

RESEARCH ARTICLE

Open Access



Attack of the clones: whole genome-based characterization of two closely related enterohemorrhagic *Escherichia coli* O26 epidemic lineages

Lucia Karnisova¹, Monika Marejkova², Hana Hrbackova³, Alexander Mellmann⁴, Helge Karch⁴, Angelika Fruth⁵, Pavel Drevinek⁶, Kveta Blahova¹, Martina Bielaszewska² and Jaroslav Nunvar^{6*}

Abstract

Background: Enterohemorrhagic *Escherichia coli* (EHEC) O26:H11/H⁻, the most common non-O157 serotype causing hemolytic uremic syndrome worldwide, are evolutionarily highly dynamic with new pathogenic clones emerging rapidly. Here, we investigated the population structure of EHEC O26 isolated from patients in several European countries using whole genome sequencing, with emphasis on a detailed analysis of strains of the highly virulent new European clone (nEC) which has spread since 1990s.

Results: Genome-wide single nucleotide polymorphism (SNP)-based analysis of 32 EHEC O26 isolated in the Czech Republic, Germany, Austria and Italy demonstrated a split of the nEC (ST29C2 clonal group) into two distinct lineages, which we termed, based on their temporal emergence, as “early” nEC and “late” nEC. The evolutionary divergence of the early nEC and late nEC is marked by the presence of 59 and 70 lineage-specific SNPs (synapomorphic mutations) in the genomes of the respective lineages. In silico analyses of publicly available *E. coli* O26 genomic sequences identified the late nEC lineage worldwide. Using a PCR designed to target the late nEC synapomorphic mutation in the *sen/ent* gene, we identified the early nEC decline accompanied by the late nEC rise in Germany and the Czech Republic since 2004 and 2013, respectively. Most of the late nEC strains harbor one of two major types of Shiga toxin 2a (Stx_{2a})-encoding prophages. The type I *stx*_{2a}-phage is virtually identical to *stx*_{2a}-phage of EHEC O104:H4 outbreak strain, whereas the type II *stx*_{2a}-phage is a hybrid of EHEC O104:H4 and EHEC O157:H7 *stx*_{2a}-phages and carries a novel mutation in Stx_{2a}. Strains harboring these two phage types do not differ by the amounts and biological activities of Stx_{2a} produced.

Conclusions: Using SNP-level analyses, we provide the evidence of the evolutionary split of EHEC O26:H11/H⁻ nEC into two distinct lineages, and a recent replacement of the early nEC by the late nEC in Germany and the Czech Republic. PCR targeting the late nEC synapomorphic mutation in *ent/sen* enables the discrimination of early nEC strains and late nEC strains in clinical and environmental samples, thereby facilitating further investigations of their geographic distribution, prevalence, clinical significance and epidemiology.

Keywords: Shiga toxin, O26, Enterohemorrhagic *Escherichia coli* (EHEC), New European clone

* Correspondence: jaroslav.nunvar@lfmotol.cuni.cz

⁶Department of Medical Microbiology, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic
Full list of author information is available at the end of the article



Background

Enterohemorrhagic *Escherichia coli* (EHEC) O26:H11/H⁻ is the most common non-O157 EHEC serotype causing diarrhea and its systemic complication, the hemolytic uremic syndrome (HUS) worldwide [1–16]. In Europe, the proportion of EHEC O26 among all EHEC patients' isolates has been increasing during past years [1] and in some countries, EHEC O26 is a more frequent cause of HUS than EHEC O157:H7 [12, 16]. HUS caused by EHEC O26 can be as severe as that caused by EHEC O157:H7 [13, 17–20] and there is no significant difference in the long-term outcome of HUS caused by EHEC O26 and EHEC O157 [18]. In addition to sporadic cases, multiple outbreaks caused by EHEC O26 have been reported throughout the world [15, 21–30].

The majority of EHEC O26 clinical isolates carry genes encoding Shiga toxins (Stx) of two major types including Stx_{1a} and/or Stx_{2a} [9, 26, 27, 31–33]. However, a subset of strains contain genes encoding Stx subtypes such as Stx_{2d} [34, 35], Stx_{2b} [15], or Stx_{2f} [36]. The presence of *stx*_{2a} and/or production of Stx_{2a} alone is a predictor for a severe disease including progression of the infection to HUS [31]. The *stx*_{2a}-harboring EHEC O26 belong to two major multilocus sequence types (STs) ST21 or ST29, whereas strains harboring *stx*_{1a} alone or together with *stx*_{2a} belong to ST21 [9, 31]. The ST29 clade is considerably heterogeneous and consists of several clonal lineages which differ by *stx*₂ alleles and plasmid gene profiles [9, 31, 37, 38]. A new highly virulent ST29 clone harboring *stx*_{2a} only was first reported in Germany in the mid-1990s [33] and since then it has rapidly spread throughout Europe [31]. Strains of this lineage termed a “new German clone” or a “new European clone” [31, 38] exhibit a particular plasmid virulence gene profile characterized by the presence of EHEC-*hlyA* (encoding EHEC hemolysin) and *etpD* (encoding a type II effector), while *katP* (encoding catalase-peroxidase) and *espP* (encoding serine protease EspP) genes are absent [31, 33]. In 2015, a distinct ST29 clonal lineage harboring *stx*_{2a} or *stx*_{2d} and lacking the plasmid-borne virulence genes (EHEC-*hlyA*, *katP*, *espP*, *etpD*) was reported in France as a “new French clone” [35, 38]. Recently, yet another ST29 clone, which harbors *stx*_{2a} only and differs from both the European and French clone by plasmid virulence gene profile (EHEC-*hlyA*⁺, *katP*⁻, *espP*⁺, *etpD*⁻) emerged in Japan [9]. This Japanese clone (designated ST29C1 clade) is phylogenetically unrelated to the new European clone (ST29C2 clade) and the French clone (ST29C3 clade) [9].

Although EHEC O26 are the most common cause of pediatric HUS in the Czech Republic [12], the clonal structure of these strains has not been investigated in detail. The aim of this study was to analyze the phylogenetic relationships and clonal structure of EHEC O26

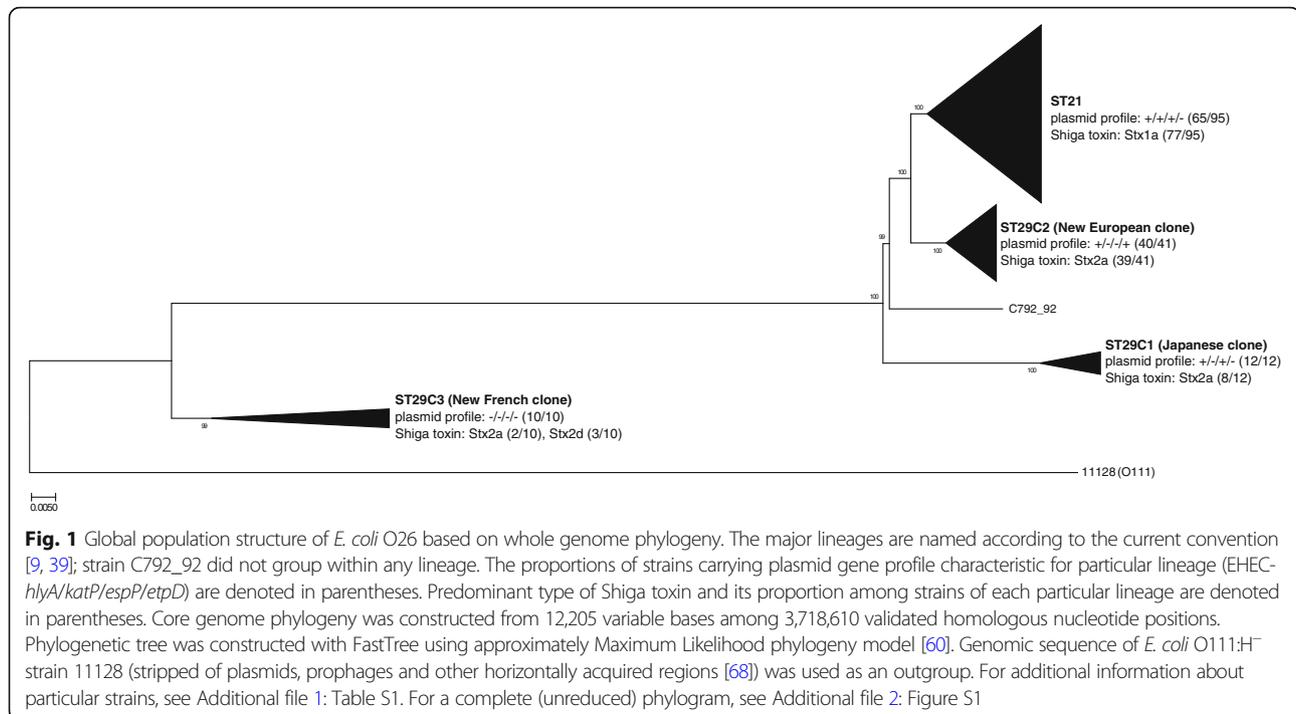
isolated from patients in the Czech Republic by using whole genome sequencing (WGS) with a particular focus on ST29 strains. To put these data into a global context, the genomic sequences of the Czech isolates were compared with those of EHEC O26 isolated in other European countries and with *E. coli* O26 genomes present in GenBank.

Results

Whole genome-based clonal structure of EHEC O26 population

Initially, we performed WGS of 16 EHEC O26 strains isolated from patients with HUS or diarrhea without HUS in the Czech Republic during the period of 2006–2016. The STs and plasmid gene profiles were determined in silico from draft genome sequences. This revealed the presence of both major O26 phylogenetic lineages, ST21 (6 isolates) and ST29 (10 isolates), among Czech isolates (Additional file 1: Table S1). Five of the six ST21 strains harbored *stx*_{2a} only, and one contained *stx*_{1a} only; all of them displayed a plasmid-borne virulence gene profile (EHEC-*hlyA*⁺, *katP*⁺, *espP*⁺, *etpD*⁻) previously identified in ST21 strains [31] (Additional file 1: Table S1). Nine of the ten ST29 EHEC O26 Czech strains corresponded to the new European clone (nEC) as defined in [31], i.e., they contained *stx*_{2a} as the sole *stx* gene and the plasmid virulence gene profile EHEC-*hlyA*⁺, *katP*⁻, *espP*⁻, and *etpD*⁺ (Additional file 1: Table S1). Remarkably, the remaining ST29 strain (15–496) harbored the plasmid gene profile typical for the nEC, but combined with the presence of *stx*_{1c} gene only instead of *stx*_{2a} (Additional file 1: Table S1). This is, to the best of our knowledge, the first report of *stx*_{1c} genotype in strains of the nEC. To gain a more detailed insight into the population structure of the nEC, we performed additional WGS of a collection of European nEC isolates including 16 strains originating from Germany (*n* = 11), Italy (*n* = 3), and Austria (*n* = 2). For comparison, genomic sequences of *E. coli* O26 available in the GenBank database were included in this analysis, making up the final set of total 159 *E. coli* O26 isolates whose genetic relationships were examined (Additional file 1: Table S1).

A genome-wide single nucleotide polymorphism (SNP)-based phylogram was constructed using the maximum likelihood model to provide high-resolution inference of *E. coli* O26 phylogeny (Fig. 1, Additional file 2: Figure S1). This phylogram clearly identified four major lineages, whose characteristics (i.e. ST, plasmid gene profile, and predominant Stx type) were in accordance with previous reports [9, 37–39]. Strikingly, the phylogenetic analysis demonstrated that strains of the nEC (syn. ST29C2 [9]) split into two distinct clusters (Fig. 2). This is in accordance with the observation of two distinct pulse-field gel electrophoresis (PFGE) clusters



among a large, representative collection of European ST29 strains isolated between 1996 and 2012 in the original description of the nEC [31]. Retrospectively, we noted tentative differences between the two ST29 PFGE clusters with respect to the distribution of isolation dates of the corresponding strains. First strains of PFGE cluster B were isolated as early as in 1996 and their majority (53.6%) until 2004. In contrast, strains of the PFGE cluster C started to emerge, with a single exception, since 2004 [31]. We therefore propose to name the WGS lineages homologous to the PFGE clusters B and C as “early” and “late” nEC, respectively (Fig. 2, Additional file 3: Table S3). This designation is further supported by a detailed analysis of a temporal shift between the nEC lineages, as shown below.

Evolution, diversification and spread of the nEC

To infer which genetic events underlied the evolutionary establishment of the nEC and its subsequent split into the “early” and “late” lineages, we sought mutations characteristic for particular lineages, i.e., SNPs absent in other *E. coli* O26 phylogroups including ST21 and ST29 non-nEC strains (synapomorphic mutations; for details, see Methods). The sets of synapomorphic mutations detected by this approach (Additional file 4: Table S2) provide an unambiguous genetic definition of the particular nEC lineages. The total numbers of SNPs synapomorphic for early nEC, late nEC and nEC as a whole were 59, 70, and 83, respectively (Fig. 2). This justifies that the early nEC and late nEC should be considered distinct phylogenetic entities. Next, we focused on

parallel evolution of missense synapomorphic mutations (which are likely to influence function of encoded proteins) within the nEC. Three cases where a single gene underwent multiple amino acid changes during various phases of nEC evolution were detected: cation transporter CusA and hypothetical proteins YehI and YggM accumulated two mutations each. Four operons accumulated two missense mutations each: *glgBXCAP* (glycogen metabolism), *wza-wzb-wzc* (capsular polysaccharide synthesis), *hisJQMP* (histidine transport) and *gcl-hyi-glx-R-ybbW-allB-ybbY-glxK* (purine metabolism) (Additional file 4: Table S2). Notably, three mutations in fimbrial biogenesis proteins (ElfG, YqiG, and HtrE) accounted for 12% of 25 missense mutations synapomorphic for the late nEC (Additional file 4: Table S2). Since fimbriae are determinants of bacterial adherence, their modification might affect persistence of late nEC strains in mammal hosts or even pathogenesis in infected humans.

To rapidly differentiate between early nEC and late nEC strains, we designed a simple SNP-based PCR, which discriminates a mutation in the *sen/ent* gene (encoding *Shigella flexneri* enterotoxin 2 [40]), which is synapomorphic for the late nEC (Additional file 4: Table S2) (see Methods). This PCR was applied to screen all nEC isolates originating from Germany and the Czech Republic (45 and 15 isolates, respectively; Additional file 3: Table S3). Despite the limited numbers of isolates, a trend of the early nEC decline with a concomitant increase in the late nEC proportion was observed in both countries (Fig. 3). This suggests that the two lineages

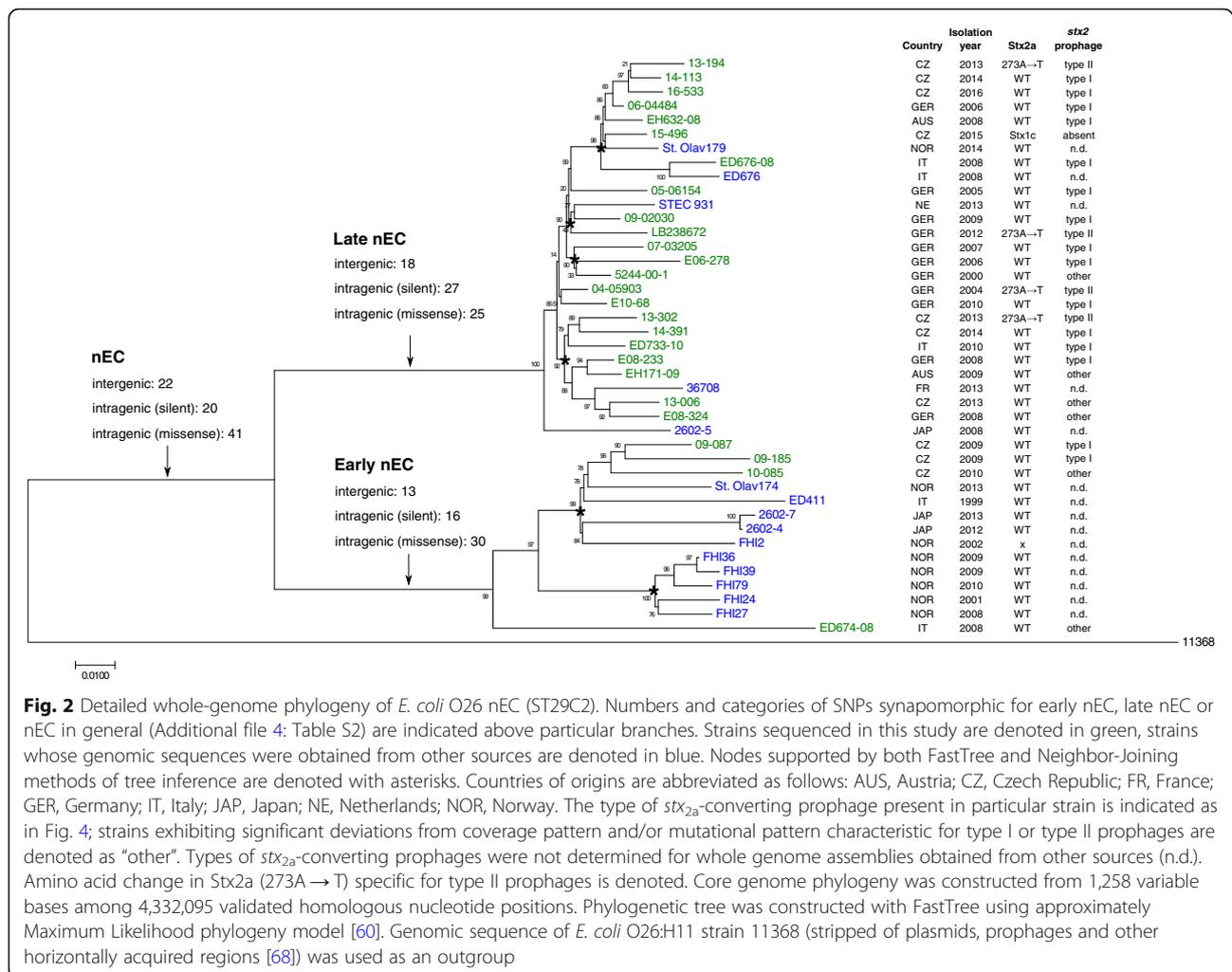


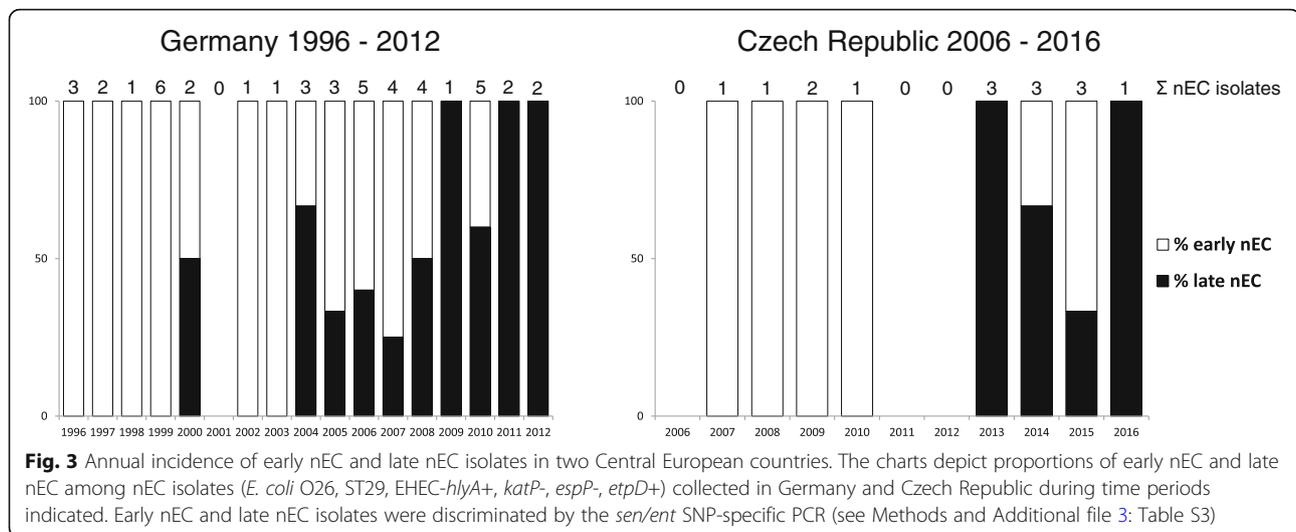
Fig. 2 Detailed whole-genome phylogeny of *E. coli* O26 nEC (ST29C2). Numbers and categories of SNPs synapomorphic for early nEC, late nEC or nEC in general (Additional file 4: Table S2) are indicated above particular branches. Strains sequenced in this study are denoted in green, strains whose genomic sequences were obtained from other sources are denoted in blue. Nodes supported by both FastTree and Neighbor-Joining methods of tree inference are denoted with asterisks. Countries of origins are abbreviated as follows: AUS, Austria; CZ, Czech Republic; FR, France; GER, Germany; IT, Italy; JAP, Japan; NE, Netherlands; NOR, Norway. The type of *stx*_{2a}-converting prophage present in particular strain is indicated as in Fig. 4; strains exhibiting significant deviations from coverage pattern and/or mutational pattern characteristic for type I or type II prophages are denoted as “other”. Types of *stx*_{2a}-converting prophages were not determined for whole genome assemblies obtained from other sources (n.d.). Amino acid change in Stx2a (273A → T) specific for type II prophages is denoted. Core genome phylogeny was constructed from 1,258 variable bases among 4,332,095 validated homologous nucleotide positions. Phylogenetic tree was constructed with FastTree using approximately Maximum Likelihood phylogeny model [60]. Genomic sequence of *E. coli* O26:H11 strain 11368 (stripped of plasmids, prophages and other horizontally acquired regions [68]) was used as an outgroup

might occupy the same niche and compete with each other. The emergence of the late nEC in Germany (regularly isolated since 2004) preceded its first occurrence in the Czech Republic (2013) by nine years (Fig. 3). Strains belonging to the late nEC were also isolated in several other European countries including Austria (2008), Italy (2008, 2010), France (2013), Netherlands (2013) and Norway (2014), as well as in Japan (2008) (Fig. 2), which documents the pan-European and inter-continental distribution of this lineage.

nEC strains harbor different *stx*_{2a}-converting prophages

Shiga toxins encoded by *stx* genes are responsible for systemic complications of EHEC infections such as HUS. The *stx* genes are located in the genomes of lambdoid prophages (*stx*-converting phages) which are very heterogeneous (reviewed in [41]). A recent report [38] demonstrated that *stx*_{2a}-converting prophages of EHEC O26 nEC strains were virtually identical with the *stx*_{2a}-prophages present in the highly virulent *E. coli* O104:H4 German

outbreak strain [42], while other O26 lineages harbored different *stx*-converting prophages [38]. To gain a deeper characterization of these elements in the early nEC and late nEC strains, we analyzed *stx*-converting prophages in all nEC genomes sequenced in this study, utilizing short sequencing reads and complete prophage sequences as references (Fig. 4). All *stx*_{2a}-converting prophages in nEC strains were integrated into the *wrbA* gene which is also the site of phage integration in *E. coli* O104:H4 [43] and EHEC O157:H7 prototype strains Sakai [44] and EDL933 [45]. In agreement with Delannoy et al. [38], our results demonstrate that *stx*_{2a}-converting prophages virtually identical with *E. coli* O104:H4 prophage (termed here “type I”) are predominant among nEC isolates. Unexpectedly, another distinct type of *stx*_{2a}-converting prophage (termed here “type II”) was identified in four late nEC strains (Fig. 2). While the left half (5′-portion) of the type II prophage sequence was identical with O104:H4 prophage, the right half (3′-portion) diverged significantly (Fig. 4). The consensus sequences obtained by mapping of



sequencing reads to the O104:H4 prophage reference were extracted from the divergent regions and searched against the BLAST database of complete *E. coli* genomes. The most closely related sequences were retrieved for *stx_{2a}-*converting prophages from EHEC O157:H7 [46]; the high level of homology was confirmed by re-mapping of the sequencing reads to the O157:H7 prophage reference (Fig. 4). The presence of two DNA segments highly homologous to either O104:H4 or O157:H7 *stx_{2a}-*converting prophages suggests a role of recombination in the evolution of type II prophages. This is in agreement with the view of recombination as the most common genetic event contributing to the mosaic structure and thus to the high genetic diversity of EHEC *stx*-converting phages [38, 41].

Remarkably, the type II prophages carry a single point mutation in the left half of their sequence, which is otherwise identical among type I, type II and O104 *stx*-converting prophages (Fig. 4). This mutation, which changes alanine (A) 273 into threonine (T) 273 in Stx2a A subunit (Fig. 2), is novel among the known diversity of Stx2 proteins and complements other mutations of A273 to amino acids with hydroxylated side chains (serine [S] and tyrosine [Y]- Additional file 5: Figure S2).

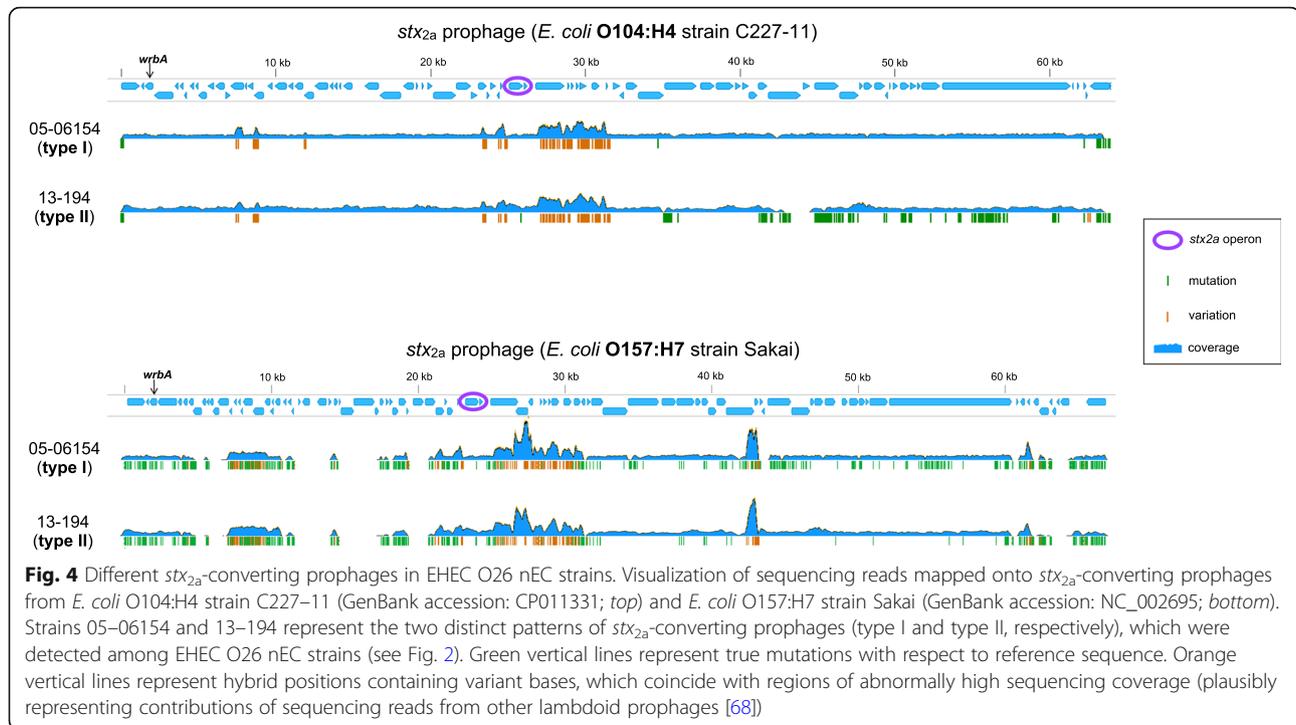
Stx production and cytotoxicity of nEC strains

To determine whether or not the A273 → T273 mutation in Stx2a A subunit encoded by the type II *stx_{2a}-*prophages influenced the toxicity of the resulting Stx2a protein, we compared the amounts, Vero cell cytotoxicity titers, and specific activities (CD₅₀/ng toxin) of Stx2a produced by nEC strains harboring the type II prophages with those of strains harboring the type I or other *stx_{2a}-*converting prophages. No significant differences were found (Fig. 5), indicating that the A273 → T273 mutation in the type II *stx_{2a}-*prophage had no effect on Stx2a biological activity. Moreover, comparison of Stx2a amounts, cytotoxicity

titers, and Stx2a specific activities between strains of the late nEC and early nEC did not reveal any differences in these characteristics within the ST29 nEC group (Fig. 5). However, all these characteristics of ST29 nEC strains were significantly higher than those of Stx2a-producing strains of ST21 clonal group (Fig. 5).

Discussion

Using genome-wide SNP-based analysis, we demonstrate that EHEC O26 strains belonging to the nEC are not phylogenetically homogeneous, but consist of two distinct lineages. This finding extends the present knowledge about the clonal structure of these pathogens and confirms previous reports [9, 31, 35, 39] that *E. coli* O26:H11/H⁻ are, from the evolutionary point of view, highly dynamic with a potential for novel virulent clones to emerge rapidly. Based on their consecutive appearance and spread within Germany and the Czech Republic, we propose to term these lineages “early” nEC and “late” nEC, respectively. nEC was reported to have arisen in Germany during the 1990s [31, 33]; currently, both nEC lineages are widely distributed in Europe and were also isolated in Japan (Fig. 2), which highlights that nEC strains are capable of rapid spreading across countries and continents. Given the genetic distinctness of the early and late nEC, their expansions are likely to have occurred independently. Although the spread via routes common to other EHEC (i.e. via livestock and/or food trade) seems to be most likely, *E. coli* O26 corresponding to the nEC have only rarely been isolated from cattle and other animals [47, 48]. Moreover, none of 10 *E. coli* O26 nEC isolates (ST29C2 clade) present among 520 whole genome sequences analyzed by Ogura et al. [39], which represent a global population of human and animal *E. coli* O26, originated from cattle or other animals. Thus, reservoirs and possible ways of the worldwide spread of *E. coli* O26 nEC in general and of the

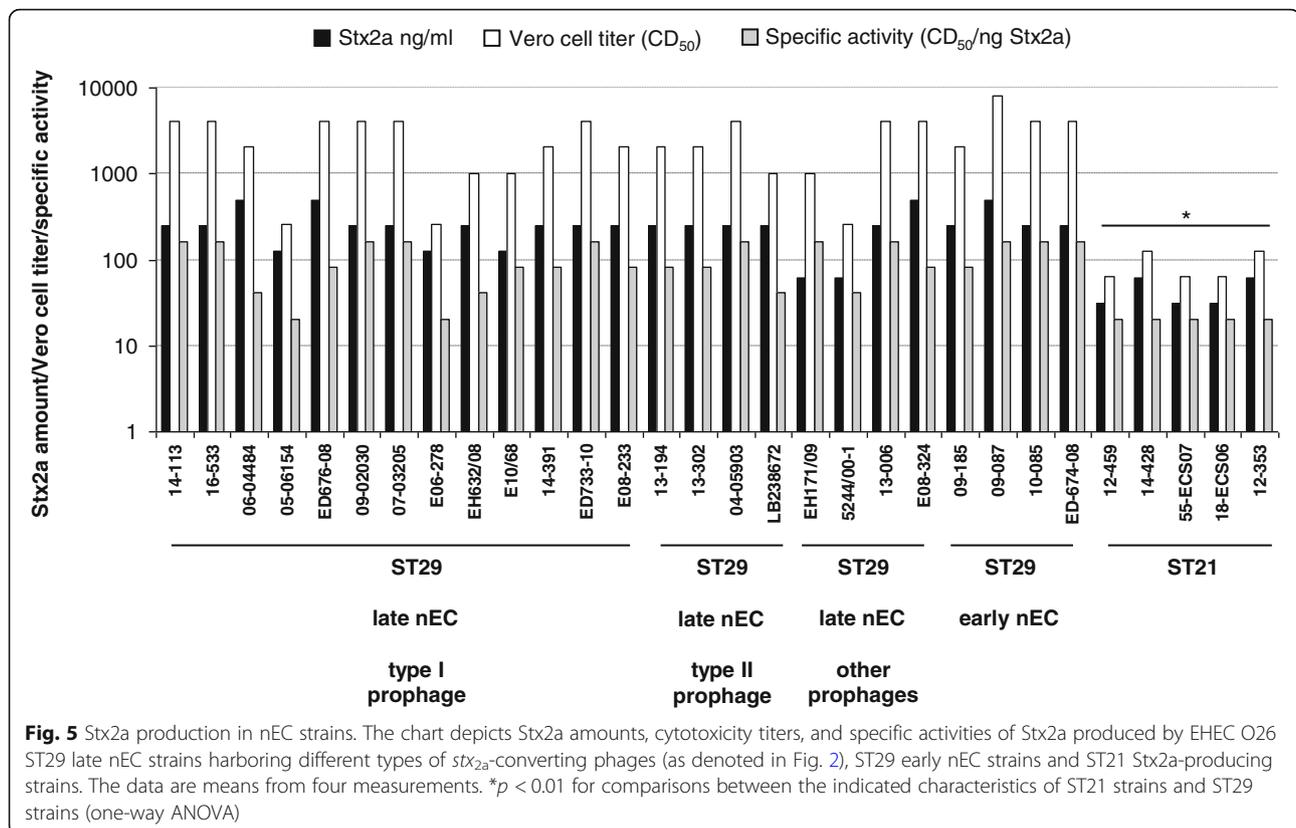


late nEC in particular need to be further elucidated. This is particularly important because of the increasing frequency of the late nEC strains as causes of human diseases in some European countries during last years (Fig. 3). To create a basis for effective measures to control further spread of these pathogens, screening of ST29 *E. coli* O26 isolates from cattle and other animals, using the *sen/ent* SNP-based PCR developed in this study, combined with plasmid gene profiling might be a useful and simple tool for identification of strains of the early and late nEC. Altogether, the emergence of the late nEC both throughout and outside Europe supports a continuous evolution of *E. coli* O26:H11/H⁻ and a high propensity of new clones to spread [9, 35, 39].

Analysis of *stx*_{2a}-converting prophages in nEC strains (Fig. 4) identified a predominant prophage (designated type I), which displayed a nearly complete sequence identity with the *stx*_{2a}-converting prophages of EHEC O104:H4 strain, which caused the large devastating outbreak of HUS in Germany in 2011 [43]. Type I prophage is widespread across the nEC diversity (Fig. 2), implicating a single acquisition during early evolution of nEC (as also suggested in [38]). Based on available bibliographic data, the emergence of EHEC O26 nEC in the 1990s [31, 33] predates the isolation of Stx2a-producing EHEC O104:H4 [32, 43, 49], which was first isolated in 2001 [32]. This allows to speculate that EHEC O26 belonging to the nEC might have served as donors of *stx*_{2a}-converting phages for enteroaggregative *E. coli* O104:H4 in the final step of

evolution of the highly virulent, Stx2a-producing EHEC O104:H4 outbreak strain. Alternatively, EHEC O26 nEC and EHEC O104:H4 outbreak strain might have acquired their *stx*_{2a}-converting phages independently from a common source. Beutin et al. [50] proposed that *stx*_{2a}-converting prophages of EHEC O104:H4 might have originated from *stx*_{2a}-converting phages present among bovine *E. coli* of various serotypes, with which they share nucleotide sequences of two specific DNA fragments, virion morphology, DNA restriction patterns, chromosomal integration site, and superinfection immunity [50]; however, this hypothesis needs to be confirmed by complete sequence analyses of the bovine phages.

Four late nEC strains were found to harbor a *stx*_{2a}-converting prophage whose sequence consisted of two regions highly homologous to EHEC O104:H4 and EHEC O157:H7 *stx*_{2a}-phages, respectively (designated type II; Fig. 4). These strains were scattered throughout the whole-genome phylogenetic tree and their most closely related strains harbored type I prophages (Fig. 2). This provides evidence that horizontal, rather than vertical pattern of type II prophages transfer was prevalent in late nEC evolution, thus suggesting an ongoing dissemination of type II phages in the *E. coli* O26 population. Notably, Stx2a encoded by type II prophages contain a novel amino acid change (A273 → T273). This mutation does not affect the potency of the toxin as demonstrated by its specific activity for Vero cells, which is similar to that of strains harboring type I prophages (Fig. 5).



However, the localization of the mutation suggests a possible connection with Stx2a maturation. A273 follows immediately after arginine [R] 272; cleavage of Stx by the Golgi protease furin takes place precisely between these two residues [51, 52]. The cleavage generates the enzymatically active form of Stx2a A subunit [53]. The furin site where A273 binding takes place is a small hydrophilic pocket [54], which could explain the propensity of A273 for substitutions with hydroxyl-containing amino acids (Additional file 5: Figure S2). If there are any biological effects conferred by this Stx2a mutation specific for type II prophages, remains to be elucidated.

Conclusions

Using genome-wide SNP-based analysis, this study presents the evidence of the split of the EHEC O26:H11/H⁻nEC, which emerged in Germany in the 1990s and has spread throughout Europe [31], into two cryptic, yet distinct clones (early nEC and late nEC, Fig. 2). For EHEC O26, the combination of MLST, plasmid gene profiling and PCR-based *stx* subtyping has been regarded as sufficient for clone discrimination [9, 31]. The fact that strains of the early nEC and late nEC are indistinguishable by this approach (Fig. 1) emphasizes the necessity of using highly discriminatory methods such as the genome-wide SNP-based analysis to distinguish the most

closely related epidemic lineages. For practical purposes, the SNP-based PCR developed in this study which targets a mutation in the *sen/ent* gene synapomorphic for the late nEC, represents a rapid and easy tool for distinguishing the early and late nEC strains in clinical microbiological laboratories and field studies. This will, in turn, enable further investigations of the geographic distribution of these pathogens, their clinical significance, and the epidemiology of human infections they cause.

Our results confirm that EHEC O26 nEC strains and *E. coli* O104:H4 German outbreak strain [43] share the same type of *stx*_{2a}-converting prophage [38]. An additional *stx*_{2a}-converting prophage (type II), present in a subset of the late nEC strains, displays several peculiarities (hybrid sequence, novel missense mutation in *stx*_{2a}, multiple horizontal transmissions among late nEC) which warrant further investigation.

In summary, the discriminatory power of whole genome sequencing allows for accurate reconstruction of evolutionary events which accompanied the emergence of novel clonal lineages of EHEC O26.

Methods

Database mining

Escherichia coli complete or draft genomic sequences were retrieved from the GenBank database [55] on

January 31st, 2017. The precompiled genome neighbors of *E. coli* strain St. Olav179 (a member of late nEC, accession JZED00000000 [56]) were extracted. Strains with symmetric identity below 92.0% were discarded since they typically represented serotypes other than O26. Among remaining strains, those belonging to non-O26 serotypes were discarded. In addition, genomic sequences of strains included in later studies which focused on genomic evolution of *E. coli* O26 [9, 38] were included in the dataset. In total, 127 strains from public sources were collected for genomic comparisons (Additional file 1: Table S1).

Whole-genome sequencing

Sixteen EHEC O26:H11/H⁻ strains were isolated from patients with HUS ($n = 10$) or with bloody ($n = 4$) or non-bloody ($n = 2$) diarrhea without HUS in the Czech Republic between 2006 and 2016. Additional 16 EHEC O26:H11/H⁻ nEC strains sequenced in this study originated from Germany ($n = 11$), Italy ($n = 3$), and Austria ($n = 2$) and were isolated from patients with HUS ($n = 10$) or non-bloody diarrhea ($n = 6$) (Additional file 1: Table S1). The Czech and German strains were from the authors' laboratories. The Italian and Austrian isolates were kindly provided by Stefano Morabito (European Union Reference Laboratory for *E. coli*, Istituto Superiore di Sanità, Rome, Italy) and Dorothea Orth-Höller (Division of Hygiene and Medical Microbiology, Innsbruck Medical University, Austria), respectively. WGS was performed as described previously [57]. Briefly, after inoculation of a single colony into nutrient broth (Heipha, Eppelheim, Germany) and overnight incubation (37 °C), genomic DNA was extracted using a MagAttract HMW DNA kit (Qiagen, Hilden, Germany) with the addition of lysozyme (Sigma-Aldrich, Taufkirchen, Germany) following the manufacturer's instructions. Subsequently, the sequencing libraries were prepared from the genomic DNA using the Nextera sample preparation kit (Illumina, Inc., San Diego, CA, USA) for a 150-bp or 250-bp paired-end sequencing run on a single MiSeq instrument (Illumina) in accordance with the manufacturer's recommendations. Libraries were scaled to reach 100-fold sequencing coverage. Subsequent quality trimming and de novo assembly were performed using the default parameters of CLC Genomics Workbench software (CLC bio, Arhus, Denmark) and the CLC Genomics Workbench de novo assembler (CLC bio). The draft genomes of strains sequenced in our study were annotated using the NCBI prokaryotic genome annotation pipeline [58] and are available in the GenBank under the accession numbers listed in Additional file 1: Table S1.

Bioinformatic analysis

Phylogeny reconstruction

Genome-wide SNP-based phylogeny was chosen for robust inference of *E. coli* O26 evolutionary history. Genomic sequences were uploaded to the CSIPhylogeny v1.4 website (<https://cge.cbs.dtu.dk/services/CSIPhylogeny/>) and automatically processed with default settings. SNP analysis was carried out using a set of algorithms as described in [59]; FastTree [60] was used for phylogram construction. Phylogenetic trees were visualized using MEGA7 [61]. Genomic sequences of *E. coli* O111:H⁻ strain 11128 and *E. coli* O26:H11 strain 11368 were included as reference genomes for EHEC O26 and nEC SNP-based phylogeny, respectively.

In silico analysis of molecular diagnostic markers

The presence of pathogenicity-related genes (plasmid genes EHEC-*hlyA*, *katP*, *espP* and *etpD*, Shiga toxin genes *stx*_{1a} and *stx*_{2a}) was assessed using TBLASTN [62] integrated in Geneious R9.1 [63], using protein query sequences from *E. coli* O26 strain 11368 (GenBank accession: AP010953) or *E. coli* O157 strain Sakai (GenBank accession: NC_002695). Presence or absence and identity of encoded proteins with respect to query sequences are listed in Additional file 1: Table S1.

Synapomorphic SNPs determination

Genomic sequences of several strains representing the genomic diversity of *E. coli* O26 (early nEC: St. Olav174, FHI24, FHI27; late nEC: ED676, St. Olav179, 36708, STEC931; ST21: ED180, ED729, STEC1117, CVM9942, non-nEC ST29: C792_92, CFSAN025102, 200C-3689, 34827) were aligned with Progressive MAUVE [64]. Presumptive synapomorphic SNPs (i.e. SNPs specific for early nEC, late nEC or nEC in general) were manually identified among the SNPs reported by MAUVE intergenomic comparison tool, based on their exclusive occurrence in corresponding lineages. For the definite validation, BLAST search [65] was performed against all genomic sequences included in the study (Additional file 1: Table S1); SNPs were considered synapomorphic if they were present in all members of a particular lineage and absent in all remaining O26 strains (Additional file 4: Table S2).

Stx-converting phage analysis

The paired-end sequencing reads were mapped onto *stx*_{2a}-converting prophage sequences present in *E. coli* O104:H4 strain C227-11 (GenBank accession: CP011331) and *E. coli* O157:H7 strain Sakai (GenBank accession: NC_002695) using Geneious R9 platform [63]. The mapping setting was as follows: max. 20% gaps per read, max. 10% mismatches per read. SNPs (true mutations) were called with coverage ≥ 5 and frequency $\geq 90\%$ (with respect

to the reference). Variable sites were called where a mixture of bases was present in assembly of sequencing reads (coverage ≥ 5 , frequency of minor variant $\geq 30\%$).

Late nEC-discriminating PCR

A SNP-based PCR was used as a quick screening method to detect a missense mutation in the *sen/ent* gene (encoding *Shigella flexneri* enterotoxin 2), specific for late nEC (102S \rightarrow Y; Additional file 4: Table S2). The following primers were designed using Primer/BLAST [66]: forward primer sen-F (5'-TCTAAATGGAAAGGTTAGTGATTG C-3') and reverse primer sen-RG (5'-GGTTATATA TAACGCTTCCCAAG-3'). In sen-RG, the 3'-terminal guanine is not complementary to the late nEC-specific SNP. The predicted amplicon length was 233 bp. PCR reactions were performed in CFX96 Touch PCR Detection System (Bio-Rad) using reagents from Top-Bio (Czech Republic), primers obtained from Generi Biotech (Czech Republic) and bacterial lysates (boiling method) as DNA templates. PCR conditions were: 94 °C for 5 min, followed by 30 cycles of denaturation (94 °C for 30 s), annealing (53 °C for 60 s), and extension (72 °C for 60 s) and a final extension step at 72 °C for 5 min. One early nEC and one late nEC strain were used as controls in each PCR. After electrophoretic separation on agarose gel and ethidium bromide staining, early nEC strains produced strong amplicons, while late nEC strains displayed substantially weaker amplicons due to a reduced pairing of 3' guanine with template DNA containing the SNP (Additional file 3: Table S3). All nEC strains analyzed by WGS in this study were tested with the late nEC-discriminating PCR; 100% correlation between the *sen/ent*-specific PCR and WGS results was achieved (Additional file 3: Table S3).

Detection of Stx production and cytotoxicity

VTEC RPLA

VTEC RPLA (Verotoxin-producing *E. coli* reverse passive latex agglutination) assay (Denka Seiken Company, Ltd., Tokyo, Japan) was performed, according to the manufacturer's instructions, with serially diluted sterile supernatants prepared from overnight shaken broth cultures by centrifugation (15 min, 4000 rpm, 4 °C) and subsequent filtration through 0.22 μ m syringe filters (P-LAB, Prague, Czech Republic). Stx1a and Stx2a titers were expressed as reciprocals of the supernatants' dilutions that produced a clear agglutination of latex particles sensitized with anti-Stx1a and anti-Stx2a antibody, respectively. Stx1a and Stx2a in the supernatants were quantified by comparison of titers produced by the supernatants with those of Stx2a (VT2) and Stx1a (VT1) standards with known toxin concentrations.

Vero cell assay

Vero-B4 cells (ACC-33; German collection of microorganisms and cell cultures, Braunschweig, Germany) were

maintained in Dulbecco's modified Eagle medium (DMEM) with 4.5 g/l of glucose and glutamine (Lonza, Cologne, Germany) supplemented with 10% of fetal bovine serum (Sigma-Aldrich, Prague, Czech Republic). In the cytotoxicity assay [67], semiconfluent Vero cell monolayers grown in 96-well plates were incubated with two-fold dilutions of supernatants prepared as above for 72 h. After removing medium with detached cells, remnant adherent cells were fixed (2% formalin), stained (0.13% crystal violet), washed, and after eluting of crystal violet with 50% ethanol, absorbance (OD₅₇₀) was measured by FLUOstar Omega (BMG Labtech, Ortenberg, Germany). Cytotoxicity titers were expressed as reciprocal values of the sample dilutions that killed 50% of cells (CD₅₀). Specific activities of Stx in supernatants were expressed as CD₅₀/ng of Stx.

Statistical analysis

Data on Stx2a amounts, cytotoxicity titers, and specific activities of Stx2a produced by different groups of strains were analyzed with one-way analysis of variance (ANOVA). *p* values < 0.05 were considered significant.

Additional files

Additional file 1: Table S1. Summary information about 159 *E. coli* O26 isolates whose genomic sequences were included for comparative analysis. Presence/absence of molecular diagnostic markers and their identity with query sequences are indicated. (DOCX 144 kb)

Additional file 2: Figure S1. Core genome-derived phylogeny of 159 *E. coli* O26 isolates. The phylogram (inferred using the CSIPhylogeny pipeline [59]) represents the full version of contracted phylogram presented in Fig. 1. (PDF 32 kb)

Additional file 3: Table S3. Discrimination of early nEC and late nEC strains by *sen/ent* SNP-specific PCR, WGS and PFGE. All nEC strains characterized in this study by WGS (Fig. 2) and/or by *sen/ent* SNP-specific PCR (Fig. 3) are listed, together with results obtained by either method and PFGE clustering [31] (n.a., not available). Representative gel image which demonstrates the performance of the *sen/ent* SNP-specific PCR in differentiating early- and late nEC strains is included. (XLSX 69 kb)

Additional file 4: Table S2. SNPs synapomorphic for the early nEC, late nEC or nEC as whole. Missense mutations in fimbrial proteins are denoted in bold. (XLS 112 kb)

Additional file 5: Figure S2. Alignment of all unique Stx2 protein sequences present in the GenBank. Residue A273 is marked with a red rectangle on the top of alignment. (PDF 6180 kb)

Acknowledgements

We thank Dr. Stephano Morabito (European Union Reference Laboratory for *E. coli*, Istituto Superiore di Sanità, Rome, Italy) and Dr. Dorothea Orth-Höller (Division of Hygiene and Medical Microbiology, Innsbruck Medical University, Austria) for providing us with EHEC O26 strains isolated in Italy and Austria, respectively. We are indebted to Drs. Nozomi Ishijima and Sunao Iyoda (Department of Bacteriology, National Institute of Infectious Diseases, Tokyo, Japan) for providing us with the draft genome assemblies of Japanese EHEC O26 strains. We are grateful to Dr. Michal Malina (Department of Paediatrics, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague) for his advice and help during the preparation of this project and to Dr. Ondrej Soucek (Laboratory of Vera Vavrova/MAL, Department of Paediatrics, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague) for help with spectrophotometry. A skilful technical assistance of

Veronika Horova (NIPH Prague), Ralph Fischer and Andrea Lagemann (Institute for Hygiene, Münster) is greatly appreciated.

Funding

This study was supported by the Grant Agency of the Charles University (194215) and by the Ministry of Health of the Czech Republic (funding project Conceptual development of research organization "The National Institute of Public Health - NIPH 75010330"). PD and JN acknowledge financial support from Czech Health Research Council of Ministry of Health of the Czech Republic, grant no. 15-28017A. AM acknowledges funding from the German Research Foundation (DFG), grant no. ME3205/2-1. The funders had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

All draft genome sequences were submitted to the GenBank database of National Center for Biotechnology Information (Bioproject accession number: PRJNA470577).

Authors' contributions

JN, LK, MB, PD and KB designed the study. LK and AM performed whole-genome sequencing. JN and LK performed bioinformatic analysis. LK, HH, MM and MB performed laboratory experiments and analyzed the data. MM, HK, AF, AM and MB collected the strains. PD and KB supervised the project. JN, LK, MM and MB drafted the manuscript. All authors have critically revised the manuscript and approved it for submission and publication.

Ethics approval and consent to participate

The use of the Czech isolates was approved by the Ethical Committee of the University Hospital Motol, Prague, Czech Republic. Written informed consent for enrolment in the study was obtained from parents of all patients. The isolates from Germany and other countries available in the collection of the Institute of Hygiene, University of Münster, were used in accordance with guidelines approved by the Ethical Committee of the Medical Faculty of the University of Münster and of the Ärztekammer Westfalen-Lippe. The informed consent of the participants was not required because the data were analyzed anonymously.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Paediatrics, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic. ²National Reference Laboratory for *E. coli* and Shigella, National Institute of Public Health, Prague, Czech Republic. ³Laboratory for Tissue Cultures, National Institute of Public Health, Prague, Czech Republic. ⁴Institute for Hygiene and the National Consulting Laboratory on Hemolytic Uremic Syndrome, University of Münster, Münster, Germany. ⁵National Reference Center for Salmonella and Other Enteric Pathogens, Robert Koch Institute, Wernigerode, Germany. ⁶Department of Medical Microbiology, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic.

Received: 22 June 2018 Accepted: 27 August 2018

Published online: 31 August 2018

References

- The European Union summary report on trends and sources of zoonoses. Zoonotic agents and food-borne outbreaks in 2016. *EFSA J.* 2017;15(12):e05077.
- Byrne L, Jenkins C, Launders N, Elson R, Adak GK. The epidemiology, microbiology and clinical impact of Shiga toxin-producing *Escherichia coli* in England, 2009-2012. *Epidemiol Infect.* 2015;143(16):3475-87.
- Caprioli A, Scavia G, Morabito S. Public health microbiology of Shiga toxin-producing *Escherichia coli*. *Microbiol Spectr.* 2014;2(6)
- Elliott EJ, Robins-Browne RM, O'Loughlin EV, Bennett-Wood V, Bourke J, Henning P, Hogg GG, Knight J, Powell H, Redmond D, et al. Nationwide study of haemolytic uraemic syndrome: clinical, microbiological, and epidemiological features. *Arch Dis Child.* 2001;85(2):125-31.
- Espié E, Grimont F, Mariani-Kurkdjian P, Bouvet P, Haeghebaert S, Filliol I, Loirat C, Decludt B, Minh NN, Vaillant V, et al. Surveillance of hemolytic uremic syndrome in children less than 15 years of age, a system to monitor O157 and non-O157 Shiga toxin-producing *Escherichia coli* infections in France, 1996-2006. *Pediatr Infect Dis J.* 2008;27(7):595-601.
- Gerber A, Karch H, Allerberger F, Verwey HM, Zimmerhackl LB. Clinical course and the role of Shiga toxin-producing *Escherichia coli* infection in the hemolytic-uremic syndrome in pediatric patients, 1997-2000, in Germany and Austria: a prospective study. *J Infect Dis.* 2002;186(4):493-500.
- Gould LH, Mody RK, Ong KL, Clogher P, Cronquist AB, Garman KN, Lathrop S, Medus C, Spina NL, Webb TH, et al. Increased recognition of non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States during 2000-2010: epidemiologic features and comparison with *E coli* O157 infections. *Foodborne Pathog Dis.* 2013;10(5):453-60.
- Hedican EB, Medus C, Besser JM, Juni BA, Koziol B, Taylor C, Smith KE. Characteristics of O157 versus non-O157 Shiga toxin-producing *Escherichia coli* infections in Minnesota, 2000-2006. *Clin Infect Dis.* 2009;49(3):358-64.
- Ishijima N, Lee KI, Kuwahara T, Nakayama-Imahiji H, Yoneda S, Iguchi A, Ogura Y, Hayashi T, Ohnishi M, Iyoda S. Identification of a new virulent clade in enterohemorrhagic *Escherichia coli* O26:H11/H- sequence type 29. *Sci Rep.* 2017;7:43136.
- Käppeli U, Hächler H, Giezendanner N, Beutin L, Stephan R. Human infections with non-O157 Shiga toxin-producing *Escherichia coli*, Switzerland, 2000-2009. *Emerg Infect Dis.* 2011;17(2):180-5.
- Karch H, Tarr PI, Bielaszewska M. Enterohaemorrhagic *Escherichia coli* in human medicine. *Int J Med Microbiol.* 2005;295(6-7):405-18.
- Marejková M, Bláhová K, Janda J, Fruth A, Petráš P. Enterohemorrhagic *Escherichia coli* as causes of hemolytic uremic syndrome in the Czech Republic. *PLoS One.* 2013;8(9):e73927.
- Pollock KG, Bhojani S, Beattie TJ, Allison L, Hanson M, Locking ME, Cowden JM. Highly virulent *Escherichia coli* O26 Scotland. *Emerg Infect Dis.* 2011; 17(9):1777-9.
- Rivas M, Miliwebsky E, Chinen I, Roldán CD, Balbi L, García B, Fiorilli G, Sosa-Estani S, Kincaid J, Rangel J, et al. Characterization and epidemiologic subtyping of Shiga toxin-producing *Escherichia coli* strains isolated from hemolytic uremic syndrome and diarrhea cases in Argentina. *Foodborne Pathog Dis.* 2006;3(1):88-96.
- Smith AM, Bosco KJ, Nicol MP, Kleynhans J, McCulloch M, Duze ST, Ismail A, Allam M, Tau NP, Keddy KH. Genome sequence for Shiga toxin-producing *Escherichia coli* O26:H11, associated with a cluster of hemolytic-uremic syndrome cases in South Africa, 2017. *Genome Announc.* 2017;5(38):e00989-17.
- Tozzi AE, Caprioli A, Minelli F, Gianviti A, De Petris L, Edefonti A, Montini G, Ferretti A, De Palo T, Gaido M, et al. Shiga toxin-producing *Escherichia coli* infections associated with hemolytic uremic syndrome, Italy, 1988-2000. *Emerg Infect Dis.* 2003;9(1):106-8.
- Ricotti GC, Buonomini MI, Merlitti A, Karch H, Luzzi I, Caprioli A. A fatal case of hemorrhagic colitis, thrombocytopenia, and renal failure associated with verocytotoxin-producing, non-O157 *Escherichia coli*. *Clin Infect Dis.* 1994; 19(4):815-6.
- Rosales A, Hofer J, Zimmerhackl LB, Jungraithmayr TC, Riedl M, Giner T, Strasak A, Orth-Höller D, Würzner R, Karch H, et al. Need for long-term follow-up in enterohemorrhagic *Escherichia coli*-associated hemolytic uremic syndrome due to late-emerging sequelae. *Clin Infect Dis.* 2012;54(10):1413-21.
- Zieg J, Dusek J, Marejkova M, Limrova P, Blazek D, Pavlicek P, Grega M, Janda J, Blahova K. Fatal case of diarrhea-associated hemolytic uremic syndrome with severe neurologic involvement. *Pediatr Int.* 2012;54(1):166-7.
- Zimmerhackl LB, Rosales A, Hofer J, Riedl M, Jungraithmayr T, Mellmann A, Bielaszewska M, Karch H. Enterohemorrhagic *Escherichia coli* O26:H11-associated hemolytic uremic syndrome: bacteriology and clinical presentation. *Semin Thromb Hemost.* 2010;36(6):586-93.
- Dallman TJ, Byrne L, Launders N, Glen K, Grant KA, Jenkins C. The utility and public health implications of PCR and whole genome sequencing for the detection and investigation of an outbreak of Shiga toxin-producing *Escherichia coli* serogroup O26:H11. *Epidemiol Infect.* 2015; 143(8):1672-80.
- Germinario C, Caprioli A, Giordano M, Chironna M, Gallone MS, Tafuri S, Minelli F, Maugliani A, Michelacci V, Santangelo L, et al. Community-wide outbreak of haemolytic uraemic syndrome associated with Shiga toxin 2-

- producing *Escherichia coli* O26:H11 in southern Italy summer 2013. *Euro Surveill.* 2016;21(38)
23. Hines JZ, Bancroft J, Powell M, Hedberg K. Case finding using syndromic surveillance data during an outbreak of Shiga toxin-producing *Escherichia coli* O26 infections, Oregon 2015. *Public Health Rep.* 2017;132(4):448–50.
 24. Kanayama A, Yahata Y, Arima Y, Takahashi T, Saitoh T, Kanou K, Kawabata K, Sunagawa T, Matsui T, Oishi K. Enterohemorrhagic *Escherichia coli* outbreaks related to childcare facilities in Japan, 2010–2013. *BMC Infect Dis.* 2015;15:539.
 25. Luna-Gierke RE, Griffin PM, Gould LH, Herman K, Bopp CA, Strockbine N, Mody RK. Outbreaks of non-O157 Shiga toxin-producing *Escherichia coli* infection: USA. *Epidemiol Infect.* 2014;142(11):2270–80.
 26. Misselwitz J, Karch H, Bielazewska M, John U, Ringelmann F, Rönnefarth G, Patzer L. Cluster of hemolytic-uremic syndrome caused by Shiga toxin-producing *Escherichia coli* O26:H11. *Pediatr Infect Dis J.* 2003;22(4):349–54.
 27. Moran-Gilad J, Rokney A, Danino D, Ferdous M, Alsana F, Baum M, Dukhan L, Agmon V, Anuka E, Valinsky L, et al. Real-time genomic investigation underlying the public health response to a Shiga toxin-producing *Escherichia coli* O26:H11 outbreak in a nursery. *Epidemiol Infect.* 2017; 145(14):2998–3006.
 28. Scavia G, Gianviti A, Labriola V, Chiani P, Maugliani A, Michelacci V, Minelli F, Tozzoli R, Caprioli A, Morabito S. A case of haemolytic uraemic syndrome (HUS) revealed an outbreak of Shiga toxin-2-producing *Escherichia coli* O26: H11 infection in a nursery, with long-lasting shedders and person-to-person transmission, Italy 2015. *J Med Microbiol.* 2018;67(6):775–82.
 29. Severi E, Vial F, Peron E, Mardh O, Niskanen T, Takkinen J. Community-wide outbreaks of haemolytic uraemic syndrome associated with Shiga toxin-producing *Escherichia coli* O26 in Italy and Romania: a new challenge for the European Union. *Euro Surveill.* 2016;21(49)
 30. Werber D, Fruth A, Liesegang A, Littmann M, Buchholz U, Prager R, Karch H, Breuer T, Tschäpe H, Ammon A. A multistate outbreak of Shiga toxin-producing *Escherichia coli* O26:H11 infections in Germany detected by molecular subtyping surveillance. *J Infect Dis.* 2002;186(3):419–22.
 31. Bielazewska M, Mellmann A, Bletz S, Zhang W, Köck R, Kossow A, Prager R, Fruth A, Orth-Höller D, Marejková M, et al. Enterohemorrhagic *Escherichia coli* O26:H11/H-: a new virulent clone emerges in Europe. *Clin Infect Dis.* 2013;56(10):1373–81.
 32. Mellmann A, Bielazewska M, Köck R, Friedrich AW, Fruth A, Middendorf B, Harmsen D, Schmidt MA, Karch H. Analysis of collection of hemolytic uraemic syndrome-associated enterohemorrhagic *Escherichia coli*. *Emerg Infect Dis.* 2008;14(8):1287–90.
 33. Zhang WL, Bielazewska M, Liesegang A, Tschäpe H, Schmidt H, Bitzan M, Karch H. Molecular characteristics and epidemiological significance of Shiga toxin-producing *Escherichia coli* O26 strains. *J Clin Microbiol.* 2000;38(6):2134–40.
 34. Delannoy S, Mariani-Kurkdjian P, Bonacorsi S, Liguori S, Ison SA, Fach P. Draft genome sequences of human-pathogenic *Escherichia coli* O26:H11 strains carrying the *stx2* gene only and circulating in France. *Genome Announc.* 2015;3(4)
 35. Delannoy S, Mariani-Kurkdjian P, Bonacorsi S, Liguori S, Fach P. Characteristics of emerging human-pathogenic *Escherichia coli* O26:H11 strains isolated in France between 2010 and 2013 and carrying the *stx2d* gene only. *J Clin Microbiol.* 2015;53(2):486–92.
 36. Grande L, Michelacci V, Bondi R, Gigliucci F, Franz E, Badouei MA, Schlager S, Minelli F, Tozzoli R, Caprioli A, et al. Whole-genome characterization and strain comparison of VT2f-producing *Escherichia coli* causing hemolytic uraemic syndrome. *Emerg Infect Dis.* 2016;22(12):2078–86.
 37. Bletz S, Bielazewska M, Leopold SR, Köck R, Witten A, Schuldes J, Zhang W, Karch H, Mellmann A. Evolution of enterohemorrhagic *Escherichia coli* O26 based on single-nucleotide polymorphisms. *Genome Biol Evol.* 2013;5(10):1807–16.
 38. Delannoy S, Mariani-Kurkdjian P, Webb HE, Bonacorsi S, Fach P. The mobilome; a major contributor to *Escherichia coli* *stx2*-positive O26:H11 strains intra-serotype diversity. *Front Microbiol.* 2017;8:1625.
 39. Ogura Y, Gotoh Y, Itoh T, Sato MP, Seto K, Yoshino S, Isobe J, Etoh Y, Kurogi M, Kimata K, et al. Population structure of *Escherichia coli* O26:H11 with recent and repeated *stx2* acquisition in multiple lineages. *Microb Genom.* 2017;3(11)
 40. Nataro JP, Seriwatana J, Fasano A, Maneval DR, Guers LD, Noriega F, Dubovsky F, Levine MM, Morris JG. Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and *Shigella* strains. *Infect Immun.* 1995;63(12):4721–8.
 41. Krüger A, Lucchesi PM. Shiga toxins and *stx* phages: highly diverse entities. *Microbiology.* 2015;161(Pt 3):451–62.
 42. Laing CR, Zhang Y, Gilmour MW, Allen V, Johnson R, Thomas JE, Gannon VP. A comparison of Shiga-toxin 2 bacteriophage from classical enterohemorrhagic *Escherichia coli* serotypes and the German *E. coli* O104: H4 outbreak strain. *PLoS One.* 2012;7(5):e37362.
 43. Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, et al. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One.* 2011; 6(7):e22751.
 44. Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han CG, Ohtsubo E, Nakayama K, Murata T, et al. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* 2001;8(1):11–22.
 45. Perna NT, Plunkett G, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, et al. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature.* 2001;409(6819):529–33.
 46. Ogura Y, Mondal SI, Islam MR, Mako T, Arisawa K, Katsura K, Ooka T, Gotoh Y, Murase K, Ohnishi M, et al. The Shiga toxin 2 production level in enterohemorrhagic *Escherichia coli* O157:H7 is correlated with the subtypes of toxin-encoding phage. *Sci Rep.* 2015;5:16663.
 47. Allerberger F, Friedrich AW, Grif K, Dierich MP, Dornbusch HJ, Mache CJ, Nachbauer E, Freilinger M, Rieck P, Wagner M, et al. Hemolytic-uremic syndrome associated with enterohemorrhagic *Escherichia coli* O26:H infection and consumption of unpasteurized cow's milk. *Int J Infect Dis.* 2003;7(1):42–5.
 48. Zweifel C, Cernela N, Stephan R. Detection of the emerging Shiga toxin-producing *Escherichia coli* O26:H11/H- sequence type 29 (ST29) clone in human patients and healthy cattle in Switzerland. *Appl Environ Microbiol.* 2013;79(17):5411–3.
 49. Zhou K, Ferdous M, de Boer RF, Kooistra-Smid AM, Grundmann H, Friedrich AW, Rossen JW. The mosaic genome structure and phylogeny of Shiga toxin-producing *Escherichia coli* O104:H4 is driven by short-term adaptation. *Clin Microbiol Infect* 2015, 21(5):468.e467–418.
 50. Beutin L, Hammerl JA, Reetz J, Strauch E. Shiga toxin-producing *Escherichia coli* strains from cattle as a source of the Stx2a bacteriophages present in enteroaggregative *Escherichia coli* O104:H4 strains. *Int J Med Microbiol.* 2013;303(8):595–602.
 51. Garred O, van Deurs B, Sandvig K. Furin-induced cleavage and activation of Shiga toxin. *J Biol Chem.* 1995;270(18):10817–21.
 52. Fagerquist CK, Sultan O. Top-down proteomic identification of furin-cleaved α -subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 using MALDI-TOF-TOF-MS. *J Biomed Biotechnol.* 2010;2010:123460.
 53. Li XP, Tumer NE. Differences in ribosome binding and sarcin/ricin loop depurination by Shiga and ricin holotoxins. *Toxins (Basel).* 2017;9(4)
 54. Tian S. A 20 residues motif delineates the furin cleavage site and its physical properties may influence viral fusion. *Biochemistry Insights.* 2009;2:BCI.S2049.
 55. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, Pruitt KD, Sayers EW. GenBank. *Nucleic Acids Res.* 2018;46(D1):D41–7.
 56. Gabrielsen C, Drablos F, Afset JE. Genome sequences of 11 Shiga toxin-producing *Escherichia coli* strains. *Genome Announc.* 2015;3(3)
 57. Mellmann A, Bletz S, Böking T, Kipp F, Becker K, Schultes A, Prior K, Harmsen D. Real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. *J Clin Microbiol.* 2016;54(12):2874–81.
 58. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 2016;44(14):6614–24.
 59. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One.* 2014;9(8):e104984.
 60. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5(3):e9490.
 61. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33(7):1870–4.
 62. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezuk Y, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* 2013;41(Web Server issue):W29–33.
 63. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012;28(12):1647–9.

64. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain loss and rearrangement. *PLoS One*. 2010;5(6):e11147.
65. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
66. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134.
67. Bauwens A, Bielaszewska M, Kemper B, Langehanenberg P, von Bally G, Reichelt R, Mulac D, Humpf HU, Friedrich AW, Kim KS, et al. Differential cytotoxic actions of Shiga toxin 1 and Shiga toxin 2 on microvascular and macrovascular endothelial cells. *Thromb Haemost*. 2011;105(3):515–28.
68. Ogura Y, Ooka T, Iguchi A, Toh H, Asadulghani M, Oshima K, Kodama T, Abe H, Nakayama K, Kurokawa K, et al. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2009;106(42):17939–44.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

