AEM Accepted Manuscript Posted Online 18 September 2020
Appl. Environ. Microbiol. doi:10.1128/AEM.01913-20

Genomic characterization of Enterotoxigenic *E. coli* F4 and F18 positive strains from post-weaning diarrhea in pigs

Vanesa García[a,b], Michela Gambino[a], Karl Pedersen[c], Sven Haugegaard[d], John Elmerdahl Olsen[a] and Ana Herrero-Fresno[a,#]

[a]Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg, Denmark

[b]Laboratorio de Referencia de *Escherichia coli* (LREC), Departamento de Microbioloxía e Parasitoloxía, Facultade de Veterinaria, Universidade de Santiago de Compostela (USC), Lugo, Spain

[c]Statens Veterinärmedicinska Anstalt, Uppsala, Sweden

[d]SEGES Danish Pig Research Centre, Denmark.

[#]Corresponding author, Ana Herrero-Fresno,

Email: ahefr@sund.ku.dk

Running title: Genome analysis of *E. coli* from post-weaning diarrhea

Keywords: pigs, diarrhea, Enterotoxigenic *E. coli*, whole genome sequencing, bioinformatics

ABSTRACT

This study aimed to characterize *in silico* Enterotoxigenic *E. coli* F4- and F18-positive-isolates (n=90) causing swine post-weaning diarrhea, including pathogenic potential, phylogenetic relationship, antimicrobial and biocide resistance (R), prophage content and metal tolerance rates. F4-strains belonged mostly to the O149 and O6 serogroups and ST100 and ST48 MLST-types. F18-strains were mainly assigned to the O8 and O147 serogroups and ST10, ST23 and ST42. The highest rates of antimicrobial-R were found against streptomycin, sulfamethoxazole, tetracycline, trimethoprim and ampicillin. No resistance was found towards ciprofloxacin, cefotaxime, ceftiofur and colistin. Genes conferring tolerance to copper (showing highest diversity), cadmium, silver and zinc were predicted in all genomes. Enterotoxins (*ltcA*; 100 % F4, 62 % F18, *astA;* 100 % F4, 38.1 % F18, *sta;*18.8 % F4, 38.1 % F18, *stb;* 100 % F4, 76.2 % F18) and fimbriae encoding genes typed as F4ac and F18ac were detected in all the strains, in addition to up to 16 other virulence genes in individual strains. Phage analysis predicted between 7 and 20 different prophage regions in each strain. High diversity of plasmid replicons was shown; IncFII, IncFIB and IncFIC were prevalent among F4-isolates while IncI1, and IncX1 were dominant among F18-strains. Interestingly, F4-isolates from the early 90´s belonged to the same clonal group detected for most of the F4-strains from 2018-2019 (ONT:H10-A-ST100-CH27-0). The low number of SNPs differences between the oldest and recent F4-ST100-isolates suggests a relative stable genome. Overall, the isolates analyzed in this study showed remarkable different genetic traits depending on the fimbria type.

IMPORTANCE

Diarrhea in the post-weaning period due to Enterotoxigenic *E. coli* (ETEC) is an economically relevant disease in pig production worldwide. In Denmark, prevention is mainly achieved by Zinc-oxide administration (discontinued by 2022). Besides, a breeding program has been implemented

aiming to reduce this illness. Treatment with antimicrobials contributes to the problem of antimicrobial resistance (AMR) development. As a novelty, this study aims to deeply understand the genetic population structure and variation among isolates diarrhea-associated by whole genome sequencing characterization. ST100-F4ac is the dominant clonal group circulating in Danish herds and showed high similarity to ETEC-ST100-isolates from China, USA and Spain. High rates of AMR and high diversity of virulence genes were detected. Characterization of ETEC diarrhea-related is important for understanding the disease epidemiology and pathogenesis, and for implementation of new strategies aiming to reduce the impact of the disease in pig production.

INTRODUCTION

Post-weaning diarrhea (PWD) affects pigs after weaning, leading to significant economic costs for the pig industry due to weight loss, mortality, as well as cost of prevention (i.e. vaccination), treatment and handling (1, 2). In addition to sudden death or profuse diarrhea, the disease is accompanied by growth retardation in surviving piglets (1, 3). During acute outbreaks, mortality due to PWD may reach 20-30% over a 1 to 2 months' time span among infected pigs (3). Besides, as it is one of the most common reasons for use of antimicrobials in pig industry worldwide, PWD significantly contributes to the problem of antimicrobial resistance (AMR) development (2, 4, 5).

Enterotoxigenic *E. coli* (ETEC) is the main etiological agent involved in PWD worldwide. The ETEC pathotypes in pigs are characterized by the expression of specific fimbrial adhesins, which mediate bacterial colonization of the gut mucosal surface. The most commonly detected types of fimbriae are F4 (previously termed K88) and F18 (F107, 2134P and 8813). Both include different antigenic variants; three for F4 (ab, ac and ad) with F4ac being the most prevalent, and two types for F18 (ab and ac), with F18ac being the one associated with PWD (6, 7). Intestinal adhesion and subsequent colonization by ETEC depends on F4 or F18 specific receptors, the presence of which therefore is essential for ETEC to cause disease (8). ETEC F4 is usually related to PWD of recently weaned piglets occurring 2-3 days after weaning (classical PWD), while F18 is commonly found associated to diarrhea 2-6 weeks after weaning. The age-dependent expression of F4 and F18 receptors in the small intestine might explain why ETEC F4 infection mainly takes place right after weaning as well as during the neonatal period, while ETEC F18 infection mainly occurs later in the post-weaning period (5).

Once ETEC bacteria have adhered and colonized the small intestine, they can produce enterotoxin(s) leading to diarrhea. Both ETEC F4 and F18 are reported to encode two classes of

77 enterotoxins; heat labile (LT) and heat stable (Sta, Stb, and EAST1 -enteroaggregative heat-stable
78 toxin 1) enterotoxins, which induce secretory diarrhea in the pigs (1, 2, 9-11). The predominant
79 serogroup of ETEC associated with classical PWD in pigs worldwide is O149-F4 (1).

80 In Denmark, and other countries, zinc oxide (ZnO) in therapeutic concentrations has been used
81 during the last decades to prevent PWD in the first 14 days after weaning. ZnO has been found to
82 improve growth performance and reduce scours (ETEC induced) in weaning piglets (12). Moreover,
83 ZnO reduces bacterial adhesion and inflammatory cytokine expression, and prevents the disruption
84 of membrane integrity caused by ETEC (13). However, due to the environmental toxicity and
85 potential co-selection for AMR, use of ZnO has been banned in pig production in EU by 2022
86 (https://www.ema.europa.eu/en/medicines/veterinary/referrals/zinc-oxide). Other strategies that
87 haven been used to tackle PWD caused by ETEC include a breeding program (DanBred) which has
88 been implemented in some Danish farms since 2003 and vaccines. The former consists in breeding
89 pigs which do not express the F4ac specific receptors on the intestinal mucosa aiming at reducing
90 occurrence of diarrhea due to ETEC F4 (5). The live attenuated vaccine Coliprotec® F4/F18
91 (https://www.ema.europa.eu/en/medicines/veterinary/EPAR/coliprotec-f4f18) has been developed
92 to diminish the incidence of PWD caused by both ETEC F4 and F18 bacteria. Besides these
93 preventive strategies, neomycin, apramycin, spectinomycin, tetracycline, amoxicillin and
94 sulphadiazine-trimethoprim are the antimicrobials commonly used to treat PWD in Denmark (14).

95 In the present work, we characterized a collection of Danish ETEC F4 and F18 positive strains
96 through whole genome sequencing in order to analyze the pathogenic potential of the strains
97 through identification of relevant virulence factors and to determine the occurrence of AMR and
98 biocide and metal resistances. Further, the analysis allowed us to understand the population genetic
99 structure and variation among strains that are associated with PWD. Moreover, we also analyzed the

100 phylogenetic relationship between the Danish strains under study and swine ETEC strains from
101 other countries of the world.

102 **RESULTS**

103 *Phylogroups, sequence types, clonotypes and serotypes*

104 Most of the ETEC isolates belonged to the phylogroup A (74 isolates, 82.2 %), and the remaining
105 were assigned to four different phylogroups; B1 (one isolate, 1.1 %), C (nine isolates, 10 %), D
106 (five isolates, 5.5 %) and E (one isolate, 1.1 %). The strains displayed eight different MLST-types
107 (STs) (ST10, ST23, ST42, ST48, ST90, ST100, ST155 and ST3524), with ST100 accounting for 56
108 (62.2 %) isolates. Based on *fumC-fimH* allele combinations (CH), eight clonotypes were identified,
109 with CH27-0 as the most prevalent type (56 isolates, 62.2 %), corresponding to strains assigned to
110 ST100 (Supplementary Table S2).

111 The SerotypeFinder tool detected six O serogroups (O6, O8, O29, O141, O147 and O149) and eight
112 different H antigens (H4, H10, H12, H14, H16, H17, H19 and H30). It was not able to predict the O
113 and H antigen in 14 and 8 isolates, respectively. Overall, 13 O:H combinations (serotypes) were
114 found (Supplementary Table S2), with O149:H10 being the most common serotype identified (49
115 isolates, 54.4 %) followed by O6:H16 (11 isolates, 12.2 %).

116 The association between clonal groups (defined by serotype, phylogroup, ST and clonotype) and
117 fimbrial type detected among the strains is shown in Table 1. Two main clonal groups were
118 identified among the 69 F4 positive isolates, O149/ONT:H10/HNT-A-ST100-CH27-0 and O6:H16-
119 A-ST48-CH11-34, representing 81.2 % (56 strains) and 15.9 % (11 strains) of the isolates,
120 respectively. The remaining F4 positive isolates belonged to the O8:H19-C-ST90-CH4-54 clonal
121 group (two isolates, 2.9 %) (Supplementary Table S2). Interestingly, five out of the six F4 positive
122 isolates isolated in the early 90 ´s also belonged to the ONT:H10-A-ST100-CH27-0 clonal group

123 shown for most of the F4 isolates collected during the period 2018-2019 (Supplementary Table S2).
124 In contrast to the homogeneity observed among F4 isolates, the 21 F18 positive isolates showed a
125 higher diversity, and were assigned to six different clonal groups, of which O141/ONT:H4-A-ST10-
126 CH11-24 (33.3 %; seven isolates) and O8:H17-C-ST23-CH4-54 (28.6 %; six isolates) were the
127 predominant ones. The remaining F18 positive isolates were assigned to the O147/ONT:H14-D-
128 ST42-CH-28-65 (five isolates, 23.8 %), O29:H12-B1-ST155-CH4-121, ONT:H19-C-ST90-CH4-0
129 (*E. coli* Nysø) and O8:H31-E-ST3524-CH23-31 (one isolate each, 4.7 %) clonal groups. F18
130 positive strains from the same herd belonged to the same clonal group including seven, six and five
131 isolates from herd D, herd B, and herd C, respectively (Supplementary Table S2).

132 *Phylogeny analysis of ETEC isolates*

133 The raw-read mapping of all 90 genomes to the reference *E. coli* K12 genome showed that
134 3,475,685 out of 4,641,652 (74.8 %) nucleotide positions in the reference genome were present in
135 all the analyzed genomes. A total of 42,172 variable nucleotide positions were detected in this core
136 genome (Supplementary Table S3).

137 Isolates clustered into four major clades (I, II, III and IV) and grouped according to their ST (Figure
138 1). Clade I included seven ST10-F18 isolates recovered from the same herd and showed 1 to 14
139 SNP differences. Isolates belonging to ST48 and encoding F4 fimbriae (eleven isolates) grouped in
140 clade II and showed from 6 (between isolates from the same herd) to 633 SNP differences. Clade III
141 encompassed isolates assigned to different STs (ST155, ST23, ST42, ST3524 and ST90) and all
142 except two ST90 isolates were F18 positive. This cluster was divided into two well-defined
143 subclades: subclade A (including five ST42 isolates, from the same herd with SNP differences
144 ranging from 4 to 18 and one ST3524 isolate) and subclade B containing six ST23 isolates (SNP
145 differences between 3 and 21), three ST90 (two of them were F4 positive) strains (SNPs differences
146 between 104 and 556) and a single ST155 isolate. Lastly, clade IV consisted of F4 positive isolates

147 belonging to ST100 (56 isolates) and was split into two subclades, with the number of different
148 SNPs spanning from 16 to 866 between all genomes and from 4 to 15 between isolates belonging to
149 the same herd. SNP differences between the five F4- ST100 isolates from the 90´s and F4-ST100
150 strains recovered during 2018-2019 ranged from 130 to 699.

151 We investigated the relationship between our strains and ETEC isolates from other countries
152 (Figure 2, Supplementary Table 11). Phylogenetic analysis based on SNP, indicated that the Danish
153 ST100 isolates are closely related to ST100 isolates from China (SNP differences between 120 and
154 443), Spain (178 to 708) and USA (196 to 466). The lowest number of SNP differences among
155 ST10 isolates were detected versus ST10 strains from USA (300 and 301), followed by Spain (1779
156 to 1852) and China (2023 to 2025). Regarding ST48, the single isolate from China included for
157 comparisons showed SNP differences between 4163 and 4291 with regards to the Danish ST48
158 isolates. Finally, our ST90 strains were more similar to the ST90 isolate from China (SNP
159 differences between 80 and 268) than to those ST90 strains from USA (SNP differences ranging
160 from 388 to 476).

161 *Antimicrobial resistance phenotypes and genotypes*

162 The antimicrobial susceptibility testing revealed that 85 (94.4 %) of the isolates were resistant to at
163 least one of the antimicrobials investigated, 60 (66.7 %) were MDR and only five were susceptible
164 to all antimicrobials. High rates of resistance were found against streptomycin (68.9 % of the
165 isolates), sulfamethoxazole (67.8 %), tetracycline (56.7 %), spectinomycin (55.6 %), trimethoprim
166 (53.3 %) and ampicillin (48.3 %). Importantly, none of the isolates was resistant to ceftiofur,
167 cefotaxime, colistin or ciprofloxacin, which are considered highly critical drugs in human medicine
168 (15) (Table 2). In addition, one and eight F4 positive isolates showed resistance to

amoxicillin/clavulanic acid and nalidixic acid, respectively, while none of the F18 isolates tested positive for these drugs (Table 2).

Results from Kappa statistical analysis showed that there was an almost perfect agreement between phenotypic resistance and the *in silico* prediction of resistance genotype (Table 3), which identified a total of 39 different AMR genes (Supplementary Table S2). The genes *bla*$_{\text{TEM-1B}}$, *tet(A)* and *dfrA1* were the most commonly detected among ampicillin, tetracycline and trimethoprim resistant isolates, respectively. Similarly, *sul1* and *sul2* were the predominant genes responsible for sulphonamide resistance. A total of 16 different genes encoding for aminoglycoside modifying enzymes were identified, with the *aph* (phosphotransferases: *aph(3')-Ia*, *aph(3')-Ib*, *aph(3'')-Ib*, *aph(4)-Ia*, *aph(6)-Id*) and *aadA* (nucleotidyltranferases: *aadA1*, *aadA2*, *aadA5*, *aadA11*, *aadA12*, *aadA17*, *aadA22*, *aadA24*) genes as the most common, consistent with the high resistance to spectinomycin and streptomycin, respectively. Regarding phenicols, *catA1*, *cmlA1* and *floR* genes were found among the resistant isolates (Table 3, Supplementary Table S2). Susceptibility to macrolides and lincosamides was not phenotypically tested, however, genes conferring resistance to both classes of drugs were detected (Table 3, Supplementary Table S2).

Notably, none of the isolates was found to harbor genes that suggested extended-spectrum beta-lactamase (ESBL) production or transferrable colistin resistance (*mcr*-class genes).

The ResFinder bioinformatics tool also allows the identification of chromosomal mutations related with antimicrobial resistance. Eight isolates, phenotypically resistant to nalidixic acid, showed a single chromosomal mutation in the *gyrA* (S83L) gene. The substitution V161G, associated with colistin resistance, was also detected in the *pmrB* gene in five isolates; however, these isolates were phenotypically susceptible to the antimicrobial. In addition, one isolate had a nucleotide change in

191 the *ampC* promoter (32T>A), which is associated with resistance to beta-lactams (Supplementary
192 Table S2).

193 *Prediction of Biocide- and metal-tolerance genes*

194 The ETEC isolates showed a large variety of biocide resistance and metal tolerance genes (Figure 3,
195 Supplementary Table S4). A total of 117 different genes were detected, and the number of genes per
196 isolate ranged from 40 to 88. Eleven genes associated with metal tolerance and/or related with metal
197 transport and metabolism as well as biocide resistance were shared by all the strains in the
198 collection: *corA* (magnesium (Mg), cobalt (Co), nickel (Ni), manganese (Mn)), *glpF* (antimony
199 (Sb), arsenic (As), *mgtA* (Co, Mg), *zntR/yhdM* (zinc (Zn)), *phoB* (benzalkonium chloride,
200 chlorhexidine), *mdtF/yhiV* , *mdtK/ydhE* (several biocides including benzalkonium chloride and
201 ethidium bromide), *ostA/lptD* (N-hexane), *emrB* (phenylmercury acetate, 2-chlorophenylhydrazine,
202 carbonylcyanide m-chlorophenyl hydrazine), *acrB* and *acrF/envD* (acriflavine). In addition,
203 different genes conferring metal tolerance to cadmium (Cd), silver (Ag), mercury (Hg), lead (Pb),
204 copper (Cu) and tellurium (Te) were predicted in each strain.

205 The highest diversity of genes conferring biocide resistance and metal tolerance were identified for:
206 copper (16 genes), hydrogen peroxide (15 genes), ethidium bromide (13 genes), sodium
207 deoxycholate (13 genes), cadmium (12 genes), acriflavine (11 genes) and zinc (11 genes).

208 Genes involved in uptake and transport of metals are ubiquitously found in bacteria since metal ions
209 play a role in many different biological processes and are essential for the bacterial survival (16).
210 All genes detected in the ETEC isolates were predicted in *E. coli* K12 with the exception of those
211 presumably conferring metal tolerate to Co (*cuiD, pcoB, pcoS*), Hg (*merA, merE, merR*), quaternary
212 ammonium compounds (*qacE, qacEdelta1, qacF*), Ag (*silA; silE; silP*), Te (*terB, terC, terD, terW,*
213 *terZ*), which were exclusively predicted in ETEC strains (data not shown).

Virulence genes carried by ETEC isolates

The VirulenceFinder bioinformatics tool corroborated the presence of F4 and F18 fimbrial genes in 69 and 21 *E. coli* isolates, respectively, and further identified them as F4ac (100 %) and F18ac (100 %). Besides, two isolates were predicted to carry F6 type fimbria together with F4. Whole genome sequencing analysis also confirmed the presence of *sta* (21 isolates, 23.3 %) and *stb* (85 isolates, 94.4 %) heat-stabile enterotoxins in those isolates positive for the toxins by PCR.

Apart from the fimbrial F4 and F18 and enterotoxin *sta* and *stb* genes, a total of 16 different virulence genes were predicted among the 90 *E. coli* isolates, such as *ltcA*, the gene associated to the heat-labile enterotoxin and present in most of the isolates (82 isolates, 91.1 %). Besides, the gene *astA* (encoding the EAST-1 heat-stable toxin) and the gene *iha* (encoding for a non-hemagglutinating adhesion protein (17)), were also present in 85 % of the isolates, each (Table 4). All isolates carried at least five different virulence genes and 60 % of the isolates harbored eight or more virulence determinants.

Notably, statistical significant differences were detected in the distribution of virulence genes depending on the fimbria type (Table 4). While all F4 positive isolates harbored *astA*, *ltcA* and *stb* genes, only a portion of F18 positive strains carried these genes (38 %, 61 % and 76 %, respectively). Besides, the genes *gad* (encoding a glutamate decarboxylase; enabling resistance to gastric acidity) (18) (35 isolates), *capU* (hexosyltransferase) (54 isolates) as well as the *cba* (33 isolates), *cma* (42 isolates) and *celb* (eight isolates) colicin-encoding genes (bacteriocins active against closely related *E. coli* bacteria or other members of the Enterobacteriaceae) (19) were also more frequently detected among F4 than among F18 positive strains. However, the genes *lpfA* (long polar fimbriae) (13 isolates) and *iss* (increased serum survival; role in extraintestinal pathogenic *E. coli* (ExPEC) virulence (20) (19 isolates) were predominantly found in F18 positive isolates. In

addition, microcin encoding genes (nine isolates) and *sepA* (*Shigella* extracellular protein A, involved in tissue invasion) (ten isolates) (21) were exclusively identified in F4 positive strains while *air* (enteroaggregative immunoglobulin protein with epithelial adhesion function) (four isolates) and *eilA* (*Salmonella* HilA homolog; transcriptional activator of SPI-1 genes) (five isolates) (22), were unique to F18 positive isolates. Hybrid variants VTEC (vero-toxin producing *E. coli*)/ETEC have not been identified in the present work.

In silico plasmid detection, and prediction of co-occurrence of AMR genes with virulence and metal tolerance genes

21 different plasmid replicons were predicted among the *E. coli* strains. The most prevalent replicon types were IncFII (92.2 % of the strains) and IncFIB (91.1 %) followed by IncI1 (72.2 %), IncFIC (54.4 %), and IncX1 (36.7 %) and the number of plasmid replicons per isolate ranged from three to eight (Table 5). At least one IncF replicon was detected in all the isolates.

With regard to fimbrial type, all F4 positive isolates contained replicons of both IncFII and IncFIB types, while all F18 positive strains harbored replicons of the type IncX1 and all but one carried IncI1 (95.2 %). The latter was found in 65.2 % of the F4 positive strains and the IncI2 type was exclusively detected in F4 isolates. Also, the replicon IncX1 was found in all F18 isolates and only in 12 F4 isolates while the replicon IncX4 was unique to three F4 isolates. In addition, IncHI, IncN and IncQ replicons were only found in F4 positive isolates (Table 5).

According to the plasmidSPAdes and ResFinder bioinformatics tools, the majority of the AMR genes detected in the ETEC genomes were predicted to be plasmid located. Specifically, bla_{TEM}, *floR, catA, cmlA,* and *sul3* genes were always detected in plasmid contigs in all isolates. Contrary, *mdf(A)* was predicted to be a chromosomally encoded gene, since it was not detected in any plasmid component (Supplementary Table S5).

Similarly, toxins were predicted to be located on plasmids in most of the strains: *sta* (71.4 %, 15 out of 21 positive isolates), *stb* (83.5 %, 71 out of 85 positive isolates), *ltcA* (87.8 %, 72 out of 82 positive isolates) and *astA* (89.6 %, 69 out of 77 positive isolates). Besides, in 58 out of the 90 genomes, the genes *astA* and *stb* were found in the same contigs, predicted to be a plasmid, and 22 of these also contained the gene *ltcA* in the same plasmid contig. Both F4 and F18 fimbriae genes appeared to be plasmid associated in all genomes too (Supplementary Table S6). Interestingly, where both AMR and virulence genes were predicted to be plasmid encoded in a strain, these two types of genes were identified in the same plasmid component, although not in the same plasmid contig.

Results from BacMet bioinformatics analysis were analyzed together with the plasmid bioformatic tool; *pitA* and *zintA/yodA* genes, conferring Zn tolerance, were predicted to be plasmid located in four and 16 out of 32 and 65 ETEC isolates that harbored these genes. This analysis also predicted that genes related to Cu tolerance were putatively plasmid located in some cases: *cuiD* in one out of six strains, *cueO* in one out of 27 strains, *pcoS* in four out of six strains, and *pcoB* in 34 out of 53 strains. Despite all these Zn and Cu tolerance genes were predicted in the same plasmid component as certain AMR genes, co-occurrence in the same plasmid contig was not detected (Supplementary Table S7). Additionally, most of the *sil*-like genes, *mer*-like genes and *ter*-like genes encoding tolerance to Ag, Hg and Te, respectively, were also predicted as plasmid-located and were also identified in the same plasmid component than several AMR genes (Supplementary Table S8).

Prediction of prophage sequences in ETEC strains

The *in silico* analysis of the 90 ETEC genomes with PHASTER resulted in the prediction of a minimum of 7 to a maximum of 20 different prophage regions in each strain (Supplementary Table S9). The genome sizes of these prophage regions spanned from 2.7 to 107.9 kb, with a GC content

283 ranging from 40.45 to 57.46 %. This is very similar to the GC content of *E. coli* that it is
284 approximatively 50 % (23). Of the total predicted prophages, 39.1 % were intact, whereas 48.2 %
285 were incomplete and the remaining 12.6 % questionable (Supplementary Figure 1). The genome
286 size, the GC content and the number of prophage regions identified did not statistically differ
287 between F4 and F18 strains. Most of the prophage regions identified in the ETEC genomes showed
288 similarity to P88 (NC_026014) and PhiP27 (NC_003356) Enterobacteriaceae prophages (from 3.9
289 % to 93.3 % and from 4.8 % to 60.9 %, respectively). P88 is an inducible prophage of the *E. coli*
290 strain K88, able to lyse avian pathogenic *E. coli* strains (24) and the coliphage PhiP27 is a Stx2e-
291 encoding phage, but since none of the ETEC isolates encoded for Shiga toxins and that the highest
292 similarity is only 60.9 %, the ETEC prophages could be part of the gene pool producing PhiP27 by
293 recombination (25, 26).

294 **DISCUSSION**

295 Characterization of ETEC causing PWD is important for understanding the disease epidemiology
296 and pathogenesis, and for implementation of new strategies aiming to reduce the impact of the
297 disease in the pig industry. In this study, we characterized *in silico* a collection of 90 ETEC F4 and
298 F18 positive isolates in order to determine the levels of antimicrobial and biocide resistance, metal
299 tolerance and virulence genes associated with PWD. Antimicrobial resistance phenotype was also
300 investigated. We also determined how many clonal groups are circulating in the production system
301 as well as the phylogenetic relationships between ETEC strains currently causing PWD in Denmark
302 and in other countries.

303 Overall, ETEC isolates showed significant different genetic traits depending on the fimbria type. It
304 should be mentioned that while 69 F4 positive strains were recovered from at least 30 different
305 herds, the 21 F18 positive isolates were collected from just five herds.

306 ETEC isolates causing PWD are commonly reported to belong to the serogroups O8, O138, O139
307 (often associated with oedema disease), O141, 147, O149, and O157 (1, 27, 28), of which O8,
308 O141, O147 and O149 were predicted among strains in the current investigation, together with O29
309 and O6, both common recognized serogroups associated to ETEC and enteroinvasive *E. coli* (EIEC)
310 in humans (29). Differences were observed concerning serogroups between F4 and F18 positive
311 isolates. O149 was the predominant serogroup among F4 isolates (72.46 % vs 0 %, $p<0.0001$),
312 while O8 was the most prevalent among F18 isolates (2.9 % vs 33.33 %, $p<0.0001$). O149 and
313 O138 (not detected here) have been the most frequent serogroups related to the classical ETEC F4
314 and F18, respectively, in Denmark (28) and other countries (30-32). As reported in Australia,
315 United States, Canada, Germany and Thailand (32, 33), ST100 was the most frequent ST among the
316 Danish porcine ETEC isolates, and ST48 and ST10 were also commonly observed. The latter STs
317 belong to the largest clonal complex 10 within *E. coli*, with isolates from both animal and human
318 sources, acting as commensals or as pathogens, and generally associated with antimicrobial
319 susceptibility and low virulence (34). ST10 is dominant in Spain among *mcr*-positive isolates and
320 China (35-37), followed by ST48 in China (37).

321 In order to study the genetic relatedness of the collected *E. coli* isolates, phylogenetic analysis based
322 on SNPs was performed. As expected, the isolates clustered depending on their ST in four well
323 defined clades, with three of them exclusively encompassing a single ST, while clade III was more
324 diverse, including isolates belonging to different STs. Isolates recovered from the same herd
325 showed the lowest number of different SNPs, suggesting that the same (or very similar) ETEC
326 strain circulates in a specific farm. F4 positive isolates appeared to be more closely related than F18
327 positive strains, since the maximum number of SNPs detected was lower in this group than among
328 F18 strains (24,593 vs 45,742 SNP) and the diversity of STs was lower among F4 isolates
329 compared to F18 strains (three vs six). Results obtained from the phylogeny comparison including

330 isolates from four different countries suggest that the ETEC ST100 isolates are closely related,
331 since the SNP differences were relatively low among them. Interestingly, a relatively low number of
332 SNPs (130-699) were identified between the F4 positive isolates from the 90´s and those recovered
333 during 2018-2019, indicating that very similar clones have been circulating for at least 30 years. To
334 confirm this hypothesis a larger number of F4 strains from previous years (from the 90 ´s up to
335 nowadays) should be analyzed. The clone O6:H16-ST48, which included eleven isolates recovered
336 from five different herds and differing in 6 to 633 SNPs, was the second most prevalent among F4
337 positive isolates. Notably, O6 is one of the most important ETEC serogroups involved in human
338 diarrhea globally, particularly among children under the age of five in developing countries (38),
339 and it is not classically associated with PWD. A recent study based on the genomic characterization
340 of 40 ETEC O6:H16/HNT human isolates collected during 1975–2016 showed significant genomic
341 diversity among them, but none was assigned to ST48 (38). The occurrence of ETEC O6:H16
342 among pig ETEC isolates may indicate a zoonotic potential, however, the strains did not harbor
343 known fimbriae genes involved in adhesion to the human intestine.

344 The World Health Organization has defined AMR a global health issue in both humans and
345 animals, and has recommended surveillance for AMR bacteria in food-producing animals, such as
346 pigs, as they represent a possible source and disseminator of AMR to humans (39). ETEC from pigs
347 are not considered zoonotic (key virulence factors required to cause disease differ between pigs and
348 human (40), but treatment with antimicrobials against PWD may select for AMR in commensal
349 intestinal bacteria, and such bacteria may transfer critical resistances to humans via the food chain.
350 To investigate this aspect, we determined both, phenotypically and *in silico*, the AMR levels among
351 the ETEC isolates. The highest rates of AMR were found against aminoglycosides and
352 sulfamethoxazole, the use of which has increased in livestock in Denmark during last years (41).
353 Resistance to tetracycline, trimethoprim and ampicillin, which are among the frequent

antimicrobials used for treatment of PWD in Denmark, were broadly detected. Notably, levels of tetracycline and ampicillin resistance were similar to those detected in a previous study on pathogenic *E. coli* from pigs in Denmark (42). All the isolates were susceptible to ciprofloxacin, cefotaxime, ceftiofur and colistin, which are critically important antimicrobials for human medicine, (15) but treatment with others antimicrobial agents may allow their co-selection. The absence of these resistances is possibly linked to the restricted use of these drug classes in the pig industry, where fluoroquinolones, cephalosporins and colistin all bear a penalty of factor 10 in the herd level registration scheme of use of antimicrobials (the Yellow card scheme) in the Danish pig production (41). Similar findings were described in a recent study in Denmark, where the highest proportions of AMR among ETEC isolates were found for ampicillin (60.7 %), sulphamethoxazole (69.7 %), tetracycline (47.2 %) and trimethoprim (69.7 %), while AMR to ciprofloxacin, ceftiofur and colistin were not detected (43). MDR in the pig industry has been linked with the wide use of aminoglycosides and beta-lactams in veterinary medicine (44). Here, MDR was detected in 60 isolates (66.7 %), far from the 94 % detected among ETEC isolates carrying *mcr-1* and causing PWD in Spain (36). It is to highlight that *in silico* prediction of AMR genes showed an almost perfect agreement with the phenotypic analysis according to Kappa statistical analysis. In addition, genes encoding for macrolides (specifically for erythromycin via MdfA) or lincosamides resistance, which were not phenotypically tested, were also predicted. Both drug classes are commonly employed in Denmark and other countries for treatment against *Lawsonia intracellularis*, an intracellular pathogen causing enteric disease in pigs (45-47). The increase in the administration of macrolides during last years and the steady use of lincosamides in the pig industry in Denmark (41) could have selected for resistance to the drugs in *E. coli*.

Studies of ETEC across several countries worldwide, including old studies from Denmark, describe ETEC F4 as the most common type associated to PWD, followed by F18 (48-52). However, in

other countries such as Poland, Cuba, Japan and Spain (36, 53-55), the highest prevalence was found for ETEC-F18. In a recent study in Denmark, the number of F4 and F18 strains detected were similar (annual report from 2018; https://diagnostik.dtu.dk/raadgivning/aarsrapporter-for-diagnostik_overvaagning/aarsrapporter-svin). As demonstrated in previous studies, the 90 strains under study were all F4ac or F18ac (6, 7). Denmark has further had a breeding strategy to reduce susceptibility of pigs against ETEC F4ac. The strategy consisted on the inactivation (based on one SNP change) of the candidate gene of the F4ac receptor *MUC4* (5, 56), however, according to our results, F4 strains type ac are still recovered. Since information on the farms is confidential, we acknowledge if the herds under study have this strategy implemented and/or piglets were vaccinated. Thus the reason why F4ac is still been detected could be that some of the isolates tested were recovered from herds where the strategy and/or vaccination has not been applied. Besides, some studies suggest that *MUC13* and not *MUC4* is the most likely gene governing susceptibility to ETEC F4ac, and this might explain also why ETEC F4ac is still the predominant causative agent of PWD right after weaning (57).

In two F4 isolates, F4 and F6 fimbriae genes were detected concurrently. ETEC isolates encoding for more than one fimbriae have been previously described (28, 48, 49, 58, 59) and such strains have been suggested to have a pathogenetic advantage (27). The most prevalent enterotoxin detected in our study was STb (85 isolates) consistent with previous studies performed in other countries (28, 48, 52). All F4-positive isolates carried *astA*, *ltcA* and *stb* genes, in line with results of previous studies, where the F4 gene was strongly associated with *lt* and *stb* (28, 48, 60). However, in our study, only a proportion of F18 positive isolates (n=8) carried *astA*, *ltcA* and *stb* genes, while the remaining strains harbored *ltcA* and/or *stb* and *sta* genes.

Additionally, 16 other virulence genes were predicted among strains, and their distribution was associated with the type of fimbria, suggesting that F4 and F18 positive isolates may use different

virulent strategies to cause disease. As suggested, ETEC strains harboring additional fimbriae adhesins, could potentially exploit other alternative pathways for colonization of the host (17). Overall, the analysis showed that ETEC from Danish pigs harbor other virulence factors besides their characteristic adhesins and toxins, but currently the role for such factors, if any, in intestinal disease is not known.

Plasmids play an important role in the spread and dissemination of both AMR and virulence genes (61). IncF, IncI and IncX replicon types were the most prevalent and at least one IncF replicon was detected in all the strains, as previously reported in porcine ETEC isolates from Australia and Spain (32, 62). In general, F4 isolates showed more plasmid replicon diversity (21 different plasmid replicons) than the F18 isolates (eight different plasmid replicons). IncF plasmid type is the most commonly described in bacteria from humans and animals, particularly in *E. coli*, and is known to carry virulence and AMR genes (63, 64). IncI1 plasmids, the most dominant among F18 strains, were also detected in all O141-F18 isolates and in 87.1% of O149:F4 isolates from Australia (32). IncI1 type plasmids are often associated with AMR and known to be ESBL carriers (65), however, in the current study, ESBL were not detected, and the link to AMR was not investigated. IncX plasmids are narrow host range plasmids of Enterobactericeae, which are known to provide additional advantages, commonly associated with AMR and biofilm formation (64, 66). Since ETEC fimbriae and toxin genes have been reported often to be plasmid located (67), we investigated the putative plasmid localization of these genes and potential co-occurrence with AMR determinants. F4 and F18 fimbria encoding genes were found to be plasmid located in all the strains), toxins were plasmid located in more than 70% of the isolates, and in addition, the majority of the AMR genes were predicted to be plasmid encoded. Interestingly, AMR and virulence genes were often predicted to be part of the same plasmid component, however, further detailed studies

425 are needed to confirm this, since the program used for predictions (plasmidSPAdes) do not separate
426 plasmids of the same type present in a single strain.

427 Recent studies have reported that AMR might be associated with tolerance to heavy metals existing
428 naturally or used in food animal production, such as zinc oxide and copper (41, 68-70). Here, we
429 predicted that aminoglycoside, tetracycline and ampicillin resistance genes and zinc or copper
430 resistance genes were located on the same plasmid component, however, as mentioned above for
431 AMR and virulence genes, more detailed studies are needed to confirm this. Co-localization would
432 imply that use of ZnO and Cu may co-select for AMR. In addition to heavy metals, biocides,
433 including disinfectants and antiseptics, widely used in farms, could also promote the spread of
434 AMR (68). All isolates analyzed here were predicted to carry genes responsible for biocidal
435 resistance, but co-occurrence with AMR was not investigated.

436 Each of the 90 isolates under study contained at least seven prophages, mostly similar to the
437 coliphages P88 (from 3.9% to 93.3%) and Phi27 (from 4.8% to 60.9%). Despite Phi27 is known to
438 encode Stx2e, none of the ETEC isolates encoded for Shiga toxins, thus indicating that, despite the
439 similarity, it is not the same phage.

440 In conclusion, the current study showed a high clonal diversity among F18 isolates, while, in
441 contrast, similar F4 clonal groups might be circulating in Danish herds. Besides, high rates of AMR
442 against aminoglycosides, sulfamethoxazole, tetracycline, trimethoprim and ampicillin as well as
443 high diversity of virulence genes were detected -including toxin genes (*ltcA*, *astA*, *sta*, *stb*), and
444 fimbriae encoding genes typed as F4ac and F18ac.

445 **MATERIALS AND METHODS**

446 *Bacterial strains and PCR detection of fimbriae types F4 and F18*

447 ETEC F4 isolates were collected from pigs with diarrhea in 2018 (n= 34, from 30 different farms),
448 2019 (n=29, from 29 different farms), and 1989-1992 (n=6, from six different farms). Presumptive
449 ETEC F18 isolates (n=20) were recovered from five different farms (collected at the same time
450 point in each farm) in 2019. In addition, *E. coli* Nysø, a well characterized ETEC strain recovered
451 in the 70´s (71), was included as F18 historical control. Strains were obtained during routine
452 diagnostic procedures and use for research purposes did not require ethical clearance, as long as
453 farm identity was not disclosed.

454 Strains were confirmed positive for F4 or F18 fimbriae using PCR with primers and conditions as
455 previously reported (52).

456 *Antimicrobial resistance phenotype*

457 MIC values for *E. coli* isolates were determined for amoxicillin–clavulanic acid (2/1-32/16 µg/ml),
458 ampicillin (1-32 µg/ml), apramycin (4-32 µg/ml), cefotaxime (0.125-4 µg/ml), ceftiofur (0.5-8
459 µg/ml), chloramphenicol (2-64 µg/ml), ciprofloxacin (0.015-4 µg/ml), colistin (1-16 µg/ml),
460 florfenicol (2-64 µg/ml), gentamicin (0.5-16 µg/ml), nalidixic acid (4-64 µg/ml), neomycin (2-32
461 µg/ml), spectinomycin (16-256 µg/ml), streptomycin (8-128 µg/ml), sulphamethoxazole (64-1024
462 µg/ml), tetracycline (2-32 µg/ml) and trimethoprim (1-32 µg/ml) by the broth microdilution method
463 using Sensititre microtiter trays (DKMVN4, Sensititre system; Thermo Fisher Scientific, West
464 Sussex, United Kingdom). *E. coli* ATCC 25922 was used for quality control. Results were
465 interpreted according to EUCAST epidemiological cut-off values, EUCAST clinical breakpoints for
466 amoxicillin-clavulanic acid (www.EUCAST.org) and DANMAP for apramycin (72). Isolates were
467 defined as "susceptible" when classified as "wild type" and "resistant" when classified as "no wild
468 type". Multidrug-resistant (MDR) strains were those resistant to one agent from three or more
469 different antimicrobial classes (73).

PCR for detection of sta and stb toxin genes

E. coli DNA was extracted from a single overnight-grown colony by the boiling lysis method as reported (52), and PCR amplification of *sta* and *stb* toxin genes was performed using the primers and PCR conditions previously described (52).

DNA extraction and whole genome sequencing (WGS)

DNA was extracted using the Maxwell® system (Promega) following the instructions provided by the Maxwell® RSC cultured cells DNA kit (Promega). Quality of the DNA was determined by NanoDrop-1000 (Thermo Fisher Scientific) and DNA quantification was performed using dsDNA BR Assay kit with the Qubit 2.0 fluorometer (Invitrogen, USA).

The libraries for sequencing were prepared using the Nextera DNA Flex Library Prep Kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer´s protocol and sequenced using Illumina NextSeq (Illumina). The paired-end raw reads were assembled using SPAdes Genome Assembler v.3.13.0 (74) and the quality of assembly was evaluated with QUAST v.5.0.2. (75). The raw sequences were submitted to the European Nucleotide Archive (ENA) under the study accession number PRJEB38608.

Whole genome characterization

The assembled contigs, with genomic size ranging between 5.1 and 5.7 Mbp (mean size 5.4 Mbp) (Supplementary Table S1), were analyzed using the bioinformatics tools of the Center for Genomic Epidemiology (CGE) for the presence of antibiotic resistance (ResFinder v.3.2) (76), virulence genes (VirulenceFinder v.2.0) (77), plasmid replicon types (PlasmidFinder v.2.1) (78), and identification of clonotypes (CHTyper v.1.0), sequence types (ST) (MLST v.2.0) (79) and serotypes (SerotypeFinder v.2.0) (75). All the CGE predictions were called using default settings.

492 Identification of antibacterial biocide- and metal- tolerance genes was assessed using BacMet-Scan
493 v.2.0 (80).The ClermonTyping tool (http://clermontyping.iame-research.center/) and PHASTER
494 webserver (http://phaster.ca/) were used to predict the phylogroups and putative prophage
495 sequences in the bacterial genomes, respectively (81, 82).

496 *In silico prediction of localization of antimicrobial resistance, virulence and metal tolerance genes*
497 *and potential co-occurrence on plasmids*

498 The putative localization of AMR and virulence genes (F4, F18, *ltcA*, *astA*, *sta* and *stb*) was
499 predicted using a combination of plasmidSPAdes v.3.13.0 (83), ResFinder and VirulenceFinder
500 tools. Briefly, plasmidSPAdes was used to identify contigs most likely belonging to plasmid DNA
501 and to assign them to components. Each component is considered as a putative plasmid consisting
502 of one or more contigs. This tool is not able to separate similar plasmids (for example similar
503 plasmids of the same type present in a single strain), and thus their contigs may be assigned to the
504 same id (same component). ResFinder and VirulenceFinder were used to analyze the presence of
505 AMR and virulence genes in all contigs identified as putative DNA regions of plasmids. The output
506 from both tools provides the contig ID and component on which the specific genes were located.
507 Genes contained in the same contig and/or component were predicted to be plasmid located, while
508 those antimicrobial and/or virulence genes not detected in plasmid DNA contigs were assumed to
509 be chromosome encoded.

510 The BacMet database (BacMet-Scan v.2.0.), which includes metal tolerance and biocide resistance
511 genes (metal tolerance genes include also those genes that are indirectly related to metals) was used
512 to investigate the presence of these genes in all the genomes. The genome of the *E. coli* K12
513 substrain MG1655 (GenBank accession number NC_000913.3) was also included in the analysis.
514 Plasmid contigs identified and assigned to components with plasmidSPAdes v.3.13.0 as described

above were analyzed with BacMet-Scan v.2.0. Next, the plasmid contigs were manually inspected for the presence of metal tolerance genes (zinc, copper, silver, mercury and tellurium) previously identified on the genome assembly by using the BacMet database (Supplementary Table S4). Co-occurrence of AMR genes (detected as indicated above) and the metal tolerance genes on the same plasmid was predicted when they were found to belong to the same contig.

Phylogenetic analysis

Phylogenetic relationships between the isolates were analyzed based on SNP trees constructed using the bioinformatics tool CSI Phylogeny v.1.4 (84) available at CGE. The genome of the *E. coli* K12 substrain MG1655 was included as a reference strain, and CGE default parameters were used during SNP analysis. The phylogenetic tree was visualized and edited by using the bioinformatics tool iTOL v5 (85).

Besides, the ETEC isolates under study were compared with 20 swine ETEC isolates from three different countries (China, USA and Spain) as mentioned above. The accession numbers and MLST-types of the 20 ETEC strains used in this phylogenetic analysis are indicated in the Supplementary Table S10.

Statistical analysis

Differences between F4 and F18 positive strains regarding serogroups, antimicrobial resistance, virulence gene content and plasmid replicons were analyzed using two-tailed Fisher's exact test with the GraphPad Prism version 8.3 software (GraphPad Inc). P values <0.05 were considered statistically significant.

Cohen's Kappa statistical analysis was used to analyze the correlation between phenotypic resistance and *in silico* gene predictions using SPSS version 26 (IBM, USA). Kappa values ≤ 0 indicate no

agreement; 0.01–0.20 none to slight; 0.21–0.40 fair; 0.41– 0.60 moderate; 0.61–0.80 substantial, and 0.81–1.00 indicate values with an almost perfect agreement (86).

Acknowledgments

The authors are grateful to Joakim Larsson and Johan Bengtsson-Palme for their assistance with the use of the Bac-Met database as well as K. Aagard and M. Carlsen for their technical assistance.

Funding

V. García acknowledges the Consellería de Cultura, Educación e Ordenación Universitaria, Xunta de Galicia for her post-doctoral grant (grant number ED481B-2018/018). This work was supported by Innovationsfonden (grant number 8088-00032B) and the Danish Ministry of Food of Environment (VetforligIII).

Conflicts of interest

All authors declare that they have no competing interests.

Data availability

The draft genome sequences of *E. coli* isolates from pigs in Denmark in this study are available in the European Nucleotide Archive (ENA) under the study accession number PRJEB38608.

References

1. Fairbrother JM, Nadeau E, Gyles CL. *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. 2005. Anim Health Res Rev 6(1):17-39.

2. Luppi A. 2017. Swine enteric colibacillosis: diagnosis, therapy and antimicrobial resistance. Porcine Health Manag 3:16.

558 3. Amezcua R, Friendship RM, Dewey CE, Gyles C, Fairbrother JM. 2002. Presentation
559 of postweaning *Escherichia coli* diarrhea in southern Ontario, prevalence of hemolytic E. coli
560 serogroups involved, and their antimicrobial resistance patterns. Can J Vet Res 66(2):73-8.
561 4. Rhouma M, Fairbrother JM, Beaudry F, Letellier A. 2017. Post weaning diarrhea in
562 pigs: risk factors and non-colistin-based control strategies. Acta Vet Scand 59(1):31.
563 5. Luise D, Lauridsen C, Bosi P, Trevisi P. 2019. Methodology and application of
564 *Escherichia coli* F4 and F18 encoding infection models in post-weaning pigs. J Anim Sci
565 Biotechnol 10:53.
566 6. Westerman RB, Mills KW, Phillips RM, Fortner GW, Greenwood JM. 1988.
567 Predominance of the ac variant in K88-positive *Escherichia coli* isolates from swine. J Clin
568 Microbiol 26(1):149-50.
569 7. Rippinger P, Bertschinger HU, Imberechts H, Nagy B, Sorg I, Stamm M, Wild P,
570 Wittig W. 1995. Designations F18ab and F18ac for the related fimbrial types F107, 2134P and 8813
571 of *Escherichia coli* isolated from porcine postweaning diarrhoea and from oedema disease. Vet
572 Microbiol 45(4):281-95.
573 8. Nagy B, Fekete PZ. 2005. Enterotoxigenic *Escherichia coli* in veterinary medicine. Int
574 J Med Microbiol 295(6-7):443-54.
575 9. Loos M, Geens M, Schauvliege S, Gasthuys F, van der Meulen J, Dubreuil JD,
576 Goddeeris BM, Niewold T, Cox E. 2012. Role of heat-stable enterotoxins in the induction of early
577 immune responses in piglets after infection with enterotoxigenic *Escherichia coli*. PLoS One
578 7(7):e41041.
579 10. Nagy B, Wilson RA, Whittam TS. 1999. Genetic diversity among *Escherichia coli*
580 isolates carrying f18 genes from pigs with porcine postweaning diarrhea and edema disease. J Clin
581 Microbiol 37(5):1642-5.

582 11. Fairbrother J, Gyles C. Colibacillosis. In: Zimmerman J, Karriker L, Ramirez A,
583 Schwartz K, Stevenson G, editors. 2012. Diseases of Swine. Tenth ed: Willey-Blackwell; p. 723-49.
584 12. Li X, Yin J, Li D, Chen X, Zang J, Zhou X. Dietary supplementation with zinc oxide
585 increases Igf-I and Igf-I receptor gene expression in the small intestine of weanling piglets. J Nutr.
586 2006;136(7):1786-91.
587 13. Roselli M, Finamore A, Garaguso I, Britti M, Mengheri E. Zinc oxide protects
588 cultures enterocytes from the damage induced by Escherichia coli The Journal of Nutrition.
589 2003;133(12):4077-82.
590 14. DANMAP. 2017. Use of antimicrobial agents and occurrence of antimicrobial
591 resistance in bacteria from food animals, food and humans in Denmark. The Danish Integrated
592 Antimicrobial Resistance Monitoring and Research Programme, Copenhagen, Denmark..
593 15. WHO. 2019. World Health Organization: Critically important antimicrobials for
594 human medicine – 6th rev. Geneva. Licence: CC BY-NC-SA 3.0 IGO.
595 16. Porcheron G, Garenaux A, Proulx J, Sabri M, Dozois CM. 2013. Iron, copper, zinc,
596 and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains,
597 site of infection and the relative importance of the different metal transport systems for virulence.
598 Front Cell Infect Microbiol ;3:90.
599 17. Wu XY, Chapman T, Trott DJ, Bettelheim K, Do TN, Driesen S,. Walker MJ, Chin J.
600 2007. Comparative analysis of virulence genes, genetic diversity, and phylogeny of commensal and
601 enterotoxigenic *Escherichia coli* isolates from weaned pigs. Appl Environ Microbiol 73(1):83-91.
602 18. Castanie-Cornet MP, Penfound TA, Smith D, Elliott JF, Foster JW. 1999. Control of
603 acid resistance in *Escherichia coli*. J Bacteriol 181(11):3525-35.
604 19. Cascales E, Buchanan SK, Duche D, Kleanthous C, Lloubes R, Postle K, Riley M,
605 Slatin S, Cavard D. 2007. Colicin biology. Microbiol Mol Biol Rev 71(1):158-229.

606 20. Johnson TJ, Wannemuehler YM, Nolan LK. 2008. Evolution of the *iss* gene in
607 *Escherichia coli*. Appl Environ Microbiol 74(8):2360-9.
608 21. Benjelloun-Touimi Z, Si Tahar M, Montecucco C, Sansonetti PJ, Parsot C. 1998.
609 SepA, the 110 kDa protein secreted by *Shigella flexneri:* two-domain structure and proteolytic
610 activity. Microbiology 144 (Pt 7):1815-22.
611 22. Sheikh J, Dudley EG, Sui B, Tamboura B, Suleman A, Nataro JP. 2006. EilA, a HilA-
612 like regulator in enteroaggregative *Escherichia coli*. Mol Microbiol 61(2):338-50.
613 23. Mann S, Chen YP.2010. Bacterial genomic G+C composition-eliciting environmental
614 adaptation. Genomics 95(1):7-15.
615 24. Chen M, Zhang L, Xin S, Yao H, Lu C, Zhang W. 2017. Inducible Prophage Mutant
616 of *Escherichia coli* Can Lyse New Host and the Key Sites of Receptor Recognition Identification.
617 Front Microbiol 8:147.
618 25. Recktenwald J, Schmidt H. 2002. The nucleotide sequence of Shiga toxin (Stx) 2e-
619 encoding phage phiP27 is not related to other Stx phage genomes, but the modular genetic structure
620 is conserved. Infect Immun 70(4):1896-908.
621 26. Smith DL, Rooks DJ, Fogg PC, Darby AC, Thomson NR, McCarthy AJ, Allison HE.
622 2012. Comparative genomics of Shiga toxin encoding bacteriophages. BMC Genomics 13:311.
623 27. Nagy B, Fekete PZ. 1999. Enterotoxigenic *Escherichia coli* (ETEC) in farm animals.
624 Vet Res 30(2-3):259-84.
625 28. Frydendahl K. 2002. Prevalence of serogroups and virulence genes in *Escherichia coli*
626 associated with postweaning diarrhoea and edema disease in pigs and a comparison of diagnostic
627 approaches. Vet Microbiol 85(2):169-82.
628 29. Nataro JP, Kaper JB. 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev
629 11(1):142-201.

630 30. Sting R, Stermann M. 2008. Duplex real-time PCR assays for rapid detection of
631 virulence genes in *E. coli* isolated from post-weaning pigs and calves with diarrhoea. Dtsch
632 Tierarztl Wochenschr 115(6):231-8.
633 31. Barth S, Schwanitz A, Bauerfeind R. 2011. Polymerase chain reaction-based method
634 for the typing of F18 fimbriae and distribution of F18 fimbrial subtypes among porcine Shiga toxin-
635 encoding *Escherichia coli* in Germany. J Vet Diagn Invest 23(3):454-64.
636 32. Abraham S, Trott DJ, Jordan D, Gordon DM, Groves MD, Fairbrother JM, Smith MG,
637 Zhang R, Chapman TA. 2014. Phylogenetic and molecular insights into the evolution of multidrug-
638 resistant porcine enterotoxigenic *Escherichia coli* in Australia. Int J Antimicrob Agents 44(2):105-
639 11.
640 33. Jiang F, Wu Z, Zheng Y, Frana TS, Sahin O, Zhang Q, Li G. 2019. Genotypes and
641 Antimicrobial Susceptibility Profiles of Hemolytic *Escherichia coli* from Diarrheic Piglets.
642 Foodborne Pathog Dis 16(2):94-103.
643 32. Manges AR, Johnson JR. 2012. Food-borne origins of *Escherichia coli* causing
644 extraintestinal infections. Clin Infect Dis 55(5):712-9.
645 35. García V, García-Meniño I, Mora A, Flament-Simon SC, Díaz-Jiménez D, Blanco JE,
646 Alonso MP, Blanco J. 2018. Co-occurrence of mcr-1, mcr-4 and mcr-5 genes in multidrug-resistant
647 ST10 Enterotoxigenic and Shiga toxin-producing *Escherichia coli* in Spain (2006-2017). Int J
648 Antimicrob Agents 52(1):104-108.
649 36. García-Meniño I, García V, Mora A, Diaz-Jiménez D, Flament-Simon SC, Alonso
650 MP, Blanco JE, Blanco M, Blanco J. 2018. Swine Enteric Colibacillosis in Spain: Pathogenic
651 Potential of mcr-1 ST10 and ST131 *E. coli* Isolates. Front Microbiol 9:2659.

652 37. Yang GY, Guo L, Su JH, Zhu YH, Jiao LG, Wang JF. 2019. Frequency of
653 Diarrheagenic Virulence Genes and Characteristics in *Escherichia coli* Isolates from Pigs with
654 Diarrhea in China. Microorganisms 7(9).
655 38. Pattabiraman V, Katz LS, Chen JC, McCullough AE, Trees E. 2018. Genome wide
656 characterization of enterotoxigenic *Escherichia coli* serogroup O6 isolates from multiple outbreaks
657 and sporadic infections from 1975-2016. PLoS One 13(12):e0208735.
658 39. WHO. 2014. World Health Organization: Antimicrobial Resistance Global Report on
659 Surveillance. http://www.who.int/drugresistance/documents/surveillancereport/en/
660 40. Zhang W, Robertson DC, Zhang C, Bai W, Zhao M, Francis DH. Escherichia coli
661 constructs expressing human or porcine enterotoxins induce identical diarrheal diseases in a piglet
662 infection model. Appl Environ Microbiol. 2008;74(18):5832-7.
663 41. DANMAP. 2018. Use of antimicrobial agents and occurrence of antimicrobial
664 resistance in bacteria from food animals, food and humans in Denmark. The Danish Integrated
665 Antimicrobial Resistance Monitoring and Research Programme, Copenhagen, Denmark. 2018.
666 42. Holmer I, Salomonsen CM, Jorsal SE, Astrup LB, Jensen VF, Hog BB, Pedersen K.
667 2019. Antibiotic resistance in porcine pathogenic bacteria and relation to antibiotic usage. BMC Vet
668 Res 15(1):449.
669 43. Rosager WN, Peter NJ, Erik Lind JS, Svend H, Matthew D, Steen PK. 2017.
670 Comparison of antimicrobial resistance in *E. coli* isolated from rectal and floor samples in pens with
671 diarrhoeic nursery pigs in Denmark. Prev Vet Med 147:42-9.
672 44. Rhouma M, Beaudry F, Thériault W, Letellier A. 2016. Colistin in Pig Production:
673 Chemistry, Mechanism of Antibacterial Action, Microbial Resistance Emergence, and One Health
674 Perspectives. Front Microbiol 7:1789.

675 45. Pyorala S, Baptiste KE, Catry B, van Duijkeren E, Greko C, Moreno MA, Pomba MC,
676 Rantala M, Ruzauskas M, Sanders P, Threlfall EJ, Torren-Edo J, Torneke K. 2014. Macrolides and
677 lincosamides in cattle and pigs: use and development of antimicrobial resistance. Vet J 200(2):230-
678 9.
679 46. Karuppannan AK, Opriessnig T. 2018. *Lawsonia intracellularis*: Revisiting the
680 Disease Ecology and Control of This Fastidious Pathogen in Pigs. Front Vet Sci 5:181.
681 47. Arnold M, Crienen A, Swam H, von Berg S, Jolie R, Nathues H. 2019. Prevalence of
682 *Lawsonia intracellularis* in pig herds in different European countries. Porcine Health Manag 5:31.
683 48. Luppi A, Gibellini M, Gin T, Vangroenweghe F, Vandenbroucke V, Bauerfeind R, .
684 Bonilauri P, Labarque G, Hidalgo A. 2016. Prevalence of virulence factors in enterotoxigenic
685 *Escherichia coli* isolated from pigs with post-weaning diarrhoea in Europe. Porcine Health Manag
686 2:20.
687 49. Vu-Khac H, Holoda E, Pilipcinec E. 2004. Distribution of virulence genes in
688 *Escherichia coli* strains isolated from diarrhoeic piglets in the Slovak Republic. J Vet Med B Infect
689 Dis Vet Public Health 51(7):343-7.
690 50. Zajacova ZS, Konstantinova L, Alexa P. 2012. Detection of virulence factors of
691 *Escherichia coli* focused on prevalence of EAST1 toxin in stool of diarrheic and non-diarrheic
692 piglets and presence of adhesion involving virulence factors in astA positive strains. Vet Microbiol
693 154(3-4):369-75.
694 51. Smith MG, Jordan D, Chapman TA, Chin JJ, Barton MD, Do TN, Fahy VA,
695 Fairbrother JM, Trott DJ. 2012. Antimicrobial resistance and virulence gene profiles in multi-drug
696 resistant enterotoxigenic *Escherichia coli* isolated from pigs with post-weaning diarrhoea. Vet
697 Microbiol 145(3-4):299-307.

698 52. Zhang W, Zhao M, Ruesch L, Omot A, Francis D. 2007. Prevalence of virulence
699 genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. Vet
700 Microbiol 123(1-3):145-52.
701 53. Osek J, Gallien P, Truszczyñski M, Protz D. 1999. The use of polymerase chain
702 reaction for determination of virulence factors of *Escherichia coli* strains isolated from pigs in
703 Poland. Comp Immunol Microbiol Infect Dis 22(3):163-74.
704 54. Blanco M, Lazo L, Blanco JE, Dahbi G, Mora A, López C, González EA, Blanco J.
705 2006. Serotypes, virulence genes, and PFGE patterns of enteropathogenic *Escherichia coli* isolated
706 from Cuban pigs with diarrhea. Int Microbiol 9(1):53-60.
707 55. Kusumoto M, Hikoda Y, Fujii Y, Murata M, Miyoshi H, Ogura Y, Gotoh Y, Iwata T,
708 Hayashi T, Akiba M. 2016. Emergence of a Multidrug-Resistant Shiga Toxin-Producing
709 Enterotoxigenic *Escherichia coli* Lineage in Diseased Swine in Japan. J Clin Microbiol 54(4):1074-
710 81.
711 56. Peng QL, Ren J, Yan XM, Huang X, Tang H, Wang YZ, Zhang B, Huang LS. 2007.
712 The g.243A>G Mutation in Intron 17 of MUC4 Is Significantly Associated With
713 Susceptibility/Resistance to ETEC F4ab/ac Infection in Pigs. Anim Genet 38(4):397-400.
714 57. Ren J, Yan X, Ai H, Zhang Z, Huang X, Ouyang J, Yang M, Yang H, Han P, Zeng W,
715 Chen Y, Guo Y, Xiao S, Ding N, Huang L. 2012. Susceptibility towards enterotoxigenic
716 *Escherichia coli* F4ac diarrhea is governed by the MUC13 gene in pigs. PLoS One 7(9):e44573.
717 58. Kwon D, Choi C, Jung T, Chung HK, Kim JP, Bae SS, Cho WS, Kim J, Chae C.
718 2002. Genotypic prevalence of the fimbrial adhesins (F4, F5, F6, F41 and F18) and toxins (LT,
719 STa, STb and STx2e) in *Escherichia coli* isolated from postweaning pigs with diarrhoea or oedema
720 disease in Korea. Vet Rec 150(2):35-7.

721 59. Chen X, Gao S, Jiao X, Liu XF. 2004. Prevalence of serogroups and virulence factors
722 of Escherichia coli strains isolated from pigs with postweaning diarrhoea in eastern China. Vet
723 Microbiol 103(1-2):13-20.
724 60. Post K, Bosworth B, Knoth J. 2000. Frequency of virulence factors in *Escherichia coli*
725 isolated from pigs with postweaning diarrhea and edema disease in North Carolina. Swine Health
726 and Production 8(3):119-20.
727 61. Carattoli A. 2009. Resistance plasmid families in Enterobacteriaceae. Antimicrob
728 Agents Chemother 53(6):2227-38.
729 62. García-Meniño I, Díaz-Jiménez D, García V, de Toro M, Flament-Simon SC, Blanco
730 J, Mora A. 2019. Genomic Characterization of Prevalent mcr-1, mcr-4, and mcr-5 *Escherichia coli*
731 Within Swine Enteric Colibacillosis in Spain. Front Microbiol 10.
732 63. Villa L, Garcia-Fernandez A, Fortini D, Carattoli A. 2010. Replicon sequence typing
733 of IncF plasmids carrying virulence and resistance determinants. J Antimicrob Chemother
734 65(12):2518-29.
735 64. Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn B, Guerra
736 B, Mevius DJ, Hordijk J. 2018. Plasmids carrying antimicrobial resistance genes in
737 Enterobacteriaceae. J Antimicrob Chemother 73(5):1121-1137.
738 65. Wang J, Stephan R, Karczmarczyk M, Yan Q, Hächler H, Fanning S. 2013. Molecular
739 characterization of blaESBL–harboring conjugative plasmids identified in multi-drug resistant
740 *Escherichia coli* isolated from food-producing animals and healthy humans. Front Microbiol 4.
741 66. Johnson TJ, Bielak EM, Fortini D, Hansen LH, Hasman H, Debroy C, Nolan LK,
742 Carattoli A. 2012. Expansion of the IncX plasmid family for improved identification and typing of
743 novel plasmids in drug-resistant Enterobacteriaceae. Plasmid 68(1):43-50.

744 67. Dubreuil JD, Isaacson RE, Schifferli DM. 2016. Animal Enterotoxigenic Escherichia
745 coli. EcoSal Plus 7(1).
746 68. Cheng G, Ning J, Ahmed S, Huang J, Ullah R, An B, Hao H, Dai M, Huang L, Wang
747 X, Yuan Z. 2019. Selection and dissemination of antimicrobial resistance in Agri-food production.
748 Antimicrob Resist Infect Control 8:158.
749 69. Pal C, Asiani K, Arya S, Rensing C, Stekel DJ, Larsson DGJ, Hobman JL. 2017.
750 Metal Resistance and Its Association With Antibiotic Resistance. Adv Microb Physiol 70:261-313.
751 70. Rensing C, Moodley A, Cavaco LM, McDevitt SF. 2018. Resistance to Metals Used
752 in Agricultural Production. Microbiol Spectr 6(2).
753 71. Larsen J. 1981. Effect of pectin on secretion in pig jejunal loops challenged to
754 enteropathogenic *E. coli* or enterotoxin (LT). A preliminary report. Nord Vet Med 33(4-5):218-23.
755 72. DANMAP. 2015. Use of antimicrobial agents and occurrence of antimicrobial
756 resistance in bacteria from food animals, food and humans in Denmark. The Danish Integrated
757 Antimicrobial Resistance Monitoring and Research Programme, Copenhagen, Denmark.
758 73. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG,
759 Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J,
760 Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-
761 resistant and pandrug-resistant bacteria: an international expert proposal for interim standard
762 definitions for acquired resistance. Clin Microbiol Infect 18(3):268-81.
763 74. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,
764 Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev
765 MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-
766 cell sequencing. J Comput Biol 19(5):455-77.

767 75. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool
768 for genome assemblies. Bioinformatics 29(8):1072-5.
769 76. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup
770 FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob
771 Chemother 67(11):2640-4.
772 77. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and
773 Easy In Silico Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. J
774 Clin Microbiol 53(8):2410-26.
775 78. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L,
776 Moller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using
777 PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58(7):3895-
778 903.
779 79. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L,
780 Sicheritz-Ponten T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-
781 genome-sequenced bacteria. J Clin Microbiol 50(4):1355-61.
782 80. Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DG. 2014. BacMet:
783 antibacterial biocide and metal resistance genes database. Nucleic Acids Res ;42.
784 81. Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. 2018.
785 ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain
786 phylotyping. Microb Genom 4(7):e000192.
787 82. Arndt D, Marcu A, Liang Y, Wishart DS. 2019. PHAST, PHASTER and PHASTEST:
788 Tools for finding prophage in bacterial genomes. Brief Bioinform 20(4):1560-7.

789 83. Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016.
790 plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics
791 32(22):3380-7.
792 84. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the problem of
793 comparing whole bacterial genomes across different sequencing platforms. PLoS One
794 9(8):e104984.
795 85. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the
796 display and annotation of phylogenetic and other trees. Nucleic Acids Res 44(W1):W242-5.
797 86. McHugh M. Interrater reliability: the kappa statistic. Biochemia Medica.
798 2012;22(3):276-82.
799

800 **Legends to Figures**

801

802 **Figure 1.** SNP-based phylogeny of the 90 ETEC isolates from pigs. Colours in the outer ring
803 correspond to phylogroups, the middle ring to fimbriae type and the inner ring to Sequence Type.
804

805 **Figure 2.** SNP-based phylogeny of the 90 ETEC isolates from pigs in Denmark and the 20 ETEC
806 isolates from different countries. Colours in the ring correspond to Sequence Type. Isolates
807 highlighted in red, orange and blue correspond to isolates from China, USA and Spain,
808 respectively.

809 **Figure 3.** Distribution of metal tolerance and biocides resistance among the 90 ETEC isolates
810 from pigs (A), the F4-positive isolates (B) and the F18-positive isolates (C).
811

812 **Table 1**. Serotype, phylogroup, Sequence Type and clonotypes association with fimbrial antigens in the 90

813 ETEC isolates from pigs.

Serotype-PG-ST-CH	Nº (%)	Fimbrial antigen
O149:H10-A-ST100-CH27-0	49 (54.4)	F4
ONT:H10-A-ST100-CH27-0	6 (6.7)	F4
O149:HNT-A-ST100-CH27-0	1 (1.1)	F4
O6:H16-A-ST48-CH11-34	11 (12.2)	F4
O8:H19-C-ST90-CH4-54	2 (2.2)	F4
O8:H17-C-ST23-CH4-54	6 (6.7)	F18
ONT:H4-A-ST10-CH11-24	5 (5.6)	F18
O141:H4-A-ST10-CH11-24	2 (2.2)	F18
O147:H14-D-ST42-CH28-65	3 (3.3)	F18
ONT:H14-D-ST42-CH28-65	2 (2.2)	F18
O29:H12-B1-ST155-CH4-121	1 (1.1)	F18
ONT:H19-C-ST90-CH4-0	1 (1.1)	F18
O8:H31-E-ST3524-CH23-31	1 (1.1)	F18

814 PG: phylogroup, ST: MLST-type, CH: clonotype.

815

816 **Table 2.** Prevalence of antimicrobial resistance among the 90 ETEC isolates from pigs.

Antimicrobial[a]	**Number of resistant isolates (%)**	**F4 (%)**[b]	**F18 (%)**[b]	**p-value**	**OR at 95% CI**
Ampicillin	47 (48.3)	35 (50.7)	12 (57.1)	0.6282	
Amoxicillin/Clavulanic acid	1 (1.1)	1 (1.5)	0	>0.9999	
Chloramphenicol/florfenicol	15 (16.7)	9 (13.4)	6 (28.5)	0.1064	
Apramycin	8 (8.9)	7 (10.14)	1 (4.7)	0.6752	
Gentamicin	6 (6.7)	8 (11.6)	1 (4.7)	0.6792	
Neomycin	23 (25.6)	16 (23.2)	7 (33.3)	0.3964	
Spectinomycin	50 (55.6)	35 (50.5)	15 (71.42)	0.1327	
Streptomycin	62 (68.9)	**53 (76.8)**	9 (42.9)	0.0061	4.417 (1.491-12.48)
Sulfamethoxazole	61 (67.8)	46 (66.7)	15 (71.4)	0.7932	
Tetracycline	51 (56.7)	42 (60.9)	9 (42.8)	0.2084	
Trimethoprim	48 (53.3)	35 (50.7)	13 (62)	0.4568	
Nalidixic acid	8 (8.9)	8 (11.6)	0	0.1901	

817 [a]The breakpoints used correspond to EUCAST Epidemiological cut-off values (ECOFFs) for ampicillin (8 µg/ml),
818 chloramphenicol/florfenicol (8 µg/ml), gentamicin (2 µg/ml), nalidixic acid (8 µg/ml), neomycin (8 µg/ml),
819 spectinomycin (64 µg/ml), streptomycin (16 µg/ml), sulphamethoxazole (64 µg/ml), tetracycline (8 µg/ml) and
820 trimethoprim (2 µg/ml), EUCAST clinical breakpoints for amoxicillin/clavulanic acid (R > 8 µg/ml) and DANMAP
821 2015 for apramycin (R >32 µg/ml)
822 [b]Percentage is estimated based on the total number of strains associated to each fimbria type.
823 OR: Odds ratio is indicated when p-value <0.05; CI: confidence interval; ∞: infinity. Significant differences (p-value
824 <0.05) are indicated in bold

825

826

827 **Table 3**. Antimicrobial resistance genes detected among the 90 ETEC isolates from pigs.

Antimicrobial group	Genes	Total (%)	F4 (%)[a]	F18 (%)[a]	kappa	p-value
Beta-lactams	*bla*$_{TEM-1A}$	4 (4.4)	4 (5.8)	0	0.933	0.0001
	bla$_{TEM-1B}$	42 (46.7)	30 (43.5)	12 (57.1)		
	bla$_{TEM-30}$	1 (1.1)	1 (1.4)	0		
Aminoglycosides[b]	*aph* (phosphotransferases)	58 (64.4)	50 (72.5)	8 (38.1)	0.927	0.0001
	aadA (nucleotidyltransferases)	57 (63.3)	42 (60.9)	15 (71.4)		
	aac (acetyltranspherases)	9 (10)	8 (11.6)	1 (4.8)		
Phenicols	*catA1*	3 (3.3)	3 (4.3)	0	1.000	0.0001
	cmlA1	8 (8.9)	2 (2.9)	6 (28.6)		
	floR	5 (5.6)	5 (7.2)	0		
Macrolides	*mdf(A)*	90 (100)	69 (100)	21 (100)	ND	ND
	mph(A)	8 (8.9)	8 (11.6)	0		
	mph(B)	7 (7.8)	7 (10.1)	0		
	erm(B)	9 (10)	9 (13)	0		
Lincosamides	*lnu(F)*	5 (5.6)	5 (7.2)	0	ND	ND
	lnu(G)	5 (5.6)	5 (7.2)	0		
Sulphonamides	*sul1*	30 (33.3)	28 (40.6)	2 (9.5)	0.898	0.0001
	sul2	42 (46.7)	35 (50.7)	7 (33.3)		
	sul3	9 (10)	3 (4.3)	6 (28.6)		
Tetracycline	*tet(A)*	40 (44.4)	32 (46.4)	8 (38.1)	1.000	0.0001
	tet(B)	13 (14.4)	12 (17.4)	1 (4.8)		
	tet(X)	1 (1.1)	1 (1.4)	0		
Trimethoprim	*dfrA1*	34 (37.8)	27 (39.1)	7 (33.3)	0.978	0.0001

	dfrA5	2 (2.2)	2 (2.9)	0
	dfrA12	8 (8.9)	2 (2.9)	6 (28.6)
	dfrA14	5 (5.6)	5 (7.2)	0
	dfrA17	2 (2.2)	2 (2.9)	0

828 [a]Percentage is estimated based on the total number of strains associated to each fimbria type. A p-value <0.05 is
829 considered significant.
830

831

832 **Table 4.** Distribution of virulence genes (other than F4 and F18-encoding fimbria) among the 90 ETEC
833 isolates.

Function	Genes	Total (%)	F4 (%)[a]	F18 (%)[a]	p-value	OD 95% CI
Toxins	*astA*	77 (85.6)	**69 (100)**	8 (38.1)	<0.0001	∞ (25.71-∞)
	ltcA	82 (91.1)	**69 (100)**	13 (62)	<0.0001	∞ (9.402-∞
	Sta	21 (23.3)	13 (18.8)	8 (38.1)	0.0820	
	Stb	85 (94.4)	**69 (100)**	16 (76.2)	0.0005	∞ (5.465-∞)
Fimbriae	*lpfA*	15 (16.7)	2 (2.9)	**13 (61.9)**	<0.0001	0.01837 (0.003635-0.08)
	fasA	2 (2.2)	2 (2.9)	0	>0.9999	
Adhesion	*iha*	77 (85.6)	57 (82.6)	20 (95.2)	0.2854	
	air	4 (4.4)	0	**4 (19)**	0.0023	0 (0.0-0.02885)
Colicin	*cba*	33 (36.7)	**33 (47.8)**	0	<0.0001	∞ (4.748-∞)
	cma	47 (52.2)	**42 (60.9)**	5 (23.8)	0.0053	4.978 (1.631-13.29)
	celb	10 (11.1)	8 (11.6)	2 (9.5)	>0.9999	
Microcin	*mchB*	9 (10)	9 (13)	0	0.1096	
	mchC	9 (10)	9 (13)	0	0.1096	
	mchF	9 (10)	9 (13)	0	0.1096	
	mcmA	9 (10)	9 (13)	0	0.1096	
Others	*capU*	55 (61.1)	**54 (78.3)**	1 (4.8)	<0.0001	72 (11.75-759.7)
	gad	40 (44.4)	**35 (50.7)**	5 (23.8)	0.0439	3.294 (1.09-8.811)
	sepA	10 (11.1)	10 (14.5)	0	0.1088	
	iss	20 (22.2)	1 (1.4)	**19 (90.5)**	<0.0001	0.001548 (0.0001506-0.0)
	eilA	5 (5.6)	0	**5 (23.8)**	0.0005	0 (0.0-0.1830)

834 [a]Percentage is estimated based on the total number of strains associated to each fimbria type.

OR: Odds ratio is indicated when p-value <0.05; CI: confidence interval; ∞: infinity. Significant differences (p-value <0.05) are indicated in bold

Applied and Environmental Microbiology

838 **Table 5.** Plasmid replicons predicted by PlasmidFinder among the 90 ETEC isolates.

Plasmid group	Replicon	Total (%)	F4 (%)[a]	F18 (%)[a]	p-value	OR 95% CI
IncF	FII	83 (92.2%)	**69 (100)**	14 (66.7)	<0.0001	∞ (7.269-∞)
	FIB	82 (91.1%)	**69 (100)**	13 (61.9)	<0.0001	∞ (9.402-∞)
	FIC	49 (54.4%)	41 (59.4)	8 (38.1)	0.1323	
	FIA	1 (1.1)	1 (1.4)	0	>0.9999	
IncI	I1	65 (72.2%)	45 (65.2)	**20 (95.2)**	0.0057	0.09375 (0.0086-0.6311)
	I2	12 (13.3)	12 (17.4)	0	0.0614	
IncX	X1	33 (36.7%)	12 (17.4)	**21 (100)**	<0.0001	0.000 (0.000-0.06711)
	X4	3 (3.3)	3 (4.3)	0	>0.9999	
IncQ	Q1	21 (23.3)	**21 (30.4)**	0	0.0024	∞ (2.202-∞)
Col-like	Col156	15 (16.7)	13 (18.8)	2 (9.5)	0.5056	
	Col4400II	2 (2.2)	2 (2.9)	0	>0.9999	
	ColRNAI	2 (2.2)	2 (2.9)	0	>0.9999	
IncHI	1A	1 (1.1)	1 (1.4)	0	>0.9999	
	1B	1 (1.1)	1 (1.4)	0	>0.9999	
	2	7 (7.7)	7 (10.1)	0	0.1933	
	2A	7 (7.7)	7 (10.1)	0	0.1933	
IncN	IncN	12 (13.3)	12 (17.4)	0	0.0614	
IncB/O/Z		10 (11.1)	3 (4.3)	7 (**33.3**)	0.0012	0.09091 (0.02421-0.415)
IncY		5 (5.5)	5 (7.2)	0	0.5869	
p0111		21 (23.3)	14 (20.3)	7 (33.3)	0.2450	
pKPC		3 (3.3)	3 (4.3)	0	>0.9999	

839 [a]Percentage is estimated based on the total number of strains associated to each fimbria type.
840 OR: Odds ratio is indicated when p-value <0.05; CI: confidence interval; ∞: infinity. Significant differences (p-value
841 <0.05) are indicated in bold

842

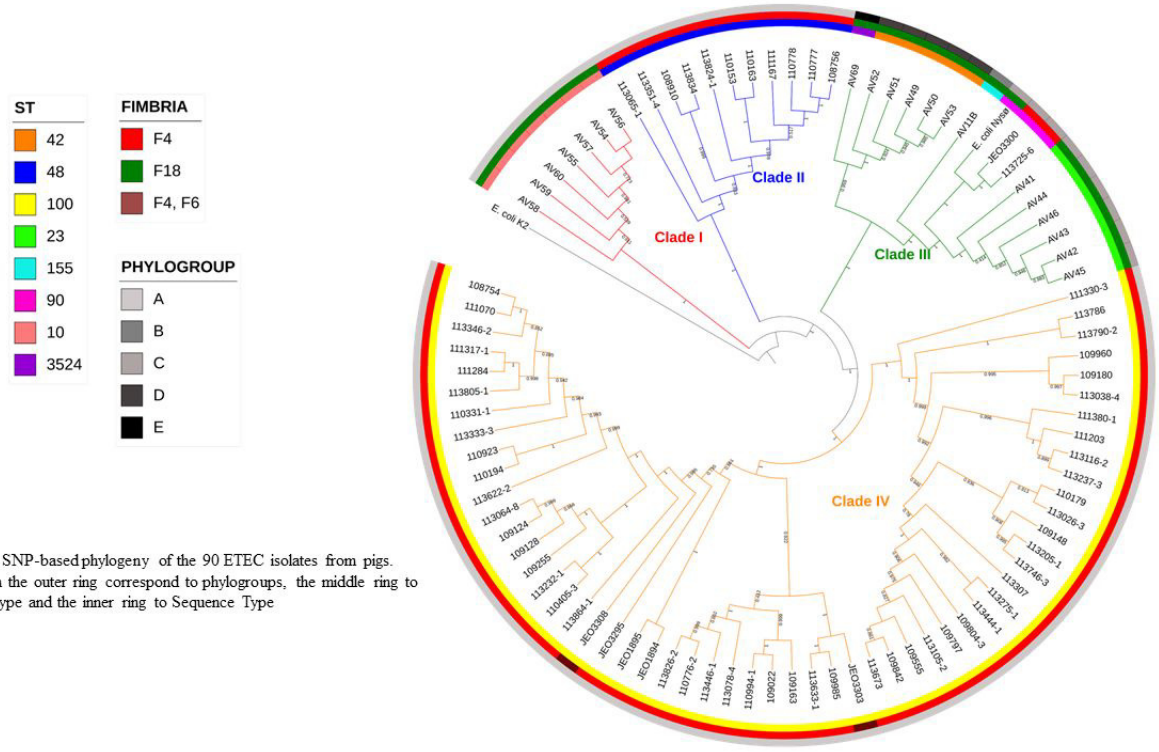

Figure 1. SNP-based phylogeny of the 90 ETEC isolates from pigs. Colours in the outer ring correspond to phylogroups, the middle ring to fimbriae type and the inner ring to Sequence Type

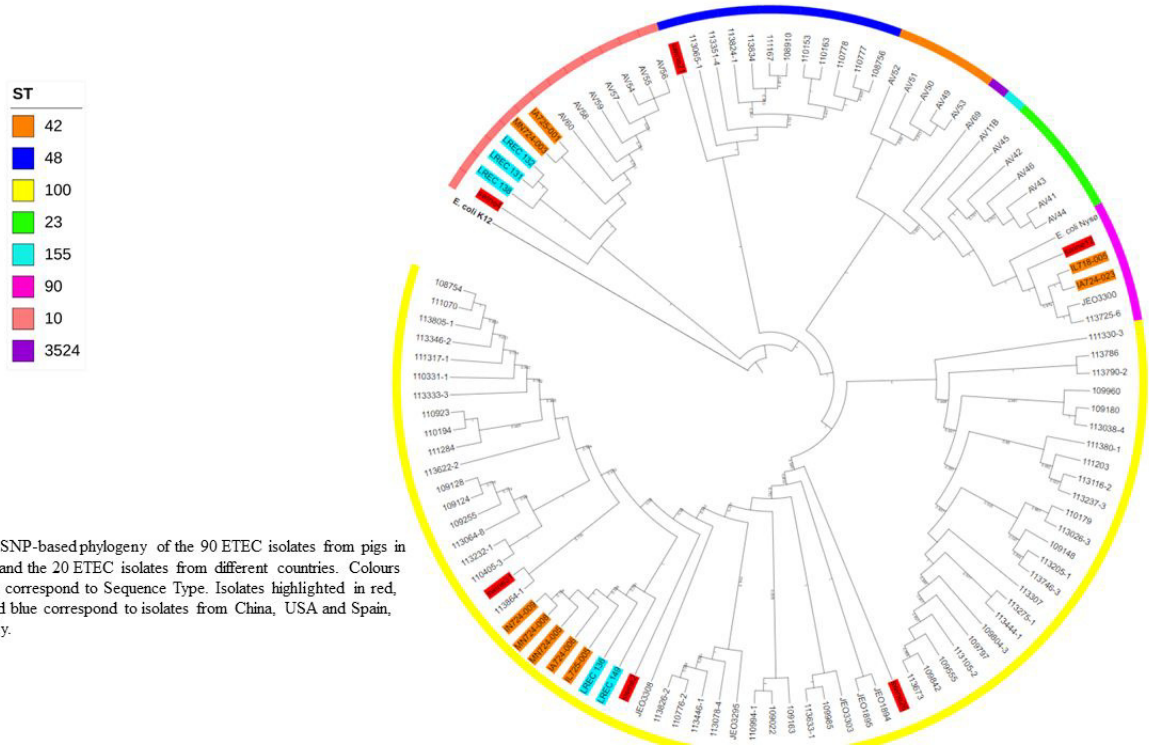

Figure 2. SNP-based phylogeny of the 90 ETEC isolates from pigs in Denmark and the 20 ETEC isolates from different countries. Colours in the ring correspond to Sequence Type. Isolates highlighted in red, orange and blue correspond to isolates from China, USA and Spain, respectively.

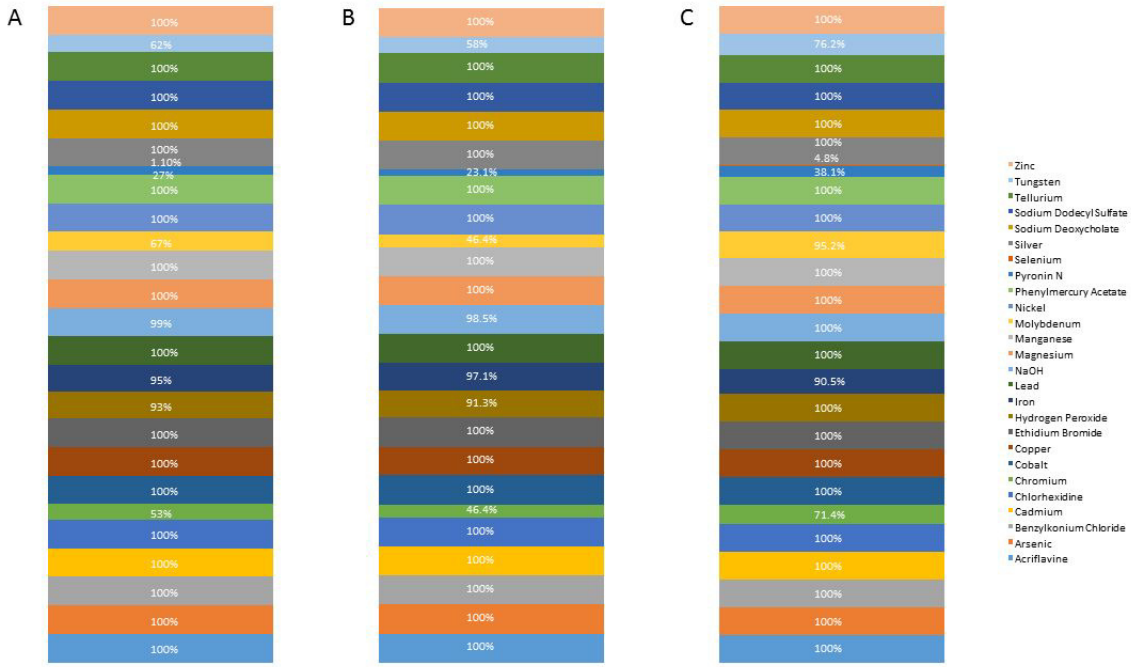

Figure 3. Distribution of metal tolerance and biocides resistance among the 90 ETEC isolates from pigs (A), the F4-positive isolates (B) and the F18-positive isolates (C)