444. https://doi.org/10.1093/infdis/jir757.

Boisen, N., Østerlund T., M., Joensen G., K., Santiago, A.E., Mandomando, I., Cravioto, A., Chattaway, M.A., Gonyar, L.A., Overballe-Petersen, S., Colin Stine, O., Rasko, D.A., Scheutz, F., Nataro, J.P., 2020. Redefining enteroaggregative *Escherichia coli* (EAEC): genomic characterization of epidemiological EAEC strains. PLoS Negl. Trop. Dis. 14, e0008613 https://doi.org/10.1371/journal.pntd.0008613.

Boudailliez, B., Berquin, P., Mariani-Kurkdjian, P., Ilef, D., Cuvelier, B., Capek, I., Tribout, B., Bingen, E., Piussan, C., 1997. Possible person-to-person transmission of *Escherichia coli* O111-associated hemolytic uremic syndrome. Pediatr. Nephrol. 11, 36–39. https://doi.org/10.1007/s004670050229.

Brown, C.L., Mullet, J., Hindi, F., Stoll, J.E., Gupta, S., Choi, M., Keenum, I., Vikesland, P., Pruden, A., Zhang, L., 2022. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. Appl. Environ. Microbiol. 88, e0099122 https://doi.org/10.1128/aem.00991-22.

Campos, L.C., Whittam, T.S., Gomes, T.A., Andrade, J.R., Trabulsi, L.R., 1994. *Escherichia coli* serogroup O111 includes several clones of diarrheagenic strains with different virulence properties. Infect. Immun. 62, 3282–3288. https://doi.org/10.1128/iai.62.8.3282-3288.1994.

CDC: Laboratory standard operating procedure for Pulsenet Nextera XT library prep and run setup for the Illumina MiSeq. ⟨https://www.cdc.gov/pulsenet/pdf/pnl32-miseq-nextera-xt.pdf⟩. November 10, 2020.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Clausen, P.T.L.C., Aarestrup, F.M., Lund, O., 2018. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinforma. 19, 307. https://doi.org/10.1186/s12859-018-2336-6.

Cohen, M.B., Nataro, J.P., Bernstein, D.I., Hawkins, J., Roberts, N., Staat, M., 2005. Prevalence of diarrheagenic *Escherichia coli* in acute childhood enteritis: a prospective controlled study. J. Pediatr. 146, 54–61. https://doi.org/10.1016/j.jpeds.2004.08.059.

Collinson, S.K., Emödy, L., Trust, T.J., Kay, W.W., 1992. Thin aggregative fimbriae from diarrheagenic *Escherichia coli*. J. Bacteriol. 174, 4490–4495. https://doi.org/10.1128/jb.174.13.4490-4495.1992.

Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Wlodarska, M., Finlay, B.B., 2013. Recent advances in understanding enteric pathogenic *Escherichia coli*. Clin. Microbiol. Rev. 26, 822–880. https://doi.org/10.1128/CMR.00022-13.

Dallman, T., Smith, G.P., O'Brien, B., Chattaway, M.A., Finlay, D., Grant, K.A., Jenkins, C., 2012. Characterization of a verocytotoxin-producing enteroaggregative *Escherichia coli* serogroup O111:H21 strain associated with a household outbreak in Northern Ireland. J. Clin. Microbiol. 50, 4116–4119. https://doi.org/10.1128/JCM.02047-12.

De Coster, W., Rademakers, R., 2023. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics 39, btad311. https://doi.org/10.1093/bioinformatics/btad311.

Decree No. 275/2010 Coll, amending Decree No. 473/2008 Coll., on the system of epidemiological vigilance for selected infections. Collection of Laws of the Czech Republic, 2010.

De Rauw, K., Jacobs, S., Piérard, D., 2018. Twenty-seven years of screening for Shiga toxin-producing *Escherichia coli* in a university hospital. Brussels, Belgium, 1987-2014. PLoS One 13, e0199968. https://doi.org/10.1371/journal.pone.0199968.

Do Nascimento, V., Day, M.R., Doumith, M., Hopkins, K.L., Woodford, N., Godbole, G., Jenkins, C., 2017. Comparison of phenotypic and WGS-derived antimicrobial resistance profiles of enteroaggregative *Escherichia coli* isolated from cases of diarrhoeal disease in England, 2015-16. J. Antimicrob. Chemother. 72, 3288–3297. https://doi.org/10.1093/jac/dkx301.

Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P., Nataro, J.P., 2006. Proteomic and microarray characterization of the AggR regulon identifies a *pheU* pathogenicity island in enteroaggregative *Escherichia coli*. Mol. Microbiol. 61, 1267–1282. https://doi.org/10.1111/j.1365-2958.2006.05281.x.

Ellis, S.J., Crossman, L.C., McGrath, C.J., Chattaway, M.A., Hölken, J.M., Brett, B., Bundy, L., Kay, G.L., Wain, J., Schüller, S., 2020. Identification and characterisation of enteroaggregative *Escherichia coli* subtypes associated with human disease. Sci. Rep. 10, 7475. https://doi.org/10.1038/s41598-020-64424-3.

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J.G., Haendiges, J., Haft, D.H., Hoffmann, M., Pettengill, J.B., Prasad, A.B., Tillman, G.E., Tyson, G.H., Klimke, W., 2021. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci. Rep. 11, 12728 https://doi.org/10.1038/s41598-021-91456-0.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.Y., Graham, M., Van Domselaar, G., Stothard, P., 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 51 (W1), W484–W492. https://doi.org/10.1093/nar/gkad326.

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C.A., Frimodt-Møller, N., Aarestrup, F.M., 2014. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. J. Clin. Microbiol. 52, 139–146. https://doi.org/10.1128/JCM.02452-13.

Hebbelstrup Jensen, B., Olsen, K.E., Struve, C., Krogfelt, K.A., Petersen, A.M., 2014. Epidemiology and clinical manifestations of enteroaggregative *Escherichia coli*. Clin. Microbiol. Rev. 27, 614–630. https://doi.org/10.1128/CMR.00112-13.

Hebbelstrup Jensen, B., Stensvold, C.R., Struve, C., Olsen, K.E.P., Scheutz, F., Boisen, N., Röser, D., Andreassen, B.U., Nielsen, H.V., Schønning, K., Petersen, A.M., Krogfelt, K. A., 2016. Enteroaggregative *Escherichia coli* in daycare - a 1-year dynamic cohort study. Front. Cell. Infect. Microbiol. 6, 75. https://doi.org/10.3389/fcimb.2016.00075.

Hebbelstrup Jensen, B., Poulsen, A., Hebbelstrup Rye Rasmussen, S., Struve, C., Engberg, J.H., Friis-Møller, A., Boisen, N., Jønsson, R., Petersen, R.F., Petersen, A.M., Krogfelt, K.A., 2017. Genetic virulence profile of enteroaggregative *Escherichia coli* strains isolated from Danish children with either acute or persistent diarrhea. Front. Cell. Infect. Microbiol. 7, 230. https://doi.org/10.3389/fcimb.2017.00230.

Hebbelstrup Jensen, B., Adler Sørensen, C., Hebbelstrup Rye Rasmussen, S., Rejkjær Holm, D., Friis-Møller, A., Engberg, J., Mirsepasi-Lauridsen, H.C., Struve, C., Hammerum, A.M., Porsbo, L.J., Petersen, R.F., Petersen, A.M., Krogfelt, K.A., 2018. Characterization of diarrheagenic enteroaggregative *Escherichia coli* in Danish adults - antibiotic treatment does not reduce duration of diarrhea. Front. Cell. Infect. Microbiol. 8, 306. https://doi.org/10.3389/fcimb.2018.0030.

Hrabák, J., 2007. Clinically important beta-lactamases of gram-negative bacteria: AmpC. Epidemiol. Mikrobiol. Imunol. 56, 155–165.

Hu, L., 2018. Prevalence of curli genes among Cronobacter species and their roles in biofilm formation and cell-cell aggregation. Int. J. Food Microbiol. 265, 65–73. https://doi.org/10.1016/j.ijfoodmicro.2017.10.031.

Huppertz, H.I., Rutkowski, S., Aleksic, S., Karch, H., 1997. Acute and chronic diarrhoea and abdominal colic associated with enteroaggregative *Escherichia coli* in young children living in western Europe. Lancet 349, 1660–1662. https://doi.org/10.1016/S0140-6736(96)12485-5.

Ileninová, Z., Klimešová, P., Schlosserová, K., Kotiš, J., Kseničová, J., Daniel, O., Bielaszewska, M., Havlíčková-Marejková, M., 2022. Laboratory diagnostics of Shiga toxin-producing *E. coli* in the National Reference Laboratory for *E. coli* and Shigella and methodological guideliness for clinical laboratories. Zprávy CEM (SZÚ, Praha) 31, 15–22.

Imuta, N., Ooka, T., Seto, K., Kawahara, R., Koriyama, T., Kojyo, T., Iguchi, A., Tokuda, K., Kawamura, H., Yoshiie, K., Ogura, Y., Hayashi, T., Nishia, J., 2016. Phylogenetic analysis of enteroaggregative *Escherichia coli* (EAEC) isolates from Japan reveals emergence of CTX-M-14-producing EAEC O25:H4 clones related to sequence type 131. J. Clin. Microbiol. 54, 2128–2134. https://doi.org/10.1128/JCM.00711-16.

Iyoda, S., Tamura, K., Itoh, K., Izumiya, H., Ueno, N., Nagata, K., Togo, M., Terajima, J., Watanabe, H., 2000. Inducible *stx*2 phages are lysogenized in the enteroaggregative and other phenotypic *Escherichia coli* O86:HNM isolated from patients. FEMS Microbiol. Lett. 191, 7–10.

Jenkins, C., 2018. Enteroaggregative *Escherichia coli*. Curr. Top. Microbiol. Immunol. 416, 427–450.. DOI 10.1007/82_2018_105.

Joffré, E., Iñiguez Rojas, V., 2020. Molecular epidemiology of enteroaggregative *Escherichia coli* (EAEC) isolates of hospitalized children from Bolivia reveal high heterogeneity and multidrug-resistance. Int. J. Mol. Sci. 21, 9543. https://doi.org/10.3390/ijms21249543.

Jønsson, R., Struve, C., Boisen, N., Mateiu, R.V., Santiago, A.E., Jenssen, H., Nataro, J.P., Krogfelt, K.A., 2015. Novel aggregative adherence fimbria variant of enteroaggregative *Escherichia coli*. Infect. Immun. 83, 1396–1405. https://doi.org/10.1128/IAI.02820-14.

Kampmeier, S., Berger, M., Mellmann, A., Karch, H., Berger, P., 2018. The 2011 German enterohemorrhagic *Escherichia coli* O104:H4 outbreak - the danger is still out there. Curr. Top. Microbiol. Immunol. 416, 117–148. https://doi.org/10.1007/82_2018_107.

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

Karch, H., Denamur, E., Dobrindt, U., Finlay, B.B., Hengge, R., Johannes, L., Ron, E.Z., Tønjum, T., Sansonetti, P.J., Vicente, M., 2012. The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. EMBO Mol. Med. 4, 841–848. https://doi.org/10.1002/emmm.201201662.

Kislichkina, A.A., Kartsev, N.N., Skryabin, Y.P., Sizova, A.A., Kanashenko, M.E., Teymurazov, M.G., Kuzina, E.S., Bogun, A.G., Fursova, N.K., Svetoch, E.A., Dyatlov, I.A., 2022. Genomic analysis of a hybrid enteroaggregative hemorrhagic *Escherichia coli* O181:H4 strain causing colitis with hemolytic-uremic syndrome. Antibiot. (Basel) 11, 1416. https://doi.org/10.3390/antibiotics11101416.

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A., 2019. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37, 540–546. https://doi.org/10.1038/s41587-019-0072-8.

Lang, C., Fruth, A., Holland, G., Laue, M., Mühlen, S., Dersch, P., Flieger, A., 2018. Novel type of pilus associated with a Shiga-toxigenic *E. coli* hybrid pathovar conveys aggregative adherence and bacterial virulence. Emerg. Microbes Infect. 7, 2003. https://doi.org/10.1038/s41426-018-0209-8.

Larsen, M.V., Cosentino, S., Lukjancenko, O., Saputra, D., Rasmussen, S., Hasman, H., Sicheritz-Pontén, T., Aarestrup, F.M., Ussery, D.W., Lund, O., 2014. Benchmarking of methods for genomic taxonomy. J. Clin. Microbiol. 52, 1529–1539. https://doi.org/10.1128/JCM.02981-13.

Li, Y., Cao, B., Liu, B., Liu, D., Gao, Q., Quili, G., Peng, X., Wu, J., Bastin, D.A., Feng, L., Wang, L., 2009. Molecular detection of all 34 distinct O-antigen forms of Shigella. J. Med. Microbiol. 58, 69–81. https://doi.org/10.1099/jmm.0.000794-0.

Llorente, M.T., Escudero, R., Ramiro, R., Remacha, M.A., Martínez-Ruiz, R., Galán-Sánchez, F., de Frutos, M., Elía, M., Onrubia, I., Sánchez, S., 2023. Enteroaggregative *Escherichia coli* as etiological agent of endemic diarrhea in Spain: A prospective multicenter prevalence study with molecular characterization of isolates. Front. Microbiol. 14, 1120285 https://doi.org/10.3389/fmicb.2023.1120285.

Marejková, M., Bláhová, K., Janda, J., Fruth, A., Petráš, P., 2013. Enterohemorrhagic *Escherichia coli* as causes of hemolytic uremic syndrome in the Czech Republic. PLoS One 8, e73927. https://doi.org/10.1371/journal.pone.0073927.

Mellmann, A., Bielaszewska, M., Köck, R., Friedrich, A.W., Fruth, A., Middendorf, B., Harmsen, D., Schmidt, M.A., Karch, H., 2008. Analysis of collection of hemolytic

uremic syndrome-associated enterohemorrhagic *Escherichia coli*. Emerg. Infect. Dis. 14, 1287–1290. https://doi.org/10.3201/eid1408.071082.

Mellmann, A., Harmsen, D., Cummings, C.A., Zentz, E.B., Leopold, S.R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y., Zhang, W., McLaughlin, S.F., Henkhaus, J.K., Leopold, B., Bielaszewska, M., Prager, R., Brzoska, P.M., Moore, R.L., Guenther, S., Rothberg, J.M., 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. https://doi.org/10.1371/journal.pone.0022751.

Morabito, S., Karch, H., Mariani-Kurkdjian, P., Schmidt, H., Minelli, F., Bingen, E., Caprioli, A., 1998. Enteroaggregative, Shiga toxin-producing *Escherichia coli* O111:H2 associated with an outbreak of haemolytic uremic syndrome. J. Clin. Microbiol. 36, 840–842. https://doi.org/10.1128/JCM.36.3.840-842.1998.

Müller, D., Greune, L., Heusipp, G., Karch, H., Fruth, A., Tschäpe, H., Schmidt, M.A., 2007. Identification of unconventional intestinal pathogenic *Escherichia coli* isolates expressing intermediate virulence factor profiles by using a novel single-step multiplex PCR. Appl. Environ. Microbiol. 73, 3380–3390. https://doi.org/10.1128/AEM.02855-06.

Nataro, J.P., Kaper, J.B., 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. 11, 142–201. https://doi.org/10.1128/CMR.11.1.142.

Nataro, J.P., Kaper, J.B., Robins-Browne, R., Prado, V., Vial, P., Levine, M.M., 1987. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. Pediatr. Infect. Dis. J. 6, 829–831. https://doi.org/10.1097/00006454-198709000-00008.

Nataro, J.P., Yikang, D., Yingkang, D., Walker, K., 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. J. Bacteriol. 176, 4691–4699. https://doi.org/10.1128/jb.176.15.4691-4699.1994.

Nüesch-Inderbinen, M.T., Hofer, E., Hächler, H., Beutin, L., Stephan, R., 2013. Characteristics of enteroaggregative *Escherichia coli* isolated from healthy carriers and from patients with diarrhoea. J. Med. Microbiol. 62, 1828–1834. https://doi.org/10.1099/jmm.0.065177-0.

Oxford Nanopore Technologies, 2022. Ligation sequencing gDNA - Native Barcoding Kit 24 V14 (SQK-NBD114.24). ⟨https://community.nanoporetech.com/docs/prepare/library_prep_protocols/ligation-sequencing-gdna-native-barcoding-v14-sqk-nbd114-24/v/nbe_9169_v114_revq_15sep2022⟩ (accessed May 15, 2024).

Oxford Nanopore Technologies, 2017. Medaka. ⟨https://github.com/nanoporetech/medaka⟩ (accessed May 15, 2024).

Pabst, W.L., Altwegg, M., Kind, C., Mirjanic, S., Hardegger, D., Nadal, D., 2003. Prevalence of enteroaggregative *Escherichia coli* among children with and without diarrhea in Switzerland. J. Clin. Microbiol. 41, 2289–2293. https://doi.org/10.1128/JCM.41.6.2289-2293.2003.

Päivärinta, M., Pohjola, L., Fredriksson-Ahomaa, M., Heikinheimo, A., 2016. Low occurrence of extended-spectrum β-lactamase-producing *Escherichia coli* in Finnish food-producing animals. Zoonoses Public Health 63, 624–631. https://doi.org/10.1111/zph.12277.

Perelle, S., Dilasser, F., Grout, J., Fach, P., 2007. Screening food raw materials for the presence of the world's most frequent clinical cases of Shiga toxin-encoding *Escherichia coli* O26, O103, O111, O145 and O157. Int. J. Food Microbiol. 113, 284–288. https://doi.org/10.1016/j.ijfoodmicro.2006.08.014.

Pérez-Pérez, F.J., Hanson, N.D., 2002. Detection of plasmid-mediated AmpC beta-lactamase genes in clinical isolates by using multiplex PCR. J. Clin. Microbiol. 40, 2153–2162. https://doi.org/10.1128/JCM.40.6.2153-2162.2002.

Prager, R., Lang, C., Aurass, P., Fruth, A., Tietze, E., Flieger, A., 2014. Two novel EHEC/EAEC hybrid strains isolated from human infections. PLoS One 9, e95379. https://doi.org/10.1371/journal.pone.0095379.

Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.S., Iliopoulos, D., Klammer, A., Peluso, P., Lee, L., Kislyuk, A.O., Bullard, J., Kasarskis, A., Wang, S., Eid, J., Rank, D., Redman, J.C., Steyert, S.R., Frimodt-Møller, J., Struve, C., Petersen, A.M., Krogfelt, K.A., Nataro, J.P., Schadt, E. E., Waldor, M.K., 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med. 365, 709–717. https://doi.org/10.1056/NEJMoa1106920.

Rodríguez, E., Gaggero, C., Laviña, M., 1999. The structural gene for microcin H47 encodes a peptide precursor with antibiotic activity. Antimicrob. Agents Chemother. 43, 2176–2182. https://doi.org/10.1128/AAC.43.9.2176.

Scavia, G., Staffolani, M., Fisichella, S., Striano, G., Colletta, S., Ferri, G., Escher, M., Minelli, F., Caprioli, A., 2008. Enteroaggregative *Escherichia coli* associated with a foodborne outbreak of gastroenteritis. J. Med. Microbiol 57, 1141–1146. https://doi.org/10.1099/jmm.0.2008/001362-0.

Scheutz, F., Teel, L.D., Beutin, L., Piérard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R., Morabito, S., Strockbine, N.A., Melton-Celsa, A.R., Sanchez, M., Persson, S., O'Brien, A.D., 2012. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. J. Clin. Microbiol. 50, 2951–2963. https://doi.org/10.1128/JCM.00860-12.

Schmidt, H., Plaschke, B., Franke, S., Rüssmann, H., Schwarzkopf, A., Heesemann, J., Karch, H., 1994. Differentiation in virulence patterns of *Escherichia coli* possessing *eae* genes. Med. Microbiol. Immunol. Berl. 183, 23–31. https://doi.org/10.1007/BF00193628.

Schmidt, H., Knop, C., Franke, S., Aleksic, S., Heesemann, J., Karch, H., 1995. Development of PCR for screening of enteroaggregative *Escherichia coli*. J. Clin. Microbiol. 33, 701–705. https://doi.org/10.1128/jcm.33.3.701-705.1995.

Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., Goesmann, A., 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb. Genom. 7, 000685 https://doi.org/10.1099/mgen.0.000685.

Shaw, K.J., Rather, P.N., Hare, R.S., Miller, G.H., 1993. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. Microbiol. Rev. 57, 138–163. https://doi.org/10.1128/mr.57.1.138-163.1993.

Stothard, P., Wishart, D.S., 2005. Circular genome visualization and exploration using CGView. Bioinformatics 21, 537–539. https://doi.org/10.1093/bioinformatics/bti054.

EUCAST guidelines for detection of resistance mechanisms and specific resistances of clinical and/or epidemiological importance. European Committee on Antimicrobial Susceptibility Testing, 2017. ⟨http://www.eucast.org/resistance_mechanisms/⟩.

The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 11.0, 2021. ⟨http://www.eucast.org⟩.

Tobias, J., Kassem, E., Rubinstein, U., Bialik, A., Vutukuru, S.R., Navaro, A., Rokney, A., Valinsky, L., Ephros, M., Cohen, D., Muhsen, K., 2015. Involvement of main diarrheagenic *Escherichia coli*, with emphasis on enteroaggregative *E. coli*, in severe non-epidemic pediatric diarrhea in a high-income country. BMC Infect. Dis. 15, 79. https://doi.org/10.1186/s12879-015-0804-4.

Tokuda, K., Nishi, J., Imuta, N., Fujiyama, R., Kamenosono, A., Manago, K., Kawano, Y., 2010. Characterization of typical and atypical enteroaggregative *Escherichia coli* in Kagoshima, Japan: biofilm formation and acid resistance. Microbiol. Immunol. 54, 320–329. https://doi.org/10.1111/j.1348-0421.2010.00210.x.

Tré-Hardy, M., De Rauw, K., Blairon, L., Simon, A., Piérard, D., 2023. Nursery outbreak caused by enteroaggregative *Escherichia coli* serogroup O111:H21. APMIS 131, 206–216. https://doi.org/10.1111/apm.13301.

Valilis, E., Ramsey, A., Sidiq, S., DuPont, H.L., 2018. Non-O157 Shiga toxin-producing *Escherichia coli*-A poorly appreciated enteric pathogen: Systematic review. Int. J. Infect. Dis. 76, 82–87. https://doi.org/10.1016/j.ijid.2018.09.002.

Wick, R.R., 2017. Filtlong. ⟨https://github.com/rrwick/Filtlong⟩ (accessed 15 May, 2024).

Wick, R.R., Holt, K.E., 2022. Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLoS Comput. Biol. 18, e1009802 https://doi.org/10.1371/journal.pcbi.1009802.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves, P.R., Maiden, M.C.J., Ochman, H., Achtman, M., 2006. Sex and virulence in *Escherichia coli:* an evolutionary perspective. Mol. Microbiol. 60, 1136–1151.. DOI: 10.1111/j.1365-2958.2006.05172.x.

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 257. https://doi.org/10.1186/s13059-019-1891-0.

Yatsuyanagi, J., Saito, S., Sato, H., Miyajima, Y., Amano, K., Enomoto, K., 2002. Characterization of enteropathogenic and enteroaggregative *Escherichia coli* isolated from diarrheal outbreaks. J. Clin. Microbiol. 40, 294–297. https://doi.org/10.1128/JCM.40.1.294-297.2002.

Zhou, Z., Alikhan, N.F., Sergeant, M.J., Luhmann, N., Vaz, C., Francisco, A.P., Carriço, J. A., Achtman, M., 2018. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 28, 1395–1404. https://doi.org/10.1101/gr.232397.117.

Zhou, Z., Alikhan, N.F., Mohamed, K., Fan, Y., Agama Study Group, Achtman, M., 2020. The EnteroBase user's guide, with case studies on Salmonella transmissions, *Yersinia pestis* phylogeny, and Escherichia core genomic diversity. Genome Res 30, 138–152. https://doi.org/10.1101/gr.251678.119.