# Whole genome sequencing based typing and characterisation of Shiga-toxin producing *Escherichia coli* strains belonging to O157 and O26 serotypes and isolated in dairy farms

**Frederique Pasquali,[1] Federica Palma,[1] Marcello Trevisani,[2] Antonio Parisi,[3] Alex Lucchi,[1] Alessandra De Cesare,[1] Gerardo Manfreda[1]**

**[1]Department of Agricultural and Food Sciences, Alma Mater Studiorum – University of Bologna; [2]Department of Veterinary Medical Sciences, Alma Mater Studiorum – University of Bologna; [3]Istituto Zooprofilattico Sperimentale della Puglia e della Basilicata, Bari, Italy**

## Abstract

In the present study, the genetic relationships as well as the virulome and resistome of newly sequenced O26 and O157 Shiga-toxin producing *E. coli* (STEC) isolates, collected from dairy farms in Italy, were investigated in comparison to publicly available genomes collected worldwide. The whole genome of Italian isolates was sequenced on Illumina MiSeq Platform. Reads quality control, *de novo* draft genome assembly, species confirmation and the 7-loci Multi-Locus Sequence Type assignment were performed using INNUca pipeline. Reference-based SNPs calling was performed on O157 and O26 genomes, separately, mapping contigs to high-quality finished genomes. Virulence and antimicrobial resistance determinants were detected *in silico* using the tool ABRicate. Phylogenetic reconstructions revealed that genomes clustered mainly based on their 7-loci MLST type. The virulome of tested genomes included 190 determinants. O157 genomes carried *chu* genes associated to heme mediated iron uptake, whereas O26 genomes harboured genes *ybt* associated to siderophore mediated iron uptake. Resistome analysis showed the presence of *tet*(34) on all but one O157 genomes and on only one O26 genomes. Only 4 genomes carried genes associated to multiresistance. In the present study, the genes *chu* and *ybt* were identified as potential biomarker for the differentiation of O157 and O26 serotypes.

## Introduction

Shiga-toxin producing *Escherichia coli* (STEC) is an important zoonotic pathogen associated with infections in humans, sometimes with severe symptoms such as haemorrhagic colitis and haemolytic uremic syndrome (HUS) (Griffin and Karmali, 2017). Cattle are considered to be one of the main reservoirs of the bacterium along with sheep and goats (EFSA and ECDC, 2017). In recent years the percentage of confirmed human cases showed a slight increase from 5,680 in 2012 to 6,378 in 2016, confirming STEC infections as the fourth most relevant zoonosis in Europe. The most identified serogroups in humans, food and animals are O157 and O26, with a recent increase in O26 detection (EFSA and ECDC, 2017). STEC O157 and O26 are among the 6 serotypes which have been regulated. After the large O104:H4 outbreak occurred in 2011, a microbiological criterion of "absence in 25 g" of STEC O157, O26, O111, O103, O145 and O104:H4 in sprouted seeds was added to Regulation (EC) No 2073/2005 (Regulation (EC) No 209/2013).

Both STEC O157 and O26 were described as harbouring different essential virulence factors: i) the Shiga-toxin genes *stx1* and *stx2*; ii) the *eae* gene coding for intimin. The genes *stx1* and *stx2* are characterised by three ($stx_{1a}$, $stx_{1c}$ and $stx_{1d}$) and seven ($stx_{2a\text{-}g}$) variants respectively, all linked to a different virulence potential with $stx_{2a\text{-}d}$ as strongly associated to sever diarrhoea and HUS (Amézquita-López *et al*., 2017). The gene *eae* is included in the locus of enterocyte effacement (LEE) and described as essential for the attachment of *E. coli* to intestinal epithelial cells (Amézquita-López *et al*., 2017). After the STEC German outbreak of 2011 associated to an *eae*-negative O104:H4 strain, it was observed that other genes might also be effectively involved in the adhesion of *E. coli* to epithelial cells: the plasmid located *aggR* gene or the chromosomally encoded *aaiC* gene. Based on these observations, the combination of *stx2* and *eae* or *stx2* and *aggR/aaiC* was established as reliable predictors of high risk of severe illness (JEMRA, 2016). Virulence genes such as *ehxA* and *hlyA*, coding for haemolysin, were additionally described. In particular *ehxA* was categorised in 4 subtypes with subtype B and C significantly associated to O157 and O26 respectively (Lorenz *et al*., 2013).

Nevertheless, these combinations of genes would have failed to predict the severe illness caused by the "French clone" described as Enterohemorrhagic *E. coli* (EHEC) strain and responsible for sporadic cases from 2010 to 2011. This clone was exclusively *stx2* positive (Delannoy *et al*., 2015; Bielaszewska *et al*., 2013). Moreover, different enterohaemorrhagic clones harbouring none of the *stx* genes were detected and named EHEC-like. Additional combinations of virulence determinants were suggested as potential biomarker predictors of severe illness of EHEC and EHEC-like O26: the *espK* gene with either *espV*, *ureD* and/or *Z2098* and $CRISPR_{O26:H11}$ (Bugarel *et al*., 2011; Delannoy *et al*., 2013; Douëllou *et al*., 2017). In particular, the *esp* genes are linked to type III secreted effector proteins of EHEC, whereas the *ureD* gene is essential for the synthesis of urease accessory protein D linked to the enhancement of acid tolerance during passage though the stomach (Steyert *et al*., 2011). As far as antimicrobial resistance (AMR) is concerned, the prevalence of AMR in STEC differs significantly among Europe. French and English studies reported an AMR prevalence below 20% with the exception of O26 English isolates showing a higher percentage of around

Correspondence: Frédérique Pasquali, Department of Agricultural and Food Sciences, *Alma Mater Studiorum*-University of Bologna, via del Florio 2, 40064 Ozzano dell'Emilia (BO), Italy.
Tel: +39.051.2097862 - Fax: +39.051.2097852.
E-mail: frederique.pasquali@unibo.it

Key words: Shiga-toxin producing *Escherichia coli*, dairy farm, Whole Genome Sequencing, typing, virulome, resistome.

Contributions: AL, AD and MT collected data and performed the lab experiment steps including culture detection, PCR screening test and DNA extraction, AP performed the whole genome sequencing, FrP, FeP and GM designed the study; FeP run the bioinformatics analyses; FrP wrote the manuscript; GM and FeP reviewed the manuscript and contribute to references search.

Conflict of interests: the authors declare no potential conflict of interests.

Funding: the work was supported by the EU founded project COMPARE (Grant Agreement N° 643476).

Received for publication: 7 July 2018.
Revision received: 15 October 2018.
Accepted for publication: 22 October 2018.


*Italian Journal of Food Safety 2018; 7:7673*
*doi:10.4081/ijfs.2018.7673*

36% (Day *et al*., 2017; Um *et al*., 2018). On the other hand, a Spanish study observed 75.3% of isolates to be carrying plasmid-mediated colistin resistance (García *et al*., 2018). A Romanian study on AMR prevalence in young livestock animals observed an increase of multidrug resistance (MDR) from 11% during the 1980s to 40% between 2000 and 2016 (Chirila *et al*., 2017).

Whole Genome Sequencing (WGS) based analyses have recently revealed their great resolution in pathogen typing as well as identification of novel or known genes related to specific phenotypes such as virulence and antimicrobial resistance (Nadon *et al*., 2017; Revez *et al*., 2017; Leopold *et al*., 20014; Oniciuic *et al*., 2018). Studies on whole genome sequencing data aimed to characterise the virulence profiles of O26 or O157 clones are emerging (Holmes *et al*., 2018; Worley *et al*., 2017; Usein *et al*., 2017; Gonzalez-Escalona *et al*., 2016). However, to the best of our knowledge, none has compared the two serotypes.

In the present study, the genetic relationships as well as the virulome and resistome of newly sequenced isolates of O26 and O157 STEC were compared to publicly available genomes.

**Table 1. Newly sequenced (labelled EC) and publicly available genomes of O26 and O157 with *stx, eae, hlyA* and *espK* genes related virulence profiles.**

| Genome | serotype | source | Country | year | $stx_{1A}$ | $stx_{1B}$ | $stx_{2A}$ | $stx_{2B}$ | *eae* | *hlyA* | *espK* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EC1 | O26 | bulk milk | Italy | 2011 | - | + | - | - | + | - | + |
| EC17 | O26 | bulk milk | Italy | 2011 | - | - | - | - | + | - | - |
| EC22 | O26 | bulk milk | Italy | 2011 | - | - | - | - | + | + | - |
| EC3 | O26 | bulk milk | Italy | 2009 | - | - | - | - | + | + | - |
| EC4 | O26 | milk filters | Italy | 2009 | - | - | - | - | + | + | - |
| INNUENDO_STEC_AU_063 | O26 | human | Austria | 2015 | - | - | + | + | + | + | + |
| INNUENDO_STEC_AU_064 | O26 | human | Austria | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_AU_065 | O26 | human | Austria | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_AU_066 | O26 | human | Austria | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_114 | O26 | human | Finland | 2014 | - | - | + | + | + | + | - |
| SAMD00064361 | O26 | human | Japan | 2013 | - | + | + | + | + | + | + |
| SAMN08724660 | O26 | cattle | US | 2009 | - | + | + | + | + | + | + |
| EC12 | O157 | milk filters | Italy | 2009 | - | - | - | - | + | + | + |
| EC2 | O157 | milk filters | Italy | 2011 | - | + | + | + | + | - | + |
| EC33 | O157 | cattle hide | Italy | 2007 | - | - | + | + | + | + | - |
| EC9 | O157 | milk filters | Italy | 2007 | - | + | - | - | + | + | - |
| SAMN01911278 | O157 | human | Japan | 1996 | + | + | + | + | + | + | + |
| SAMN06349171 | O157 | cattle | Canada | 2002 | - | + | + | + | + | + | + |
| SAMN06349172 | O157 | cattle | Canada | 2002 | - | + | + | + | + | + | + |
| SAMN06349173 | O157 | cattle | Canada | 2002 | - | + | + | + | + | + | + |
| SAMN07224767 | O157 | cattle | Francia | 2015 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_003 | O157 | cattle | Finland | 2014 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_007 | O157 | cattle | Finland | 2012 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_015 | O157 | cattle | Finland | 2012 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_020 | O157 | cattle | Finland | 2013 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_033 | O157 | environment | Finland | 2013 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_042 | O157 | environment | Finland | 2014 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_067 | O157 | human | Finland | 2014 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_071 | O157 | human | Finland | 2012 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_077 | O157 | human | Finland | 2014 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_084 | O157 | human | Finland | 2010 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_088 | O157 | human | Finland | 2013 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_092 | O157 | human | Finland | 2010 | - | + | + | + | + | + | + |
| INNUENDO_STEC_FI_094 | O157 | human | Finland | 2009 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_102 | O157 | human | Finland | 2011 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_106 | O157 | human | Finland | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_109 | O157 | human | Finland | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_111 | O157 | human | Finland | 2013 | - | - | + | + | + | + | + |
| INNUENDO_STEC_FI_116 | O157 | human | Finland | 2014 | - | - | + | + | + | + | + |
| SAMN06159501 | O157 | human | US | 2016 | - | - | - | - | - | - | - |

## Materials and Methods

In the present study, 9 *E. coli* isolates were included. The isolates belong to O157 and O26 serotypes and were collected from bulk milk (n=4), milk filters (n=4) and cattle hide (N=1) (Table 1), between 2007 and 2011 and whole genome sequenced. Part of these isolates was included in a previous study on the detection of STEC in bovine dairy herds in Northern Italy (Trevisani *et al*., 2014). As previously assessed by PCR-based methods, selected isolates carried one or more of three genes: the $stx_1$ and/or $stx_2$ genes and/or the *eae* gene (Trevisani *et al*., 2014). For a wider comparison between the two serotypes, 31 publicly available draft (n=29) and complete (n=2) high-quality genomes, belonging to O157 (n=24) and O26 (n=7) STEC serotypes and collected worldwide from humans and cattle sources, were included in the study along with the newly sequenced ones (Table 1). Publicly available genomes were retrieved from NCBI as well as the INNUENDO Sequence Dataset (https://zenodo.org/record/1323690#.W73LV4huYdW) (BioProject n° PRJEB27020). Whole-genomic DNA of Italian isolates was extracted using the MagAttract HMW DNA Kit (Qiagen, Hilden, Germany). The purified DNA concentration and the quality parameter ratio 260/280 were measured by BioSpectrometer fluorescence (Eppendorf). Libraries were built using the TruSeq DNA sample Prep Kit (Illumina, Milan, Italy) and the whole genome of selected isolates was paired-end sequenced using the MiSeq platform (Illumina). Reads of 250 bp on average, were quality checked and *de novo* assembled using the INNUca v1.2 pipeline, which includes SPAdes v3.9 (https://github.com/B-UMMI/INNUca). The pipeline also includes a tool for the *in-silico* characterisation of the 7-loci Multilocus Sequence Type. Reads were submitted to Enterobase (http://enterobase.warwick.ac.uk) under accession numbers: ESC_FA4394AA (EC1); ESC_FA4390AA (EC2); ESC_FA4384AA (EC3); ESC_FA4387AA (EC4); ESC_FA4385AA (EC9); ESC_FA4395AA (EC12); ESC_FA4389AA (EC17); ESC_FA4391AA (EC22); ESC_JA4691AA (EC33).

SNP calling was performed on O157 and O26 draft genomes, separately using the open source snippy v3.2 pipeline with default settings (https://github.com/tseemann/snippy). High-quality complete genomes *E. coli* O157:H7 str. Sakai (EHEC) (Ref Seq NC_002695) and *E. coli* O26:H11 str. 11368 (Ref Seq NC_013361.1) were used as references for SNP calling of O157 and O26 genomes respectively. For each serotype, an alignment of core genome SNPs was generated by snippy and used to infer a Maximum Likelihood (ML)-based high-resolution phylogeny using the iQTree software (Nguyen *et al*., 2015). In order to evaluate the genetic distance among different STEC ST, phylogenetic trees were graphically represented with iTOL viewer (https://itol.embl.de/). The most genetically distant genome for each serotype, counting several tens of thousands of SNPs, was used to root the ML-trees.

Analyses of virulome and resistome of all genomes were performed using ABRicate (https://github.com/ tseemann/abricate/). With this tool, a BLAST search of genes included in the Virulence Factors Database (VFDB) and the Resfinder database was performed on *de novo* assemblies of newly sequenced as well as publicly available selected genomes (http://www.mgc.ac.cn/VFs/main.htm; https://cge.cbs.dtu.dk/services/ResFinder/). In particular, the VFDB database includes 2,606 curated genes related to virulence factors whereas the Resfinder database includes 1,723 genes related to antimicrobial resistance (Chen *et al*., 2016; Zankari *et al*. 2013).

## Results and Discussion

The draft genome sequences of 9 newly sequenced STEC isolates, collected from bovine dairy farms in Italy over four years (2007-2011), passed the QA/QC measures defined by INNUca pipeline. Draft genomes included from 93 to 345 contigs with a final coverage between 42X and 79X and N50 values ranging from 76377 to 208613 (Table S1).

In order to evaluate the genetic relationships among newly sequenced Italian genomes of cattle origin in comparison to public genomes isolated from cattle as well as humans worldwide, SNPs-based phylogenetic analyses were performed using snippy on the whole genome of the O157 and O26 *E. coli* strains separately. The resulting ML-trees show genomes essentially clustered based on their ST with ST11 as the most prevalent O157 ST and ST21 and ST29 as the most prevalent O26 STs (Figure 1). These STs have been already described as associated to EHEC strains (Bielaszewska *et al*., 2013; Wang *et al*., 2014).

Overall, the SNPs differences encountered within ST11 genomes of O157 isolates in relation to the reference were between 854 and 2,861 (Figure 1A). A smaller genetic distance was observed among O26 genomes of ST21 against the selected reference, with a SNPs count between 257 and 1,615 (Figure 1B), while ST29 genomes cluster showed a SNPs count ranging from 4492 to 5336 (Figure 1B). Overall, within each serotype a high genetic diversity was observed with no specific clustering of genomes based on the year or country of isolation as well as source.

The virulome and resistome of a total of 40 STEC genomes belonging to O26 and O157 serotypes were compared. The virulome of draft genomes included 190 virulence determinant genes. Among the 40 genomes tested, only one O157 isolate was positive for $stx_{1a}$ whereas 22 were positive for $stx_{1b}$, 32 for $stx_{2a}$ and $stx_{2b}$, 39 for *eae*, 36 for *hlyA* and 32 for *espK*. Five genomes did not carry $stx_1$ or $stx_2$ but carried the *eae* and/or *hlyA* genes with one genome additionally carrying the *espK* gene suggesting these five as potential EHEC-like isolates. One publicly available O157 genome did not carry any of the five virulence genes (Table 1). Similar patterns were observed in cattle and human genomes confirming the cattle genomes as potentially pathogenic for humans as previously suggested (EFSA and ECDC, 2017; Trevisani *et al*., 2014).

The heatmap of all identified 190 virulence genes is reported in Figure 2. Comparing O157 and O26 genomes, all O157 and none O26 carried *chu* genes, homologous to *shu* genes of *Shigella dysentariae*, related to the use of the heme group of haemoglobin as iron source. This mechanism was described as an efficient strategy for iron acquisition during an ongoing infection (Torres and Payne, 1997; Wyckoff *et al*., 1998; Braun, 2001). Both heme and haemoglobin were already described as significantly stimulating the growth of *E. coli* O157 and production of enterohaemolysin in comparison to non-O157 strains (Law *et al*., 1995). This observation suggests that the heme mediated iron uptake is specific for O157 serotype. Moreover, O157 genomes carried a higher number of *esp* genes: *espR*, *espX* e *espY* genes, involved in the secretion system of type III associated to the survival of the pathogen within the host cell (Galán and Wolf-Watz, 2006). On the other hand, O26 carried *flhA*, a gene associated to the biosynthesis of flagella involved in the first steps of adhesion and invasion (Haiko and Westerlund-Wikström, 2013), which was not detected in any O157 genomes. Finally, all O26 and none O157 genomes carried eight *ybt* genes (*ybtA*, *ybtE*, *ybtP*, *ybtQ*, *ybtS*, *ybtT*, *ybtU* e *ybt*X) associated to the acquisition of iron from yersinabactin, a highly relevant siderophore for the hyper-virulence of *Yersinia enterocolitica* (Pelludat *et al*., 1998). Further stud-

ies should be performed in order to evaluate whether the siderophore mediating iron uptake is significantly linked to the recent emergence of O26 as the serotype most frequently associated to haemolytic uremic syndrome (HUS) in children (EFSA and ECDC, 2017). Finally, all O26 genomes but one carried the *fyuA* gene along with the *irp1* and *irp2* genes. These *irp* genes are included in the gene cluster related to the biosynthesis of yersinabactin while *fyuA* encodes for the outmembrane receptor for this siderophore (Pelludat *et al*., 1998).

Regarding the resistome, a limited number of AMR-associated genes to antimicrobial resistance was detected (Table 2). In particular, among the newly sequenced genomes, two O26 genomes carried the *tet*(C) gene, whereas one O26 and three O157 genomes carried the *tet*(34) gene with a gene coverage of 76,34%. Among the publicly available genomes, *tet*(34) gene was detected on 23 O157 genomes, along with *aph*(6)-Id, *strA*, *sul2* and *tet*(B) in three of them. These three genomes were related to isolates collected from cattle in Canada in 2002. These genes are associated to aminoglycoside, streptomycin, sulphonamides and tetracycline resistance respectively. An additional potentially multiresistant isolate belonging to O26 serotype and collected from humans in Japan, carried the *aph*(3'')-Ib, *bla*TEM-30 and *sul2* conferring resistance to aminoglycosides, beta-lactams and sulphonamides respectively. The low detection rate of AMR genes observed across the 40 screened genomes, regardless of their serotype, was in accordance to previously reported data (Day *et*

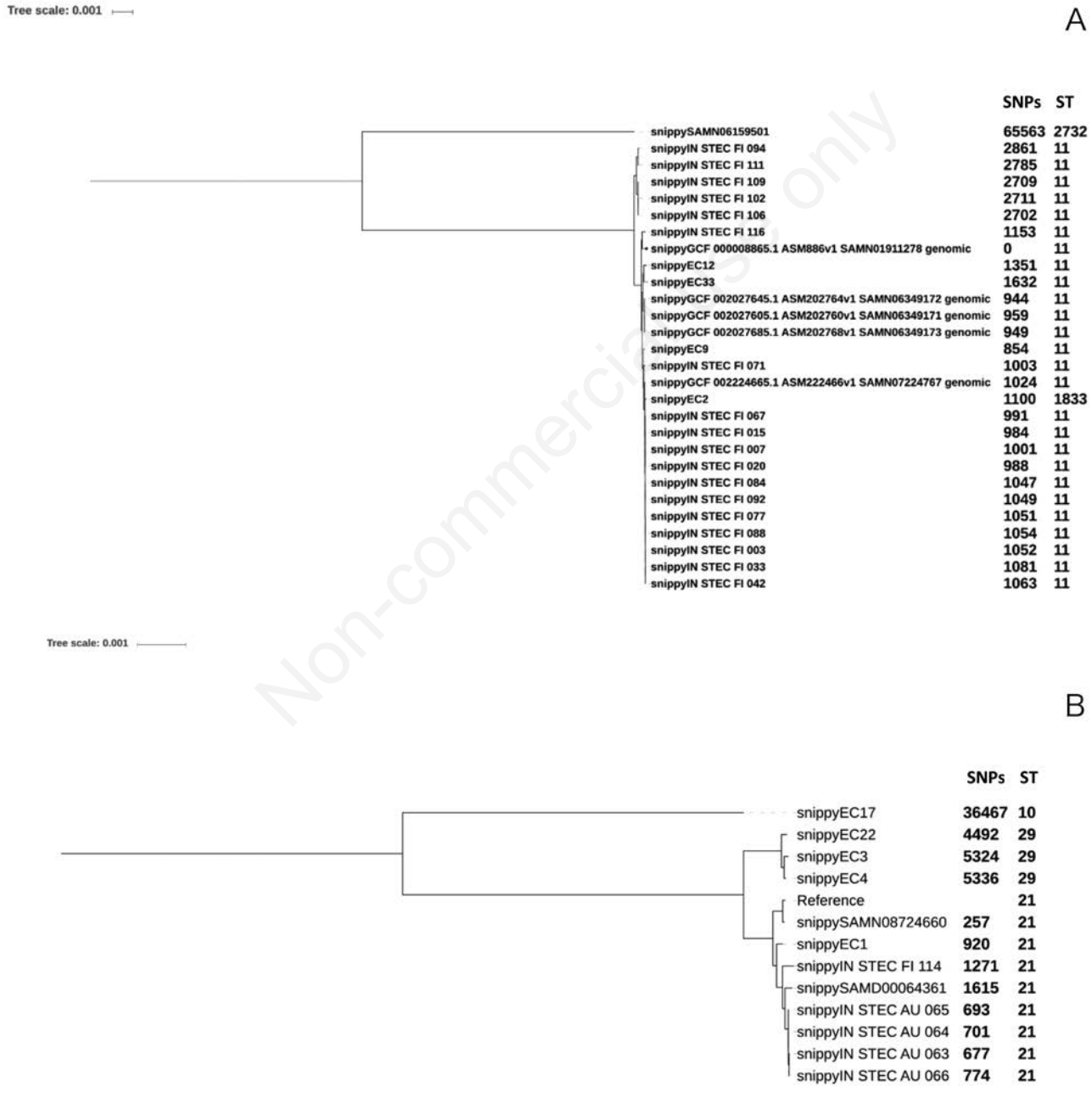




**Figure 1. Maximum likelihood phylogeny based on whole genome SNPs of O26 genomes (A) and O157 genomes (B). The number of SNPs differences to the reference is reported in the first column (SNPs) and 7-*loci* MLST type on the second column (ST).**

*al*., 2017).

Overall, all but one O157 and none but one O26 genomes harboured the *tet*(34) gene. No differences were observed both on the virulome and resistome of human *versus* cattle STEC genomes confirming the phylogenetic tree outputs which clustered together genomes from both sources.

## Conclusions

In order to compare draft genomes of the two mostly isolated serotypes associated to haemorrhagic colitis and haemolytic uremic syndrome in humans, the genetic relationships as well as the virulome and resistome of 40 STEC genomes were assessed. Newly sequenced as well as publicly available genomes of O26 and O157 serotypes isolated from cattle and humans worldwide were included in the study showing no differences in terms of genetic distance as well as of virulome and resistome compositions of human *versus* cattle genomes. Based on the virulome analysis, the presence of different virulence genes in O26 and O157 associated to siderophore and heme mediated iron uptake systems, respectively, was

**Table 2. Antimicrobial resistance determinant genes of newly sequenced (labelled EC) and publicly available genomes of O26 and O157.**

| Genome | *aph(3'')-Ib* | *aph(6)-Id* | *blaTEM-1C* | *blaTEM-30* | *strA* | *sul2* | *tet(34)** | *tet(B)* | *tet(C)* |
|---|---|---|---|---|---|---|---|---|---|
| EC1 | - | - | - | - | - | - | - | - | - |
| EC17 | - | - | - | - | - | - | + | - | - |
| EC22 | - | - | - | - | - | - | - | - | - |
| EC3 | - | - | - | - | - | - | - | - | + |
| EC4 | - | - | - | - | - | - | - | - | + |
| INNUENDO_STEC_AU_063 | - | - | - | - | - | - | - | - | - |
| INNUENDO_STEC_AU_064 | - | - | - | - | - | - | - | - | - |
| INNUENDO_STEC_AU_065 | - | - | - | - | - | - | - | - | - |
| INNUENDO_STEC_AU_066 | - | - | - | - | - | - | - | - | - |
| INNUENDO_STEC_FI_114 | - | + | - | - | + | - | - | - | - |
| SAMD00064361 | + | | + | - | - | + | - | - | - |
| SAMN08724660 | - | - | - | - | - | - | - | - | - |
| EC12 | - | - | - | - | - | - | + | - | - |
| EC2 | - | - | - | - | - | - | + | - | - |
| EC33 | - | - | - | - | - | - | + | - | - |
| EC9 | - | - | - | - | - | - | + | - | - |
| SAMN01911278 | - | - | - | - | - | - | + | - | - |
| SAMN06349171 | - | + | - | - | + | + | + | + | - |
| SAMN06349172 | - | + | - | - | + | + | + | + | - |
| SAMN06349173 | - | + | - | - | + | + | + | + | - |
| SAMN07224767 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_003 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_007 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_015 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_020 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_033 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_042 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_067 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_071 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_077 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_084 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_088 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_092 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_094 | - | - | - | + | - | - | + | - | - |
| INNUENDO_STEC_FI_102 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_106 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_109 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_111 | - | - | - | - | - | - | + | - | - |
| INNUENDO_STEC_FI_116 | - | - | - | - | - | - | + | - | - |
| SAMN06159501 | - | - | - | - | - | - | - | - | - |

**tet(34)*: detected with 76.34 % of coverage

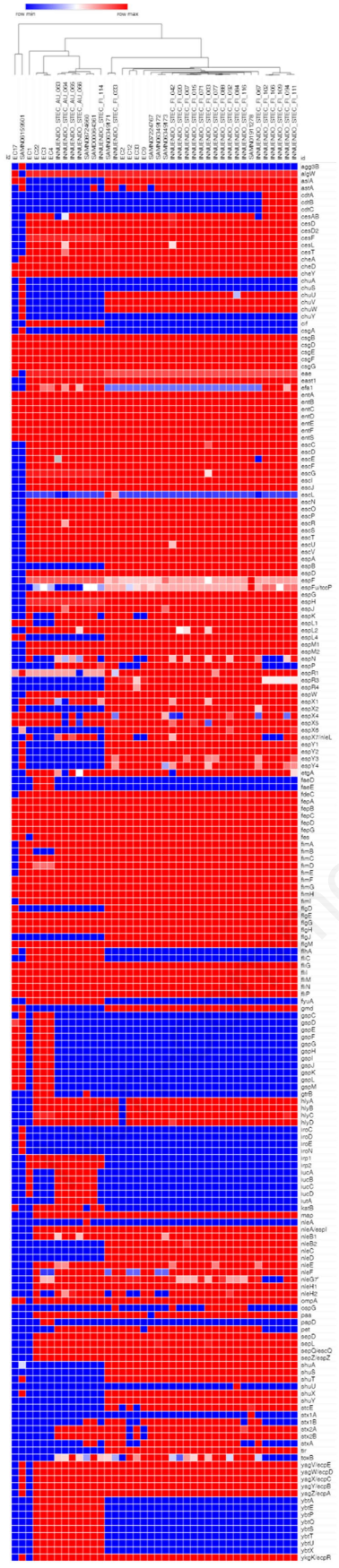

**Figure 2. Heatmap of the virulome of O157 and O26 genomes. Colours from blue to red correspond to a range of 0-100% coverage of the gene sequences of the Virulence Finder Database.**

observed. These genes, *chu* and *ybt*, could be used as biomarkers for the identification of O157 and O26 genomes. Further analyses are also required to better understand how the iron and other metals (*i.e.* copper) uptake system impact on the pathogenicity of the strain belonging to these two haemorrhagic serotypes. Additional investigations based on genome-wide association study (GWAS) including a higher number of publicly available genomes should be performed in order to confirm the statistically significant relevance of the serotype-specific genes identified in this study.

## References


Amézquita-López BA, Soto-Beltrán M, Lee BG, Yabao JC, Quiñones B, 2017. Isolation, genotyping and antimicrobial resistance of Shiga toxin-producing Escherichia coli, J Microbiol Immunol Infect [Epub ahead of print]

Bielaszewska M, Mellmann A, Bletz S, Zhang W, Köck R, Kossow A, Prager R, Fruth A, Orth-Höller D, Marejková M, Morabito S, Caprioli A, Piérard D, Smith G, Jenkins C, Čurová K, Karch H, 2013. Enterohemorrhagic Escherichia coli O26:H11/H−: A New Virulent Clone Emerges in Europe. Clin Infect Dis 56:1373-81.

Bielaszewska M, Zhang W, Mellmann A, Karch H, 2007. Enterohaemorrhagic Escherichia coli O26:H11/H: a human pathogen in emergence. Berl Munch Tierarztl Wochenschr 120:279-87.

Braun V, 2001. Iron uptake mechanisms and their regulation in pathogenic bacteria. Int J Med Microbiol 291:67-79.

Buchrieser C, Prentice M, Carniel E, 1998. The 102-Kilobase Unstable Region of Yersinia pestis Comprises a High-Pathogenicity Island Linked to a Pigmentation Segment Which Undergoes Internal Rearrangement. J Bacteriol 180:2321-9.

Bugarel M, Beutin L, Scheutz F, Loukiadis E, Fach P, 2011. Identification of genetic markers for differentiation of Shiga toxin-producing, enteropathogenic, and avirulent strains of Escherichia coli O26. Appl Environ Microbiol 77:2275-81.

Bugarel, M, Beutin, L, Fach, P, 2010. Low-density macroarray targeting non-locus of enterocyte effacement effectors (nle genes) and major virulence factors of Shiga toxin-producing Escherichia coli (STEC): a new approach for molecular risk assessment of STEC isolates. Appl Environ Microbiol 76:203-11.

Chen L, Zheng D, Liu B, Yang J, Jin Q, 2016. VFDB 2016: hierarchical and

refined dataset for big data analysis—10 years on. Nucleic Acids Research. 44:D694-7

Chirila F, Tabaran A, Fit N, Nadas G, Mihaiu M, Tabaran F, Cătoi C, Reget OL, Dan SD, 2017. Concerning increase in antimicrobial resistance in Shiga toxin-producing Escherichia coli isolated from young animals during 1980-2016. Microbes Environ 32:252-9.

Commission Regulation (EU) No 209/2013 of 11 March 2013 amending Regulation (EC) No 2073/2005 as regards microbiological criteria for sprouts and the sampling rules for poultry carcases and fresh poultry meat Text with EEA relevance. OJ L 68, 12.3.2013, p. 19-23.

Day M, Doumith M, Jenkins C, Dallman TJ, Hopkins KL, Elson R, Godbole G, Woodford N, 2017. Antimicrobial resistance in Shiga toxin-producing Escherichia coli serogroups O157 and O26 isolated from human cases of diarrhoeal disease in England. J Antimicrob Chemother 72:145-52.

Delannoy S, Beutin L, Fach P, 2013. Towards a molecular definition of enterohemorrhagic Escherichia coli (EHEC): detection of genes located on O island 57 as markers to distinguish EHEC from closely related enteropathogenic E. coli strains. J Clin Microbiol 51:1083-8.

Delannoy S, Mariani-Kurkdjian P, Bonacorsi S, Liguori S, Fach P, 2015. 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 53:486-92.

Douëllou T, Delannoy S, Ganet S, Fach P, Loukiadis E, Montel MC, Sergentet-Thevenot D, 2017. Molecular characterization of O157:H7, O26:H11 and O103:H2 Shiga toxin-producing Escherichia coli isolated from dairy products. Int J Food Microbiol 253:59-65.

EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control), 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. EFSA Journal 15:5077.

García V, García-Meniño I, Mora A., Flament-Simon SC, Díaz-Jiménez D., Blanco JE, Alonso MP, Blanco J, 2018. Co-occurrence of mcr-1, mcr-4 and mcr-5 genes in multidrug-resistant ST10 Enterotoxigenic and Shiga toxin-producing Escherichia coli in Spain (2006-2017). Int J Antimicrobial agents 52:104-8.

Galán JE, Wolf-Watz H, 2006. Protein delivery into eukaryotic cells by type III secretion machines. Nature 444:567-573.

Gonzalez-Escalona N, Toro M, Rump LV, Cao G, Nagaraja TG, Meng J, 2016. Virulence gene profiles and clonal relationships of Escherichia coli O26:H11 isolates from feedlot cattle as determined by Whole Genome Sequencing. Appl Environ Microbiol 82:3900-12.

Griffin PM, Karmali MA, 2017. Emerging Ppublic health challenges of Shiga toxin-producing Escherichia coli related to changes in the pathogen, the population, and the Environment, Clin Infect Dis 64:371-76.

Haiko J, Westerlund-Wikström B, 2013. The role of the bacterial flagellum in adhesion and virulence. Biology 2:1242-67.

Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L, 2018. Validation of whole-genome sequencing for identification and characterization of Shiga toxin-producing Escherichia coli to produce standardized data to enable data sharing. J Clin Microbiol 56:e01388-17.

JEMRA, 2016. Joint FAO/WHO Core Expert Group Meeting on VTEC/STEC.

Koh EI, Robinson AE, Bandara N, Rogers BE, Henderson JP, 2017. Copper import in Escherichia coli by the yersiniabactin metallophore system. Nat Chem Biol 13:1016-21.

Law D, Kelly J, 1995. Use of heme and hemoglobin by Escherichia coli O157 and other Shiga-like-toxin-producing E. coli serogroups. Infect Immun 63:700-2.

Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A, 2014. Bacterial Whole-Genome Sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. J Clin Microbiol 52:2365-70.

Lorenz SC, Son I, Maounounen-Laasri A, Lin A, Fischer M, Kase JA, 2013. Prevalence of hemolysin genes and comparison of ehxA subtype patterns in Shiga toxin-producing Escherichia coli (STEC) and non-STEC strains from clinical, food, and animal sources. Appl Environ Microbiol 79:6301-11.

Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Kam Kai M, Perez E, Trees E, Kubota K, Takkinen J, Nielsen EM, Carleton H, FWD-NEXT Expert Panel, 2017. PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill. 22:pii=30544.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ, 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol 32:268-74.

Oniciuc EA, Likotrafiti E, Alvarez-Molina A, Prieto M, Santos JA, Alvarez-Ordóñez A, 2018. The present and future of Whole Genome Sequencing (WGS) and Whole Metagenome Sequencing (WMS) for surveillance of antimicrobial resistant microorganisms and antimicrobial resistance genes across the food chain. Genes 9:268.

Pelludat C, Rakin A, Jacobi CA, Schubert S, Heesemann J, 1998. The yersiniabactin biosynthetic gene cluster of Yersinia enterocolitica: organization and siderophore-dependent regulation. J Bacteriol 180:538-46.

Rakin A, Schneider L, Podladchikova O, 2012. Hunger for iron: the alternative siderophore iron scavenging systems in highly virulent Yersinia. Frontiers Cell Infect Microbiol 2:151.

Revez J, Espinosa L, Albiger B, Leitmeyer KCe, Struelens MJ, ECDC National Microbiology Focal Points and Experts Group, 2017. Survey on the use of Whole-Genome Sequencing for infectious diseases surveillance: rapid expansion of European national capacities, 2015-2016. Frontiers in Public Health 5:347.

Steyert SR, Rasko DA, Kaper JB, 2011. Functional and Phylogenetic Analysis of ureD in Shiga Toxin-Producing Escherichia coli. J Bacteriol 193:875-886.

Torres AG, Payne SM, 1997. Heme iron-transport system in enterohaemorrhagic Escherichia coli O157:H7. Mol Microbiol 23:825-3.

Trevisani M, Mancusi R, Delle Donne G, Bacci C, Bassi L. Bonardi S, 2014. Detection of Shiga toxin (Stx)-producing Escherichia coli (STEC) in bovine dairy herds in Northern Italy. Int J Food Microbiol 184:45-49.

Um MM, Brugère H, Kérourédan M, Oswald E, Bibbal D, 2018. Antimicrobial resistance profiles of Enterohemorrhagic and Enteropathogenic Escherichia coli of serotypes O157:H7, O26:H11, O103:H2, O111:H8, O145:H28 compared to Escherichia coli isolated from the same adult cattle. Microb Drug

Resist. [Epub ahead of print]
Usein CR, Ciontea AS, Militaru CM, Condei M, Dinu S, Oprea M, Cristea D, Michelacci V, Scavia G, Zota LC, Zaharia A, Morabito S, 2017. Molecular characterisation of human Shiga toxin-producing Escherichia coli O26 strains: results of an outbreak investigation, Romania, February to August 2016. Euro Surveill 22:pii=17-00148.
Wang S, Zhang S, Liu Z, Liu P, Shi Z, Wei J, Shao D, Li B, Ma Z, 2014. Molecular characterization of Enterohemorrhagic E. coli O157 isolated from animal fecal and food samples in Eastern China. The Scientific World Journal 2014:946394.
Worley JN, Flores KA, Yang X, Chase JA, Cao G, Tang S, Meng J, Atwill ER, 2017. Prevalence and genomic characterization of Escherichia coli O157:H7 in cow-calf herds throughout California. Appl Environ Microbiol 83:e00734-17.
Wyckoff EE, Duncan D, Torres AG, Mills M, Maase K, Payne SM, 1998. Structure of the Shigella dysenteriae heme transport locus and its phylogenetic distribution in enteric bacteria. Mol Microbiol 28:1139-52.
Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agersø Y, Lund O, Larsen MV, Aarestrup FM, 2013. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother 68:771-7.