



OPEN

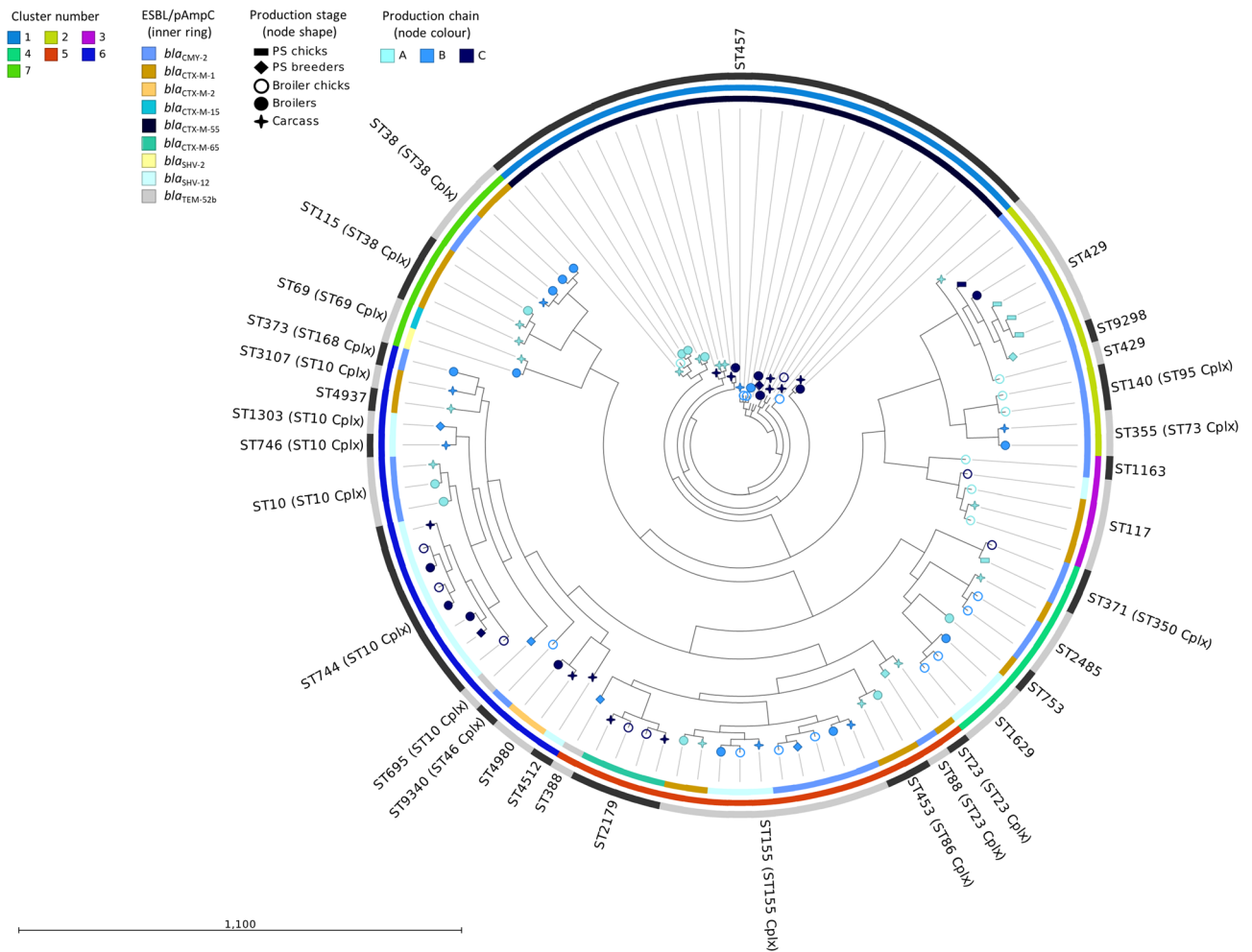
# High-resolution characterisation of ESBL/pAmpC-producing *Escherichia coli* isolated from the broiler production pyramid

Ilias Apostolakos<sup>1</sup>, Claudia Feudi<sup>2</sup>, Inga Eichhorn<sup>2</sup>, Nicola Palmieri<sup>3</sup>, Luca Fasolato<sup>1</sup>, Stefan Schwarz<sup>2</sup> & Alessandra Piccirillo<sup>1</sup>✉

The presence of extended-spectrum  $\beta$ -lactamase (ESBL) or plasmid-mediated AmpC  $\beta$ -lactamase (pAmpC)-producing *Escherichia coli* (ESBL/pAmpC-EC) in livestock is a public health risk given the likelihood of their transmission to humans via the food chain. We conducted whole genome sequencing on 100 ESBL/pAmpC-EC isolated from the broiler production to explore their resistance and virulence gene repertoire, characterise their plasmids and identify transmission events derived from their phylogeny. Sequenced isolates carried resistance genes to four antimicrobial classes in addition to cephalosporins. Virulence gene analysis assigned the majority of ESBL/pAmpC-EC to defined pathotypes. In the complex genetic background of ESBL/pAmpC-EC, clusters of closely related isolates from various production stages were identified and indicated clonal transmission. Phylogenetic comparison with publicly available genomes suggested that previously uncommon ESBL/pAmpC-EC lineages could emerge in poultry, while others might contribute to the maintenance and dissemination of ESBL/pAmpC genes in broilers. The majority of isolates from diverse *E. coli* lineages shared four dominant plasmids (IncK2, IncI1, IncX3 and IncFIB/FII) with identical ESBL/pAmpC gene insertion sites. These plasmids have been previously reported in diverse hosts, including humans. Our findings underline the importance of specific plasmid groups in the dissemination of cephalosporin resistance genes within the broiler industry and across different reservoirs.

Third-generation-cephalosporins (3GCs) are critically important antimicrobial agents in human medicine as they are amongst the few available treatment options in infections caused by multi-drug resistant *Enterobacteriaceae*<sup>1</sup>. Their role in the treatment of food-producing animal diseases (e.g. respiratory infections in ruminants, foot rot and mastitis in cattle) is as well considered as crucial<sup>2</sup>. To preserve the efficacy of 3GCs, their administration in poultry within the European Union is not authorised<sup>3</sup> but there is evidence of off-label use in hatcheries, either *in ovo* or subcutaneously in day-old chicks, to prevent early mortality due to *Escherichia coli* infections<sup>4</sup>. Nevertheless, several studies suggested that the broiler production system acts as a reservoir of 3GC-resistant bacteria, such as extended-spectrum  $\beta$ -lactamase (ESBL) or plasmid-encoded AmpC  $\beta$ -lactamase (pAmpC)-producing *E. coli* (ESBL/pAmpC-EC), given that their prevalence is higher in the broiler production industry compared to other animal sectors<sup>5</sup>. Further, presence of ESBL/pAmpC genes on highly transmissible plasmids, which enables their dissemination in diverse reservoirs<sup>6</sup>, initiated a debate on the exposure risk for consumers via the food chain<sup>7,8</sup>. However, the connection between human exposure and presence of ESBL/pAmpC-EC in broilers was supported by circumstantial evidences showing the presence of identical ESBL/pAmpC genes in closely related *E. coli* of

<sup>1</sup>Department of Comparative Biomedicine and Food Science, University of Padua, 35020 Padua, Italy. <sup>2</sup>Institute of Microbiology and Epizootics, Centre for Infection Medicine, Department of Veterinary Medicine, Freie Universität Berlin, 14163 Berlin, Germany. <sup>3</sup>Department for Farm Animals and Veterinary Public Health, University Clinic for Poultry and Fish Medicine, University of Veterinary Medicine, 1210 Vienna, Austria. ✉email: alessandra.piccirillo@unipd.it



**Figure 1.** Phylogenetic analysis of the 100 sequenced isolates. The SNP tree was created with CSI phylogeny. Nodes shape and colour indicate the broiler production stage and production chain, respectively. The inner ring specifies the ESBL/pAmpC resistance genes identified. The middle ring indicates hierBAPS clusters. The outer ring with alternating dark and light grey segments, marks different STs. Scale bar refers to the branch lengths, which are measured in the number of substitutions per site.

both hosts inferred by low-resolution typing methods<sup>9,10</sup>. The importance of high-resolution, whole genome sequencing (WGS) studies in informing and reducing the uncertainty of quantitative microbial risk assessment (QMRA) models employed in the risk analysis of foodborne antimicrobial resistance was recently stressed<sup>11,12</sup>. Such an assessment for the broiler chain would require the thorough characterisation of ESBL/pAmpC-EC and their transmission patterns across the production pyramid<sup>5</sup>.

In a recent study, we detected a high prevalence of ESBL/pAmpC-EC in an integrated broiler production chain<sup>13</sup>. Here, we proceeded with the WGS-assisted characterisation of a selection of ESBL/pAmpC-EC isolates collected within the formerly published framework. The main aim of the study was to generate high-resolution data on ESBL/pAmpC-EC, advance our understanding on their characteristics and epidemiology in the broiler production and thus provide valuable information to (parts of) QMRA efforts. Our specific objectives were (i) to fully explore the antimicrobial resistance (AMR) and virulence profiles of the isolates (crucial for hazard identification) (ii) to characterise ESBL/pAmpC gene-carrying plasmids, study their epidemiology within the broiler production chain and assess their relevance in other reservoirs (exposure assessment), (iii) to infer the phylogeny of ESBL/pAmpC-EC and identify transmission events within the broiler production (exposure assessment), and (iv) set the isolates of this collection in context with other publicly available genomes and thereby assess their presence in other hosts.

## Results and discussion

**Phylogenetic analysis.** In silico typing revealed a complex genetic background in the 100 sequenced *E. coli* isolates by assignment to 31 unique Sequence Types (STs) (Fig. 1). Two novel STs were identified with one representing isolate each. The ST9298 isolate was obtained from a breeder and is closely related to ST429 as they differ by only one SNP in the *icd* allele (T109A, *icd13* → *icd1072*). The ST9340 isolate was recovered from a day-old broiler chick and is closely related to ST46 of the ST Complex (Cplx) ST10 (C336T, *gyrB1* → *gyrB812*).

Overall, ST457 and ST155 (comprising 25% and 10% of all isolates, respectively) were the most dominant STs followed by ST744 (ST10 Cplx; 8%) and ST429 (7%) (Table 1). The majority of the remaining isolates were assigned to various STs, represented by less than three isolates each, while ST38 (ST38 Cplx), ST117 and ST2179 were represented by four isolates each.

Genomes were mapped against the reference genome of *E. coli* strain MRY15-131 (Accession no. NZ\_AP017620.1). All isolates covered at least 66.3% of the reference genome sequence resulting in a core genome alignment of 3.34 Mbp and 148,830 informative single nucleotide polymorphisms (SNPs) (File S2). Cluster analysis with hierBAPS revealed the presence of seven clusters. The examined isolates clustered primarily in accordance to their ST and subsequently to their ST complex (Fig. 1). Each of the four most dominant STs (ST457, ST155, ST744 and ST429/ST9298) was assigned to an independent cluster. For each of these STs, a separate in-depth SNP analysis was conducted with suitable reference genomes (File S2) to identify clonal transmission events within the studied broiler production pyramid (Fig. S1) and gain insights from the comparative genomic analysis with publicly available genomes (Fig. S2).

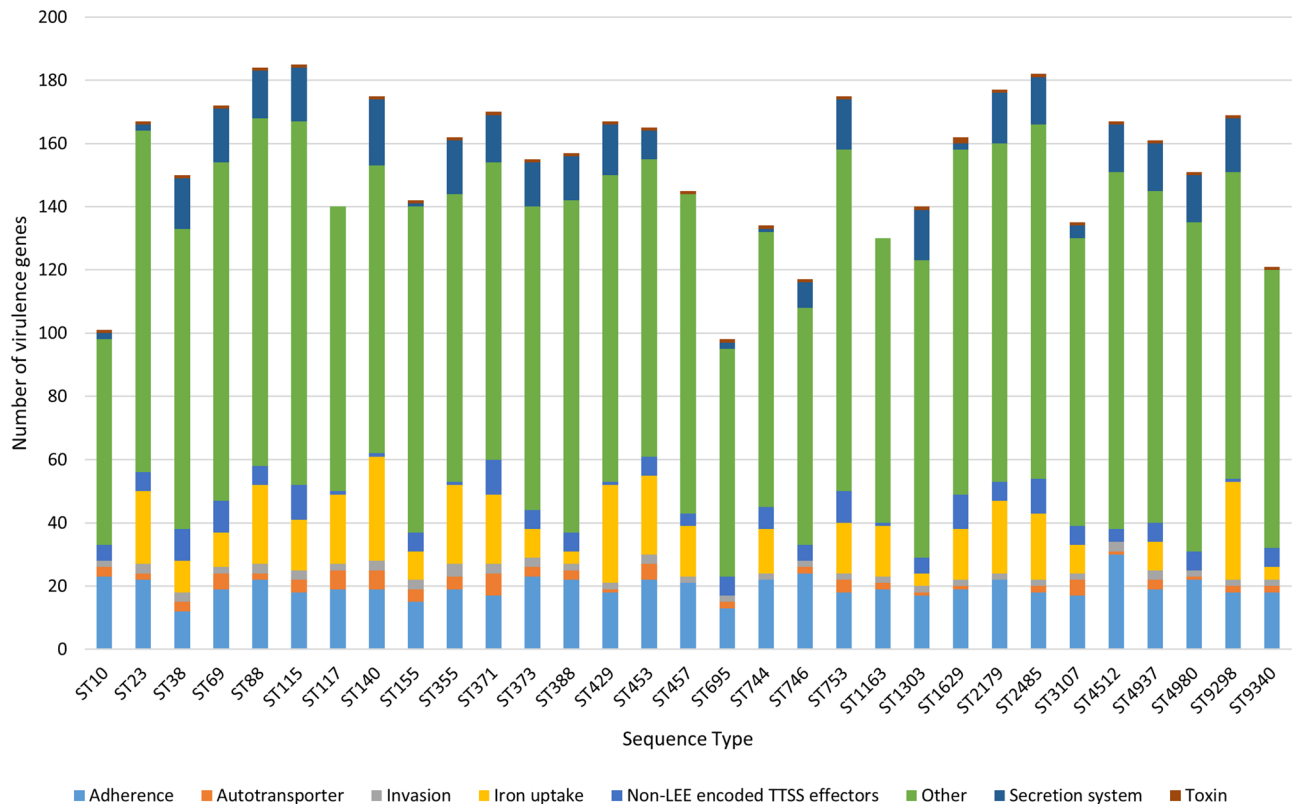
Among all isolates, ST457 was the most widespread ST ( $n = 25$  isolates), representing isolates from all chains and stages of the production pyramid. All ST457 isolates harboured *bla*<sub>CTX-M-55</sub> and exclusively formed Cluster 1 (Fig. 1). In-depth SNP analysis distinguished two clusters, Cluster 1 and Cluster 2, among the ST457 isolates, with a median of pairwise differences (MPD) among all isolates at 3,855 SNPs (Fig. S1a). The cluster 1 ( $n = 19$  isolates) was formed by closely related isolates (MPD 13 SNPs), retrieved from various stages of all three production chains, whereas Cluster 2 ( $n = 6$  isolates) consisted of closely related isolates of chain A, including broiler chicks, fattened broilers and carcasses (MPD 19 SNPs) (Fig. S1a). Given that in previous reports epidemiologically linked isolates had 0–23 SNPs<sup>14</sup>, our results indicate a clonal expansion of ST457 in the majority of the production chains within the studied broiler production system. ST457 is a rare ST, first reported in the United Kingdom as a nosocomial isolate obtained from a urine sample<sup>15</sup>. A *bla*<sub>KPC-3</sub>-*bla*<sub>CTX-M-55</sub>-carrying *E. coli* isolated from a human bloodstream infection in Italy<sup>16</sup> was amongst the few recently reported ST457 isolates. An additional study conducted in the United States, showed the concurrent presence of *mcr-1* and *bla*<sub>CTX-M-55</sub> on the same plasmid carried by an *E. coli* ST457 obtained from a human urinary tract infection<sup>17</sup>. Fifteen out of 22 ST457 genomes at Enterobase harboured ESBL/pAmpC genes, but none of them carried *bla*<sub>CTX-M-55</sub> or was isolated from poultry. In the SNP analysis, six of our ST457 isolates clustered closely with two human *bla*<sub>CTX-M-1</sub> carrying isolates from the Netherlands (unpublished) and one *bla*<sub>CMY-2</sub>-carrying isolate from Germany<sup>14</sup> (Cluster 3, Fig. S2a), and these nine isolates had a MPD of 142 SNPs. In contrast, the remaining ST457 isolates clustered more distantly (MPD 638 SNPs) with livestock isolates (Fig. S2a).

In the overall phylogenetic analysis, Cluster 5, the second most diverse cluster in terms of varying ST content after Cluster 6 (Fig. 1), was mainly represented by ST155 ( $n = 10$  isolates). Isolates of ST155 formed three clusters according to their ESBL/pAmpC gene content (*i.e.* *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-1</sub> and *bla*<sub>SHV-12</sub>) and the within-cluster SNP differences (MPD 5–11 SNPs) suggested clonal dissemination (Fig. S1b). SNP analysis including the respective Enterobase genomes showed that ESBL/pAmpC isolates of ST155 are common in poultry<sup>8,14</sup> ( $n = 22/35$  isolates were of poultry origin). The isolates of this study were grouped in a large and diverse cluster (MPD 1,355 SNPs) mainly of poultry origin, although three human isolates were included in the same cluster (Cluster 1, Fig. S2b).

The second largest and most diverse cluster in terms of included ST groups (Cluster 6;  $n = 21$  isolates,  $n = 4$  STs), primarily included ST10 Cplx isolates ( $n = 15$ ) with ST744 as the most prevalent ( $n = 8$  isolates) (Fig. 1). ST744-*bla*<sub>SHV-12</sub> *E. coli* ( $n = 8/100$ ) isolated from breeders, broilers and carcasses of chain C formed two clusters with an MPD of respectively six and 21 SNPs (Fig. S1c). The ESBL/pAmpC isolates of ST744 at Enterobase ( $n = 79$ ) were mainly of human origin, and none of the isolates carried *bla*<sub>SHV-12</sub> or was of poultry origin. Our ST744 isolates clustered with a set of human isolates, which mainly carried *bla*<sub>CTX-M-15</sub> (Cluster 4, Fig. S2c). We found further evidence for the distribution of ST744 in animals<sup>18</sup> and humans<sup>14,19</sup>, which corroborates the omnipresence of ST10 Cplx members in both populations<sup>19</sup>.

All isolates of Cluster 2, including the novel ST9298 isolate and the closely related ST429 isolates ( $n = 7$ ), harboured *bla*<sub>CMY-2</sub> (Fig. 1). Four out of five isolates from one-day-old, imported Parent Stock (PS) chicks were assigned to ST429. The remaining ST429 isolates ( $n = 3$ ), together with one isolate of the novel ST9298 isolated from a PS breeder, were subsequently recovered from broilers and a carcass and belonged to chains A and C (Fig. S1d). The MPD of SNPs (43–76 SNPs) among the isolates of each of the three identified clusters (Fig. S1d) were generally larger compared to the aforementioned ST groups, even for a set of three isolates that were isolated from the same PS breeder flock (Cluster 2, Fig. S1d). Hence, instead of direct clonal transmission, the repeated introduction and subsequent circulation of closely related ST429-*bla*<sub>CMY-2</sub> isolates through the import of PS chicks in each broiler production cycle may be a more likely scenario. SNP analysis of ST429/ST9298 isolates of this study and Enterobase ( $n = 30$ ) revealed that almost all isolates harboured *bla*<sub>CMY-2</sub>, clustered together with an MPD of 76 SNPs and were predominantly of poultry origin (Cluster 1, Fig. S2d). The only exception in the ST429 dataset was one human *bla*<sub>TEM-52</sub>-carrying *E. coli*, which showed an MPD of 2,976 SNPs to the isolates of Cluster 1. The Enterobase isolates originated from various stages of the broiler production, including isolates from breeders and poultry meat (not displayed). This finding highlights the existence of a conserved ST429-*bla*<sub>CMY-2</sub> lineage, which seems to have spread in the European poultry production.

**Resistance genes other than ESBL/pAmpC genes.** Sulphonamide resistance genes were present in 89% ( $n = 89/100$ ) of isolates with the *sul2* gene as most prevalent (Table 1). Aminoglycoside acetyltransferase genes (*aac(3)*-like), nucleotidyltransferase genes (*aadA*-like) and phosphotransferase genes (*strA*, *strB*, *aph(3)*-like) were found in 87% of isolates. The tetracycline resistance genes *tet(A)* and *tet(B)* were found in 76% and 12% of the *E. coli*, respectively. Moreover, WGS data showed trimethoprim resistance conferred by *dfrA*-like genes in 45% of isolates, with *dfrA14* being the most prevalent. The chloramphenicol exporter gene *floR* was pre-



**Figure 2.** Distribution of virulence genes by function in the sequence types identified in this study.

sent in 21 out of the 43 isolates, which carried phenicol resistance genes, while *cmlA1*, *catA1* and *catB3* were also present. Plasmid-mediated quinolone resistance (PMQR) genes were identified in 20% of the *E. coli*, specifically *qnrS1* (13%), *qnrS2* (4%) and *qnrB19* (2%). In addition, quinolone resistance was mediated by chromosomal mutations in quinolone resistance determining regions (QRDR) in 49% of isolates. The most prevalent mutations were those resulting in the amino acid exchanges Ser83Leu in GyrA (47%), followed by Ser80Ile in ParC (34%). Twelve isolates had combinations of PMQR genes and mutations in the *gyrA* and/or *parC* genes. In addition, 14 isolates harboured at least one of the macrolide resistance genes *mph(A)* or *mph(B)*, which were found in twelve and two isolates, respectively. Furthermore, the rifampicin resistance gene *arr-3*, and the *aac(6')Ib-cr* gene, which confers resistance to both fluoroquinolones and aminoglycosides, were found in all four ST2179 isolates of this study (Table 1). Although we did not confirm resistance phenotypes by susceptibility testing, previous studies showed high accordance between resistance phenotype and genotype<sup>20,21</sup>. Multi-drug resistance, which is frequently observed in ESBL/pAmpC-EC<sup>22,23</sup>, is crucial for their selection and dissemination by the use of non- $\beta$ -lactam antimicrobial agents and thus may help to explain the high prevalence of ESBL/pAmpC-EC in broilers even in the absence of cephalosporin use<sup>24</sup>. Indeed in our case, all PS and fattening broiler flocks, except for three, received standard antimicrobial treatments for therapeutic reasons, which mainly included the administration of amoxicillin for ~ 5 days although occasional use of enrofloxacin and oxytetracycline was recorded as well<sup>13</sup>. Therefore, antimicrobial administration may have led to the selection of resistant isolates in the chicken gut and enabled their further dissemination<sup>24</sup>. However, alternative explanations may exist since ESBL/pAmpC genes can be maintained in bacterial populations even in the absence of antimicrobial selection pressure through conjugation and as a result of their low fitness cost<sup>25</sup>. Nevertheless, antimicrobial administration seems to play a central role in the complex mechanisms that modulate the abundance of ESBL/pAmpC plasmids in a population through the modulation of pathways that lead to their positive or negative frequency-dependent selection<sup>25</sup>.

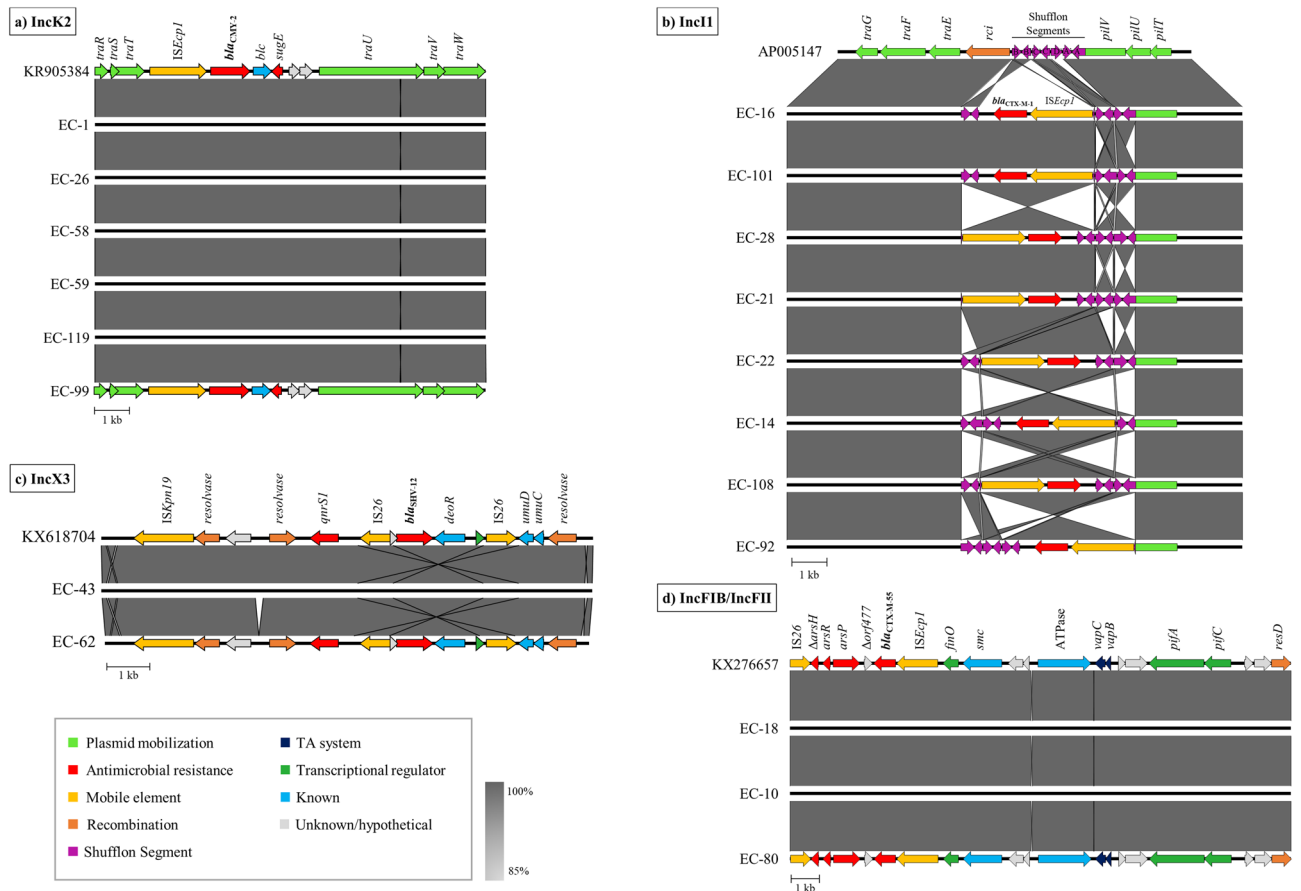
**Virulence genes.** Overall, 320 different virulence genes (VGs) were identified in the 100 sequenced *E. coli* and individual isolates carried a median of 170 VGs. A total of 56 virulence genes were ubiquitous. These included the curlin fibers genes (*csg*), common-pilus genes (*ecp*), enterobactins (*ent*), ferrienterobactins (*fep*), the fimbrial proteins *fimD* and *fimG*, flagellar proteins of the families *flg* and *fli* and the flagellar motor proteins *motAB*. Only a few of the VGs showed an association with specific STs. While screening for a total of 89 adherence related VGs, we detected an average of 20 adherence-related VGs per isolate (Fig. 2). The adherence-conferring *pil* genes (*pilC*, *pilQ*, *pilR*, *pilT*, *pilV*, *pilW*, *pilX*, *pilY*, *pilZ*) were found in 13/100 isolates and are in general connected to uropathogenic *E. coli* (UPEC)<sup>26</sup>. Moreover, the *etp* gene cluster (which codes for a type II secretion system) and the *stcE* protease gene, mostly associated with enterohaemorrhagic *E. coli* O157:H7<sup>27,28</sup>, were found in two and one isolates, respectively (File S1). Autotransporters showed a heterogeneous distribution among the various STs. From this category, the *upaG* adhesion gene, associated with UPEC, was found in 41 isolates<sup>29</sup>. Additionally, the temperature-sensitive hemagglutinin autotransporter *tsh* gene was identified in ST371 isolates (n = 2/100) and

has previously been linked to avian pathogenic *E. coli* (APEC)<sup>30</sup>. With regard to the 33 iron uptake VGs analysed, the *sit*-like iron transporter and *iuc*-like aerobactin genes were associated with 83.8% (n=26/31) and 61.3% (n=19/31) of the STs identified in this study, respectively. Furthermore, the toxin haemolysin E (*hlyE*) gene was present in the majority of STs (n=29/31) and in 83/100 isolates. Most isolates showed the presence of VGs with miscellaneous functions (Fig. 2), such as the increased serum survival (*iss*) gene (n=87/100), which is a crucial VG of extra-intestinal pathogenic *E. coli* (ExPEC)<sup>31</sup>, as well as *hma* (n=9/100) that encodes a haem acquisition protein and plays a critical role in the colonisation of the urinary tract<sup>32</sup>. Additionally, the ESBL/pAmpC-EC were assigned to pathotypes based on previously defined criteria (File S1)<sup>33–36</sup>. Interestingly, 56 isolates belonging to 18 different STs were classified as atypical enteroaggregative *E. coli* (aEAEC) as they possessed *aatA* but lacked the *aggR* regulon<sup>35,37</sup>. About one third (34%) of the aEAEC group was represented by ST457 isolates (n=19/25). In the study of Bamidele et al.<sup>36</sup>, eight out of 31 adults suffering from travellers' diarrhoea carried aEAEC that harboured only *aatA* and not *aggR* or *aaiC* and were positive in the Hep-2 cell adherence test, which remains the gold-standard in the assessment of aEAEC/EAEC pathogenicity<sup>36,37</sup>. Both typical and atypical EAEC have been predominantly associated with paediatric diarrhoea in developing countries<sup>35,37</sup>. Further, 51 isolates from 12 STs were characterised as ExPEC [presence of  $\geq 2$  of *pap* (P fimbriae), *sfa/foc* (S/F1C fimbriae), *afa/dra* (Dr binding adhesins), *iutA* (aerobactin receptor), and *kpsMT II* (group 2 capsule synthesis)]<sup>31,33</sup>. All ST457 isolates (n=25/25) made up half (49%) of the ExPEC group. Moreover, 39 isolates, belonging to 13 different STs, were classified as avian pathogenic *E. coli* (APEC) [presence of *iutA*, *hlyF* (haemolysin F), *iss*, *iroN* (siderophore receptor), *ompT* (outer membrane protein T)]<sup>34</sup>. The APEC group mainly (38.5%) consisted of ST744 (n=8/8) and ST429/ST9298 (n=7/8) isolates. Interestingly, 42 isolates showed hybrid pathotypes, with 30 isolates from seven STs displaying the aEAEC/ExPEC pathotype (File S1). This latter subgroup was predominantly represented by the isolates of ST457 (n=19/25), suggesting an association of this ST with the aEAEC/ExPEC pathotypes, followed by ST155 (n=4/10). While the possession of ExPEC VGs was expected for the APEC and ExPEC pathotypes, it was less anticipated for aEAEC, since it is a potential intestinal pathogenic *E. coli* (InPEC) pathotype. A recent study suggested that such hybrid pathotypes are dominant in human faecal microbiota and are usually undetected by routine diagnostic tests<sup>38</sup>. Although prediction of pathogenicity is fundamental for QMRA models, the distinction between commensals and pathogens has become a bigger conundrum in the WGS era. For example, disease-associated VGs were identified in probiotic bacteria with a long history of safe use<sup>39</sup>. Another study suggested that pathogenicity may be linked to genome reduction rather than the acquisition of certain VGs<sup>40</sup>. It is therefore likely that traditional pathotyping schemes will be modified in the future and that multi-omic pathogenicity predictive models that consider the presence and expression of all genes rather than small sets will gain more ground as genome databases grow<sup>41</sup>. Nevertheless, the virulence gene repertoire of ESBL/pAmpC-EC reported here and in other studies, may explain their adaptation to and persistence in various niches, such as the poultry environment<sup>41</sup>.

**Genomic localisation of ESBL/pAmpC genes.** ESBL/pAmpC genes were often associated with plasmids belonging to specific incompatibility groups and in some cases associations with certain STs were also observed. The most dominant ESBL/pAmpC gene/plasmid association was the presence of *bla*<sub>CMY-2</sub> on IncK2 plasmids, which was found in 27% of the sequenced isolates, belonging to a variety of different STs (11 out of 31 STs) although predominantly in ST429 (7 out of the 27 *bla*<sub>CMY-2</sub>-carrying IncK2 plasmids) (Table 1). Another striking association was the occurrence of *bla*<sub>CTX-M-55</sub> on IncFIB/IncFII plasmids (25%), which was uniquely present in all isolates of ST457. The *bla*<sub>CTX-M-1</sub> gene was mainly found on IncI1 plasmids (13%), but IncI2, IncFIB/IncFII and IncHI2 plasmids also harboured this gene (Table 1). Moreover, *bla*<sub>SHV-12</sub> was found to be carried by IncX3 or IncI1 plasmids in 12% and 6% of the isolates, respectively (Table 1). For the four *bla*<sub>CTX-M-65</sub>-carrying ST2179 isolates, the gene was not associated with any plasmid replicons and we thus concluded its putative chromosomal localisation, which will be further investigated. The genetic context of the most dominant plasmid types is described below.

**IncK2 plasmids carrying *bla*<sub>CMY-2</sub>.** In 27 of the thirty *E. coli* genomes harbouring *bla*<sub>CMY-2</sub>, the resistance gene was associated with the novel incompatibility group IncK2<sup>42</sup>. An approximately 12 kb large region including *bla*<sub>CMY-2</sub> was highly conserved in all plasmids and was almost identical (99.81–100% nucleotide sequence identity) to that of plasmid pDV45 (GenBank accession no. KR905384.1). *ISEcp1-bla*<sub>CMY-2</sub>-*blc-sugE* was integrated into the *tra* locus between the *traU* and *traT* genes in all cases (Fig. 3a). The two imperfect inverted repeats (IV-L: TGACGGTGATCCT; IV-R: AGCATCTCCGTC) described by Seiffert et al.<sup>42</sup> flanking this mobile element were also identified in those sequences. Recent studies reported the dissemination of the highly conserved cluster of IncK2 plasmids in diverse *E. coli* of various reservoirs<sup>14,43</sup>, although they seem to be dominant in broilers and broiler meat<sup>44</sup>. In our study, IncK2 plasmids were identified in isolates assigned to eleven different STs recovered from different stages of the production pyramid (Table 1, Fig. 3a). Interestingly, a further comparative analysis with ABACAS of all ST429-*bla*<sub>CMY-2</sub> isolates of poultry origin present in Enterobase using plasmid pDV45 as reference, showed in all cases the presence of the aforementioned *ISEcp1-bla*<sub>CMY-2</sub>-*blc-sugE* element integrated in IncK2 plasmids (data not shown). It is also noteworthy that plasmids with almost identical IncK2 backbones have been identified in *E. coli* isolated from hospitalised patients<sup>42,44</sup>. Furthermore, Donà et al.<sup>45</sup> reported a similar IncK2 plasmid harbouring *mcr-1*, *bla*<sub>TEM-1</sub> and *sul2* but not *bla*<sub>CMY-2</sub> in *E. coli* from retail chicken meat in Switzerland. In our study, *bla*<sub>CMY-2</sub> was the only resistance gene present on IncK2 plasmids with the exception of four plasmids that carried also *aac(3)-VIa*, *aadA1* and *sul1* (Table 1).

**IncI1 plasmids carrying *bla*<sub>CTX-M-1</sub>.** The *bla*<sub>CTX-M-1</sub> gene was mainly detected on IncI1 plasmids (n=13 of 17), which together with IncN plasmids, contribute to the spread of *bla*<sub>CTX-M</sub> genes<sup>46</sup>. All *bla*<sub>CTX-M-1</sub>-carrying IncI1



**Figure 3.** Linear comparison of (a) IncK2, (b) IncI1, (c) IncX3 and (d) IncFIB/IncFII plasmid regions carrying *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-1</sub>, *bla*<sub>SHV-12</sub> and *bla*<sub>CTX-M-55</sub> genes, respectively, identified in our study. The open reading frames are represented with arrows, with the arrowhead indicating the direction of transcription. Their role in transfer, recombination, antimicrobial resistance, their association to mobile genetic elements and known or unknown functions are colour-coded. Areas shaded in grey indicate the percentage of nucleotide sequence identity.

plasmids belonged to plasmid MLST type pST3 or to closely related variants (Table 1). The genetic region containing *bla*<sub>CTX-M-1</sub> (~12 kb) was highly conserved amongst all plasmids (99.96% nucleotide sequence identity) and the well-studied IncI1 plasmid R64 (GenBank accession no. AP005147) served as a reference (Fig. 3b). The mobile element *ISEcp1-bla*<sub>CTX-M-1</sub> was integrated in the shufflon region, between the *rci* shufflon-specific recombinase and the *pilV* gene, which represent a hotspot for DNA rearrangements<sup>46</sup>. *Rci* binds to specific recombination sites (*sfx* repeats) and causes the rearrangement of shufflon segments, which leads to variable arrangements of the shufflon region, even within the same plasmid preparation<sup>47</sup>, in what seems to be a non-random process<sup>48</sup>. It is, therefore, highly probable that the observed re-arrangements in the integration site of *bla*<sub>CTX-M-1</sub> (Fig. 3b) do not represent multiple insertion events but rather recombination events due to *Rci* activity. The insertion of the module *ISEcp1-bla*<sub>CTX-M-1</sub> into the shufflon region seems to be common for IncI1 plasmids belonging to pST3, as shown in previous studies identifying this plasmid type from various sources<sup>46,49–51</sup>. A recent study reported identical *bla*<sub>CTX-M-1</sub>-carrying IncI1 pST3 plasmids from chicken, chicken meat, and bloodstream infections in humans<sup>21</sup>. Furthermore, additional resistance genes, mainly *tet(A)* and *sul2* (Table 1), were found on the IncI1-*bla*<sub>CTX-M-1</sub> plasmids and their co-localisation with *bla*<sub>CTX-M-1</sub> has been described before<sup>21,45</sup>.

**IncX3 plasmids carrying *bla*<sub>SHV-12</sub>.** The resistance gene *bla*<sub>SHV-12</sub> was detected in 18 of the sequenced isolates and was carried by either IncX3 (n = 12/18) or IncI1 (n = 6/18) plasmids (Table 1). The genetic context of *bla*<sub>SHV-12</sub> in IncX3 plasmids included two IS26 elements in opposite orientation, with *qnrS1* being inserted upstream (Fig. 3c). The most similar sequences identified by BLASTn analysis of the ~11 kb region against the NCBI nucleotide database were ~48 kb large IncX3-*bla*<sub>SHV-12</sub>-*qnrS1* plasmids isolated from chicken faeces and a human patient in the Netherlands, reported by Liakopoulos et al. in 2018<sup>52</sup> (e.g. GenBank accession no. KX618704, 99.82–100% nucleotide sequence identity). Similar IncX3 plasmids with the same *bla*<sub>SHV-12</sub> surrounding region have been isolated from different countries and sources, including humans<sup>52,53</sup>. Despite of *bla*<sub>SHV-12</sub> being mainly associated with IncI1 plasmids<sup>54</sup>, in our study we observed a higher occurrence of the IncX3 plasmid compared to IncI1 (Table 1). This shift has also recently been reported in the Netherlands, where a slow decrease of *bla*<sub>SHV-12</sub>-encoding IncI1 plasmids in favour of IncX3 plasmids was observed<sup>52</sup>. Except for *bla*<sub>SHV-12</sub> and *qnrS1*, no additional resistance genes were co-located on IncX3 plasmids.

ESBL/pAmpC gene	Plasmid Replicon type [no. of isolates]	pMLST/RST	MLST [no. of isolates]	Additional resistance genes located on ESBL/pAmpC gene-carrying plasmids <sup>c</sup>	Other resistance genes <sup>d</sup>						
<i>bla</i> <sub>CMY-2</sub>	IncB/O/K/Z [27] <sup>a</sup>	–	ST429 [7]	<i>aac(3)-VIa*</i> , <i>aadA1*</i> , <i>sul1*</i>	<i>tet(A)</i>						
			ST9298 [1]	–	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>tet(A)</i> , <i>sul1</i>						
			ST155 [5]	–	<i>bla</i> <sub>TEM-1b*</sub> , <i>aac(3)-VIa</i> , <i>aadA1</i> , <i>tet(A)</i> , <i>sul1</i>						
			ST10 [3]	–	<i>bla</i> <sub>TEM-1b*</sub> , <i>aac(3)-VIa*</i> , <i>aadA1</i> , <i>dfrA1*</i> , <i>tet(A)*</i> , <i>sul1*</i>						
			ST140 [2]	<i>aac(3)-VIa*</i> , <i>aadA1*</i> , <i>sul1*</i>	<i>tet(A)</i>						
			ST2485 [2]	–	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>tet(A)</i> , <i>sul1</i>						
			ST371 [2]	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>sul1</i>	<i>tet(A)</i>						
			ST38 [2]	–	<i>aac(3)-VIa*</i> , <i>aadA1*</i> , <i>aph(3')-Ic*</i> , <i>strA</i> , <i>strB</i> , <i>tet(B)</i> , <i>sul1*</i> , <i>sul2</i>						
			ST1163 [1]	–	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>sul1</i>						
			ST373 [1]	–	–						
			ST9340 [1]	–	<i>bla</i> <sub>TEM-1b</sub> , <i>aac(3)-VIa</i> , <i>aadA1</i> , <i>tet(A)</i> , <i>sul1</i> , <i>dfrA1</i>						
			ST355 [2]	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>strA</i> , <i>strB</i> , <i>floR</i> , <i>tet(A)</i> , <i>sul1</i> , <i>sul2</i>	<i>aac(3)-VIa</i> , <i>aadA1</i> , <i>strA</i> , <i>strB</i> , <i>floR</i> , <i>tet(A)</i> , <i>sul1</i> , <i>sul2</i>						
			ST88 [1]	–	<i>bla</i> <sub>TEM-1b</sub> , <i>aac(3)-VIa</i> , <i>aadA1</i> , <i>aadA2</i> , <i>strA</i> , <i>strB</i> , <i>cmlA1</i> , <i>floR</i> , <i>tet(A)</i> , <i>sul1</i> , <i>sul2</i> , <i>sul3</i>						
<i>bla</i> <sub>CTX-M-55</sub>	IncFIB/IncFII [25]	F18:A-B1	ST457 [25]	<i>bla</i> <sub>TEM-1b*</sub> , <i>aph(3')-Ia*</i> , <i>tet(A)*</i> , <i>sul2*</i> , <i>dfrA14*</i>	<i>aac(3)-IIa*</i> , <i>aadA1*</i> , <i>strA*</i> , <i>strB*</i> , <i>floR*</i> , <i>sul3*</i>						
			<i>bla</i> <sub>CTX-M-1</sub>	IncI1 [13]	pST3 <sup>b</sup>	ST115 [3]	<i>aadA5*</i> , <i>tet(A)*</i> , <i>sul2*</i> , <i>dfrA17*</i>	–			
						ST117 [3]	<i>tet(A)</i> , <i>sul2</i>	–			
						ST155 [2]	<i>tet(A)</i> , <i>sul2</i>	<i>bla</i> <sub>TEM-1b</sub> , <i>aadA5</i> , <i>aph(3')-Ia</i> , <i>strA</i> , <i>strB</i> , <i>catA1</i> , <i>tet(B)</i> , <i>dfrA17</i>			
						ST453 [2]	<i>tet(A)</i> , <i>sul2</i>	<i>bla</i> <sub>TEM-1b</sub> , <i>strA</i> , <i>strB</i> , <i>dfrA14</i>			
						ST23 [1]	<i>tet(A)</i> , <i>sul2</i>	<i>bla</i> <sub>TEM-1b</sub> , <i>aadA1</i> , <i>sul3</i>			
						ST2485 [1]	<i>tet(A)</i> , <i>sul2</i>	<i>bla</i> <sub>TEM-1b</sub>			
						ST753 [1]	<i>tet(A)</i> , <i>sul2</i>	<i>bla</i> <sub>TEM-1b</sub> , <i>qnrS1</i>			
						ST38 [2]	–	<i>aadA1</i> , <i>strA</i> , <i>strB</i> , <i>tet(A)</i> , <i>mph(B)</i> , <i>sul1</i> , <i>sul2</i> , <i>dfrA1</i>			
						ST4937 [1]	–	<i>aadA1</i> , <i>strA</i> , <i>strB</i> , <i>catA1</i> , <i>tet(A)</i> , <i>mph(A)</i> , <i>sul1</i> , <i>dfrA1</i>			
						ST3107 [1]	–	<i>aac(3)-IId</i> , <i>tet(A)</i> , <i>sul2</i> , <i>dfrA14</i> , <i>dfrA17</i>			
						<i>bla</i> <sub>SHV-12</sub>	IncX3 [12]	–	ST744 [8]	<i>qnrS1</i>	<i>bla</i> <sub>TEM-1b</sub> , <i>aadA5</i> , <i>aph(3')-Ia*</i> , <i>strA</i> , <i>strB</i> , <i>catA1</i> , <i>tet(B)</i> , <i>mph(A)</i> , <i>sul1</i> , <i>sul2</i> , <i>dfrA17</i>
									ST1629 [3]	<i>qnrS1</i>	<i>tet(A)</i> , <i>sul1</i>
ST746 [1]	<i>qnrS1</i>	<i>cml</i> , <i>qnrA1</i> , <i>sul1</i> , <i>dfrA5</i>									
ST155 [3]	–	<i>aadA1</i> , <i>aadA2</i> , <i>cmlA1</i> , <i>tet(A)</i> , <i>sul3</i>									
ST4512 [1]	–	<i>aadA1</i> , <i>aadA2</i> , <i>cmlA1</i> , <i>tet(A)</i> , <i>sul3</i>									
ST117 [1]	–	<i>aadA1</i> , <i>aadA2</i> , <i>cmlA1</i> , <i>tet(A)</i> , <i>sul3</i>									
ST1303 [1]	–	<i>bla</i> <sub>TEM-1b</sub> , <i>tet(A)</i> , <i>qnrB19</i>									
<i>bla</i> <sub>CTX-M-65</sub>	Chromosome [4]	–	ST2179 [4]	–	<i>bla</i> <sub>OXA-1</sub> , <i>bla</i> <sub>TEM-1a*</sub> , <i>bla</i> <sub>TEM-1b*</sub> , <i>aadA1</i> , <i>aadA2</i> , <i>aadA5*</i> , <i>strA*</i> , <i>strB*</i> , <i>catB3</i> , <i>cmlA1</i> , <i>tet(A)*</i> , <i>qnrS2</i> , <i>mph(A)*</i> , <i>sul1*</i> , <i>sul2*</i> , <i>sul3</i> , <i>dfrA17*</i> , <i>arr-3*</i> , <i>aac(6')-Ib-c</i>						
			ST388 [1]	–	<i>qnrB19</i>						
<i>bla</i> <sub>TEM-52b</sub>	IncX1 [2]	–	ST695 [1]	–	–						
			ST4980 [2]	–	<i>bla</i> <sub>TEM-1b</sub> , <i>aadA1</i> , <i>strA</i> , <i>strB</i> , <i>tet(A)</i> , <i>sul1</i> , <i>sul2</i> , <i>dfrA1</i>						
<i>bla</i> <sub>CTX-M-2</sub>	IncQ1 [2]	–	ST69 [1]	–	<i>bla</i> <sub>TEM-1b</sub> , <i>strA</i> , <i>strB</i> , <i>tet(A)</i> , <i>qnrS1</i> , <i>sul2</i>						
<i>bla</i> <sub>CTX-M-15</sub>	IncY [1]	–	ST69 [1]	–	<i>aadA1</i> , <i>aadA2</i> , <i>cmlA1</i> , <i>tet(A)</i> , <i>sul3</i>						
<i>bla</i> <sub>SHV-2</sub>	Untypable [1]	–	ST69 [1]	–	–						

**Table 1.** Characteristics of 100 ESBL/pAmpC-EC isolates subjected to whole genome sequencing. <sup>a</sup>The IncB/O and IncK replicons were identified by PBRT. Further analysis indicated that these plasmids belong to the novel incompatibility group IncK2<sup>12</sup>. <sup>b</sup>Four isolates had truncated *ardA* genes in their assemblies but the other four alleles (*pilL2*, *repI12*, *sogS1* and *trbA4*) were those of pST3. One isolate had also a truncated *ardA* gene and a novel *sogS* allele, differing from *sogS1* by the mutation C24T. One isolate represented a novel IncI1 pST, closely related to pST3 (*ardA1* → *ardA7*). <sup>c</sup>Resistance genes in this column were found on the ESBL/pAmpC gene-carrying plasmids after analysis with ABACAS. The following plasmid types, found on 76% of isolates, were studied: *bla*<sub>CMY-2</sub>-IncB/O/K/Z, *bla*<sub>CTX-M-55</sub>-IncFIB/IncFII, *bla*<sub>CTX-M-1</sub>-IncI1, *bla*<sub>SHV-12</sub>-IncX3. Resistance genes marked with an asterisk indicate that they were not present in all the corresponding plasmids. <sup>d</sup>Resistance genes marked with an asterisk indicate that they were not present in all the isolates of the corresponding ST.

**IncFIB/FII plasmids carrying *bla*<sub>CTX-M-55</sub>** Interestingly, *bla*<sub>CTX-M-55</sub> was found only on IncFIB/FII plasmids of replicon sequence type F18:A-B1 (n = 25/25). This variant of the CTX-M enzyme has been mainly reported in Asian countries, where it represents the second most common variant in *E. coli* isolated from companion and livestock animals<sup>55,56</sup> and was only recently described in *E. coli* isolates of animal origin in France<sup>18</sup>.

BLASTn analysis of the region including *bla*<sub>CTX-M-55</sub> yielded homologous (99–100% nucleotide sequence identity) *bla*<sub>CTX-M-55</sub>-carrying F18:A-B1 plasmids isolated from clinical *E. coli* isolates of human origin in the United States (GenBank accession nos. CP041997.1, KX276657.1, CP029748.1) and wastewater in Canada (GenBank accession no. MK878525.1). The human related plasmids additionally carried *mcr* colistin resistance genes. Here we used the *mcr-1*-carrying plasmid pMR0516mcr isolated from a clinical ST457 isolate (GenBank accession no. KX276657.1)<sup>17</sup> as a reference sequence (Fig. 3d). In all of our ST457 isolates, the *bla*<sub>CTX-M-55</sub> gene is associated with *ISEcp1*, within one of the most common genetic modules, *ISEcp1-bla*<sub>CTX-M-55-orf477</sub><sup>57</sup>, and is located between the *finO* gene, which is involved in the regulation of plasmid conjugation, and the arsenic resistance operon *ars* (Fig. 3d). Other co-located resistance genes (e.g. *bla*<sub>TEM-1b</sub>, *dfrA14*, *sul2*) were found in varying combinations in the majority of our IncFIB/FII plasmids (Table 1).

## Conclusions and outlook

Our WGS approach revealed a complex landscape of *E. coli* genotypes present in the broiler production. We demonstrated that 76% of the sequenced isolates shared four dominant plasmid types that harboured ESBL/pAmpC genes. The finding of diverse STs from various production stages, which share closely related plasmid backbones with identical ESBL/pAmpC gene insertions, is a strong indication that horizontal gene transfer via the exchange of plasmids that have been previously reported to be transmissible<sup>42,46,52</sup>, plays a fundamental role in the dissemination of ESBL/pAmpC genes across the broiler production pyramid<sup>5</sup>. In addition to  $\beta$ -lactamase genes, the sequenced isolates carried genes which conferred resistance to four other antimicrobial classes, on average, a fact that is relevant for public health and definitely needs to be considered in the hazard identification step of a QMRA. The co-localisation of several other resistance genes on ESBL/pAmpC-carrying plasmids might explain their persistence and propagation through selection pressure by non-3GC antimicrobial agents<sup>49</sup>. In addition to horizontal gene transfer, clonal expansion of several closely related ESBL/pAmpC-EC along the production pyramid was also observed and has been previously reported<sup>58</sup>. The identification of phylogenetically related isolates in distantly located farms, but also in subsequent production levels of the same production chain, further supports the argument of clonal transmission from the top of the pyramid to the lower levels of production. However, at least part of the ESBL/pAmpC-EC dissemination may be explained by indirect transmission events within the studied integrated production system (e.g. hatchery-to-farm or farm-to-farm transmission), which underlines the need to maintain and continuously monitor biosecurity protocols and thus prevent transmission of ESBL/pAmpC-EC by e.g. the movement of vehicles, equipment and personnel<sup>5</sup>. The importance of such measures is reflected by the fact that in accordance with other studies<sup>22,59</sup>, in our recent prevalence study<sup>13</sup> we encountered the lowest occurrence of ESBL/pAmpC-EC in PS breeders, a production stage where increased biosecurity standards are implemented due to the higher value of these flocks compared to fattening broilers. Also, as Hiroi et al.<sup>60</sup> showed, contamination of the environment of broilers with ESBL/pAmpC-EC may be a more important factor for their colonisation than the use of antimicrobial agents.

We previously discussed a potential introduction of *bla*<sub>CMY-2</sub> isolates through the import of day-old PS chicks, based on the finding that all isolates from this production stage had the phylogroup B2-*bla*<sub>CMY-2</sub> genotype<sup>13</sup>. Here, we show that at least a part of these isolates belong to ST429/ST9298, a highly conserved *bla*<sub>CMY-2</sub>-positive lineage, which seems to be prevalent in the European poultry production<sup>14</sup>. Results from the comparative analysis with ST429 genomes at Enterobase suggest the circulation of a conserved ST429/ST9298 lineage carrying *bla*<sub>CMY-2</sub> on IncK2 plasmids in the poultry production of at least five European countries, which needs to be studied further in future investigations, considering also the pathogenic potential of this ST for poultry and its role in colibacillosis (n = 7/8 isolates were APEC). Additionally, these findings further support the notion that ESBL/pAmpC-EC may spread through the interconnected breeder stock supply chains<sup>5,14,58</sup>, which warrants their eradication in pedigree flocks and in hatcheries. This can be achieved by e.g. the complete cessation of ceftiofur administration in these flocks for prophylaxis<sup>61</sup>. Further, comparison with publicly available data proved that the occurrence of ESBL/pAmpC-EC in poultry is a dynamic phenomenon, since STs and ESBL/pAmpC genes not common in the European context, such as ST457 and *bla*<sub>CTX-M-55</sub>, can emerge and dominate in poultry. The fact that our *E. coli* isolates of ST457 and ST744 were closer related to human ones than to those originating from livestock (Fig. S2), raises questions about their ability to colonise the human gut upon exposure, transmit their resistance determinants, and/or exert their virulence potential. The latter aspect is especially relevant for ST457, which has a documented record of association with human disease and its isolates were assigned to defined pathotypes (ExPEC and/or EAEC) in our virulence analysis. Using an in vitro colonisation simulation system, Anjum et al.<sup>62</sup> showed that *E. coli* isolated from human faeces and poultry meat were able to adapt and transmit their AMR plasmids to other *Enterobacteriaceae* in the human gut. Nevertheless, a recent study demonstrated that the dissemination of ESBL/pAmpC genes between animals and humans seems to be mediated by specific plasmid types rather than by expansion of successful *E. coli* clones<sup>9</sup>. This hypothesis is further supported by the detection of closely related plasmid backbones with identical ESBL/pAmpC gene insertion sites between the isolates of this study and those of human origin<sup>17,21,42,52</sup>.

Our isolate selection procedure was suitable for describing the diversity of ESBL/pAmpC-EC genotypes that may be encountered in the broiler production, studying the underlying transmission mechanisms and providing solid evidence of transmission events along the broiler production pyramid. However, our study has limitations. Quantitative estimates on transmission pathways as well as statistical associations between the identified genotypes, which are essential for QMRAs, cannot be inferred with our analysis since this would require the



sequencing of a larger and more representative part of our initial collection of isolates<sup>11</sup>. Such approaches are hampered by the decreasing but still significant cost of WGS. The second limitation concerns the findings of our comparative genomic analysis, which should be interpreted with caution since potential bias in Enterobase data may exist. Our objective was not to make deductions on potential transmission events between the analysed hosts, which would also require thorough epidemiological data, but to gain insights from the host distribution and phylogenetic relatedness of dominant ESBL/pAmpC-EC lineages in the context of an extensive collection of publicly available genomes supported by basic metadata. The two major conclusions on the highly conserved, poultry-related ST429-IncK2-*bla*<sub>CMY-2</sub> and the emerging ST457-*bla*<sub>CTX-M-55</sub> were noteworthy outcomes of this analysis that ought to be investigated further.

Our results provide valuable insights into the characteristics of ESBL/pAmpC-producing *E. coli*, their plasmid contents and distribution in the broiler production chain. These results contribute to the growing amount of available high throughput data that will serve as the baseline for future food chain QMRA models. For example, the AMR profile, the genetic context of resistance genes and the virulence potential of isolates described in this study are relevant for guiding the risk identification and characterisation steps<sup>11</sup>. Given the key role of plasmids in the dissemination of ESBL/pAmpC genes elucidated here and in other studies<sup>9</sup>, exposure assessment modelling should definitely include, if not prioritise on, the transfer probability of ESBL/pAmpC plasmids, especially of those with a wide host-spectrum. The output of such thorough QMRAs will be the risk estimate of human exposure to resistance determinants via the (broiler) food chain and the identification of intervention measures with the maximum mitigation effect<sup>11</sup>.

## Methods

**Bacterial isolates.** One-hundred isolates from a collection of ESBL/pAmpC-EC isolated from three production chains, namely A (n=37), B (n=32), and C (n=31), of an integrated broiler company in Italy were included in this study<sup>13</sup>. Briefly, we sampled with cloacal swabs one PS flock per production chain at two time points; (i) one-day-old PS chicks upon arrival at the rearing farms and (ii) PS breeders at ~21 weeks of age during the laying period at the production farms. Afterwards, the offspring of the PS flock were sampled in four fattening farms per production chain at the start (one-day-old) and the end (~30-days-old) of the production cycle (broiler chicks and broilers, respectively). Carcasses from the sampled broilers were collected at the slaughterhouse after chilling and isolation of ESBL/pAmpC-EC was performed after rinsing the whole carcass with buffer peptone water<sup>13</sup>. *E. coli* phylogenetic groups and ESBL/pAmpC genes have been previously determined for these isolates by PCR and sequencing<sup>13</sup>. At least one isolate per production chain, production stage, ESBL/pAmpC gene, and phylogroup combination (e.g. isolate EC-1 in File S1, chain A/PS chicks/*bla*<sub>CMY-2</sub>/phylogroup B2) was randomly selected for sequencing to fully explore the diversity of ESBL/pAmpC-EC genotypes across the production pyramid and identify potential transmission events. The 100 selected isolates originated from the following production stages: PS chicks (n=5), PS breeders (n=8), broiler chicks (n=26), broilers (n=28) and slaughterhouse/carcass (n=33) (File S1).

**Whole genome sequencing and in silico typing.** DNA extraction and sequencing were carried out at two different sites. At the first site, DNA was isolated from 32 isolates with the Master Pure Genomic DNA-Purification Kit (Epicenter, USA), libraries were prepared with the Nextera XT library preparation kit (Illumina, USA) and sequencing on an Illumina MiSeq platform using the MiSeq v3 reagent kit (Illumina) with 2×300 bp paired-end reads. At the second site, DNA was isolated from 68 isolates with the Invisorb Spin Tissue Mini Kit (Invitek, Germany), library preparation was done with the Nextera XT library preparation kit and sequencing on an Illumina HiSeqX platform with 2×150 bp paired-end reads at a private company (Macrogen, Korea).

Raw reads were directly submitted for processing at the Enterobase database<sup>63</sup>. The Enterobase backend pipeline assures high quality assemblies by including reads pre-processing and post-correction of SPAdes<sup>64</sup> assemblies (<https://enterobase.readthedocs.io/>). Assemblies that did not meet the quality control criteria were not used for downstream analysis. Assembly statistics (coverage, N50, genome size and contig number) and accession numbers (Enterobase barcodes) for each strain can be found in File S1. Further, sequenced isolates were genotyped in silico with regard to acquired resistance genes and chromosomal mutations (ResFinder 3.2 and PointFinder 3.1, respectively)<sup>20</sup>, as well as plasmid replicon types (PlasmidFinder 2.0)<sup>65</sup> and pSTs when available (pMLST 2.0)<sup>65</sup>, using the tools at default settings, whereas *E. coli* STs were extracted from the Enterobase automated pipeline, which follows the Achtman MLST scheme.<sup>66</sup> Presence of VGs was assessed with BLAST<sup>67</sup> alignments ( $E < 10^{-5}$ ) to the Virulence Factors Database (VFDB)<sup>68</sup> extended with VGs from literature (*hlyF*, *iss*, *ompT*)<sup>31,34</sup>. Based on their virulence gene content, isolates were assigned to known pathotypes<sup>33-36</sup>. A VG was considered to be associated to a given ST if at least 80% of the isolates included in that ST carried the specific VG.

**Plasmid characterisation.** In several isolates, the ESBL/pAmpC gene-carrying plasmid was identified by the presence of the resistance gene and the plasmid replicon on the same contig. Otherwise, the contigs containing ESBL/pAmpC genes but missing the plasmid replicon were compared by BLASTn<sup>67</sup> (default settings) to the most similar, complete reference plasmids available in GenBank database. To identify all the contigs corresponding to particular plasmids, the draft genome of each isolate and the reference plasmid were aligned with ABACAS<sup>69</sup> as previously described<sup>70</sup>. Identification of the genetic context of ESBL/pAmpC genes was done by aligning the identified plasmid contigs against the reference plasmid in Geneious 10.1.3, using the multiple alignment tool and default settings. EasyFig<sup>71</sup> was used to visualise aligned plasmid regions of interest.

To confirm results obtained in silico, plasmid transformation experiments were performed for at least one isolate with a specific ST/ESBL/pAmpC gene combination, resulting in a total of 32 plasmid transformations. Briefly, plasmid DNA was extracted with the alkaline lysis method<sup>72</sup>, electro-competent DH5α *E. coli* cells

(Invitrogen, Denmark) were prepared according to Smith et al.<sup>73</sup> and transformed by electroporation utilizing an Eppendorf Eporator (Eppendorf, Germany). Potential transformants harbouring an ESBL/pAmpC-carrying plasmid were isolated on LB agar (Microbiol, Italy) containing 1 mg/L cefotaxime. The incompatibility group of plasmids transferred from the wild type isolates was determined by PCR-based replicon typing (PBRT)<sup>74,75</sup>.

**Phylogenetic analysis.** The phylogeny of the sequenced isolates was deduced by a SNP-based mapping analysis. A suitable reference was selected among all complete *E. coli* genomes (n=983) (<https://www.ncbi.nlm.nih.gov/genome/167>) using Mash<sup>76</sup> to find the reference genome with the least Mash distance from (i) all sequenced isolates and (ii) all isolates of each dominant ST group. The phylogenetic analysis was extended with all publicly available genomes at Enterobase, belonging to the most dominant STs of our collection and with a confirmed presence of ESBL/pAmpC genes by local ResFinder analysis. *E. coli* genomes were mapped to the selected reference (File S2) and a maximum likelihood SNP tree was created with CSI phylogeny<sup>77</sup> using default settings. Cluster calculation was performed with the hierBAPS module of the Bayesian Analysis of Population Structure (BAPS) software v6.0<sup>78</sup>. BAPS clusters were assigned based on the output SNP alignment of CSI phylogeny with 2 levels of hierarchy and a maximum number of cluster (K) equal to the number of ST groups (n = 31) in the phylogenetic analysis of sequenced isolates, and equal to the number of isolates in the phylogenetic analysis of each ST group. All phylogenetic trees (Figs. 1, S1 and S2) generated with CSI phylogeny were visualised as phylograms and annotated with metadata in CLC Genomics Workbench version 12.0 (Qiagen, Denmark). The Mash-based ranking of complete reference genomes and their NCBI accession numbers, the barcodes of Enterobase genomes used for the comparative genomic analysis and the statistics of the SNP analysis (e.g. % reference coverage, number of SNPs) can be found in File S2.

**Ethics statement.** Collection of cloacal swabs from live birds in the study of Apostolakos et al.<sup>13</sup>, was done by veterinarians of the private poultry company as part of the routine monitoring of the flock health status and conducted in compliance with good veterinary practices. This routine monitoring does not require approval by an institutional and/or licensing committee.

### Data availability

All data generated or analysed during this study are included in this published article and its Supplementary Information files.

Received: 22 February 2020; Accepted: 29 May 2020

Published online: 07 July 2020

### References

1. WHO. Critically important antimicrobials for human medicine—6th revision. (2019).
2. OIE. *List of antimicrobial agents of veterinary importance*. (2015).
3. Heinrich, K., Chan, D., Fussell, R. J., Kay, J. F. & Sharman, M. Can the unauthorised use of ceftiofur be detected in poultry?. *Food Addit. Contam. A* **30**, 1733–1738 (2013).
4. Baron, S. et al. Impact of third-generation-cephalosporin administration in hatcheries on fecal *Escherichia coli* antimicrobial resistance in broilers and layers. *Antimicrob. Agents Chemother.* **58**, 5428–5434 (2014).
5. Dame-Korevaar, A., Fischer, E. A. J., van der Goot, J., Stegeman, A. & Mevius, D. Transmission routes of ESBL/pAmpC producing bacteria in the broiler production pyramid, a literature review. *Prev. Vet. Med.* **162**, 136–150 (2019).
6. Brolund, A. & Sandegren, L. Characterization of ESBL disseminating plasmids. *Infect. Dis. (Auckl)* **48**, 18–25 (2016).
7. Evers, E. G. et al. Comparative exposure assessment of ESBL-producing *Escherichia coli* through meat consumption. *PLoS ONE* **12**, e0169589 (2017).
8. Lazarus, B., Paterson, D. L., Mollinger, J. L. & Rogers, B. A. Do human extraintestinal *Escherichia coli* infections resistant to expanded-spectrum cephalosporins originate from food-producing animals? A systematic review. *Clin. Infect. Dis.* **60**, 439–452 (2015).
9. de Been, M. et al. Dissemination of cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by specific plasmid lineages. *PLoS Genet.* **10**, e1004776 (2014).
10. Leverstein-van Hall, M. A. et al. Dutch patients, retail chicken meat and poultry share the same ESBL genes, plasmids and strains. *Clin. Microbiol. Infect.* **17**, 873–880 (2011).
11. Collineau, L. et al. Integrating whole-genome sequencing data into quantitative risk assessment of foodborne antimicrobial resistance: A review of opportunities and challenges. *Front. Microbiol.* **10**, 1–18 (2019).
12. Dorado-García, A. et al. Molecular relatedness of ESBL/AmpC-producing *Escherichia coli* from humans, animals, food and the environment: a pooled analysis. *J. Antimicrob. Chemother.* **73**, 339–347 (2018).
13. Apostolakos, I., Mughini-Gras, L., Fasolato, L. & Piccirillo, A. Assessing the occurrence and transfer dynamics of ESBL/pAmpC-producing *Escherichia coli* across the broiler production pyramid. *PLoS ONE* **14**, e0217174 (2019).
14. Pietsch, M. et al. Whole genome analyses of CMY-2-producing *Escherichia coli* isolates from humans, animals and food in Germany. *BMC Genomics* **19**, 601 (2018).
15. Lau, S. H. et al. Major uropathogenic *Escherichia coli* strain isolated in the Northwest of England identified by multilocus sequence typing. *J. Clin. Microbiol.* **46**, 1076–1080 (2008).
16. Accogli, M. et al. Emergence of *Escherichia coli* ST131 sub-clone H30 producing VIM-1 and KPC-3 carbapenemases, Italy. *J. Antimicrob. Chemother.* **69**, 2293–2296 (2014).
17. McGann, P. et al. *Escherichia coli* harboring *mcr-1* and *bla*<sub>CTX-M</sub> on a novel IncF plasmid: first report of *mcr-1* in the United States. *Antimicrob. Agents Chemother.* **60**, 5107–5107 (2016).
18. Lupo, A., Saras, E., Madec, J.-Y. & Haenni, M. Emergence of *bla*<sub>CTX-M-55</sub> associated with *fosA*, *rmtB* and *mcr* gene variants in *Escherichia coli* from various animal species in France. *J. Antimicrob. Chemother.* **73**, 867–872 (2018).
19. Matamoros, S. et al. Global phylogenetic analysis of *Escherichia coli* and plasmids carrying the *mcr-1* gene indicates bacterial diversity but plasmid restriction. *Sci. Rep.* **7**, 15364 (2017).
20. Zankari, E. et al. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).
21. Valcek, A. et al. Inc11 ST3 and Inc11 ST7 plasmids from CTX-M-1-producing *Escherichia coli* obtained from patients with bloodstream infections are closely related to plasmids from *E. coli* of animal origin. *J. Antimicrob. Chemother.* **74**, 2171–2175 (2019).

22. Dierikx, C. M., van der Goot, J. A., Smith, H. E., Kant, A. & Mevius, D. J. Presence of ESBL/AmpC-producing *Escherichia coli* in the broiler production pyramid: a descriptive study. *PLoS ONE* **8**, e79005 (2013).
23. EFSA/ECDC. The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2017/2018. *EFSA J.* **18**, e06007 (2020).
24. Mo, S. S. *et al.* Emergence of AmpC-producing *Escherichia coli* in the broiler production chain in a country with a low antimicrobial usage profile. *Vet. Microbiol.* **171**, 315–320 (2014).
25. Dimitriu, T. *et al.* Negative frequency dependent selection on plasmid carriage and low fitness costs maintain extended spectrum  $\beta$ -lactamases in *Escherichia coli*. *Sci. Rep.* **9**, 17211 (2019).
26. Lugerling, A. The Pix pilus adhesin of the uropathogenic *Escherichia coli* strain X2194 (O2:K:H6) is related to Pap pili but exhibits a truncated regulatory region. *Microbiology* **149**, 1387–1397 (2003).
27. Grys, T. E., Walters, L. L. & Welch, R. A. Characterization of the stcE protease activity of *Escherichia coli* O157:H7. *J. Bacteriol.* **188**, 4646–4653 (2006).
28. Szabady, R. L., Lokuta, M. A., Walters, K. B., Huttenlocher, A. & Welch, R. A. Modulation of neutrophil function by a secreted mucinase of *Escherichia coli* O157:H7. *PLoS Pathog.* **5**, 2 (2009).
29. Totsika, M. *et al.* Molecular characterization of the ehaG and upaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl. Environ. Microbiol.* **78**, 2179–2189 (2012).
30. Stathopoulos, C., Provence, D. L. & Curtiss, R. Characterization of the avian pathogenic *Escherichia coli* hemagglutinin Tsh, a member of the immunoglobulin a protease-type family of autotransporters. *Infect. Immun.* **67**(2), 772–781 (1999).
31. Johnson, J. R. *et al.* Accessory traits and phylogenetic background predict *Escherichia coli* extraintestinal virulence better than does ecological source. *J. Infect. Dis.* **219**, 121–132 (2019).
32. Hagan, E. C. & Mobley, H. L. T. Haem acquisition is facilitated by a novel receptor Hma and required by uropathogenic *Escherichia coli* for kidney infection. *Mol. Microbiol.* **71**, 79–91 (2009).
33. Johnson, J. R. *et al.* Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrob. Agents Chemother.* **47**, 2161–2168 (2003).
34. Johnson, T. J. *et al.* Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. *J. Clin. Microbiol.* **46**, 3987–3996 (2008).
35. Guerrieri, C. G. *et al.* Typical and atypical enteroaggregative *Escherichia coli* are both virulent in the *Galleria mellonella* model. *Front. Microbiol.* **10**, 1791 (2019).
36. Bamidele, O., Jiang, Z. D. & Dupont, H. Occurrence of putative virulence-related genes, *aatA*, *aggR* and *aaIC*, of Enteroaggregative *Escherichia coli* (EAEC) among adults with travelers' diarrhea acquired in Guatemala and Mexico. *Microb. Pathog.* **128**, 97–99 (2019).
37. Lozer, D. M. *et al.* Genotypic and phenotypic analysis of diarrheagenic *Escherichia coli* strains isolated from Brazilian children living in low socioeconomic level communities. *BMC Infect. Dis.* **13**, 418 (2013).
38. Lindstedt, B.-A., Finton, M. D., Porcellato, D. & Brandal, L. T. High frequency of hybrid *Escherichia coli* strains with combined Intestinal Pathogenic *Escherichia coli* (IPEC) and extraintestinal pathogenic *Escherichia coli* (ExPEC) virulence factors isolated from human faecal samples. *BMC Infect. Dis.* **18**, 544 (2018).
39. Wassenaar, T. *et al.* Virulence genes in a probiotic *E. coli* product with a recorded long history of safe use. *Eur. J. Microbiol. Immunol.* **5**, 81–93 (2015).
40. Merhej, V., Georgiades, K. & Raoult, D. Postgenomic analysis of bacterial pathogens repertoire reveals genome reduction rather than virulence factors. *Brief. Funct. Genomics* **12**, 291–304 (2013).
41. Pitout, J. D. D. Extraintestinal pathogenic *Escherichia coli*: A combination of virulence with antibiotic resistance. *Front. Microbiol.* **3**, 1–7 (2012).
42. Seiffert, S. N. *et al.* Plasmids carrying *bla*<sub>CMY-2/4</sub> in *Escherichia coli* from poultry, poultry meat, and humans belong to a novel IncK subgroup designated IncK2. *Front. Microbiol.* **08**, 1–9 (2017).
43. Manga, I. *et al.* Fecal carriage and whole-genome sequencing-assisted characterization of CMY-2 beta-lactamase-producing *Escherichia coli* in calves at Czech dairy cow farm. *Foodborne Pathog. Dis.* **16**, 42–53 (2019).
44. Berg, E. S. *et al.* Norwegian patients and retail chicken meat share cephalosporin-resistant *Escherichia coli* and IncK/*bla*<sub>CMY-2</sub> resistance plasmids. *Clin. Microbiol. Infect.* **23**(407), e9–407.e15 (2017).
45. Donà, V. *et al.* Heterogeneous genetic location of *mcr-1* in colistin-resistant *Escherichia coli* isolates from humans and retail chicken meat in Switzerland: Emergence of *mcr-1*-carrying IncK2 plasmids. *Antimicrob. Agents Chemother.* **61**, 1–6 (2017).
46. Irrgang, A. *et al.* Diversity of CTX-M-1-producing *E. coli* from German food samples and genetic diversity of the *bla*<sub>CTX-M-1</sub> region on IncI1 ST3 plasmids. *Vet. Microbiol.* **221**, 98–104 (2018).
47. Brouwer, M. S. M. *et al.* IncI shufflons: assembly issues in the next-generation sequencing era. *Plasmid* **80**, 111–117 (2015).
48. Brouwer, M. S. M. *et al.* The shufflon of IncI1 plasmids is rearranged constantly during different growth conditions. *Plasmid* **102**, 51–55 (2019).
49. Zurfluh, K. *et al.* Vertical transmission of highly similar *bla*<sub>CTX-M-1</sub>-harboring IncI1 plasmids in *Escherichia coli* with different MLST types in the poultry production pyramid. *Front. Microbiol.* **5**, 519 (2014).
50. Wang, J. *et al.* Nucleotide sequences of 16 transmissible plasmids identified in nine multidrug-resistant *Escherichia coli* isolates expressing an ESBL phenotype isolated from food-producing animals and healthy humans. *J. Antimicrob. Chemother.* **69**, 2658–2668 (2014).
51. Smith, H. *et al.* Characterization of epidemic IncI1-Iy plasmids harboring Ambler Class A and C genes in *Escherichia coli* and *Salmonella enterica* from animals and humans. *Antimicrob. Agents Chemother.* **59**, 5357–5365 (2015).
52. Liakopoulos, A. *et al.* Genomic and functional characterisation of IncX3 plasmids encoding *bla*<sub>SHV-12</sub> in *Escherichia coli* from human and animal origin. *Sci. Rep.* **8**, 7674 (2018).
53. Alonso, C. A. *et al.* Analysis of *bla*<sub>SHV-12</sub>-carrying *Escherichia coli* clones and plasmids from human, animal and food sources. *J. Antimicrob. Chemother.* **72**, 1589–1596 (2017).
54. Accogli, M. *et al.* IncI1 plasmids associated with the spread of CMY-2, CTX-M-1 and SHV-12 in *Escherichia coli* of animal and human origin. *Clin. Microbiol. Infect.* **19**, 2–4 (2013).
55. Zhang, J. *et al.* Nationwide high prevalence of CTX-M and an increase of CTX-M-55 in *Escherichia coli* isolated from patients with community-onset infections in Chinese county hospitals. *BMC Infect. Dis.* **14**, 1–10 (2014).
56. Jiang, W. *et al.* Prevalence of plasmid-mediated fosfomycin resistance gene *fosA3* among CTX-M-Producing *Escherichia coli* isolates from chickens in China. *Foodborne Pathog. Dis.* **14**, 210–218 (2017).
57. Zhao, W. H. & Hu, Z. Q. Epidemiology and genetics of CTX-M extended-spectrum  $\beta$ -lactamases in Gram-negative bacteria. *Crit. Rev. Microbiol.* **39**, 79–101 (2013).
58. Myrenäs, M. *et al.* Clonal spread of *Escherichia coli* resistant to cephalosporins and quinolones in the Nordic broiler production. *Vet. Microbiol.* **213**, 123–128 (2018).
59. Dame-Korevaar, A. *et al.* Dynamics of CMY-2 producing *E. coli* in a broiler parent flock. *Vet. Microbiol.* **203**, 211–214 (2017).
60. Hiroi, M. *et al.* Factors for occurrence of extended-spectrum  $\beta$ -lactamase-producing *Escherichia coli* in broilers. *J. Vet. Med. Sci.* **74**, 1635–1637 (2012).

61. Verrette, L., Fairbrother, J. M. & Boulianne, M. Effect of cessation of ceftiofur and substitution with lincomycin-spectinomycin on extended-spectrum- $\beta$ -lactamase/ampc genes and multidrug resistance in *Escherichia coli* from a Canadian broiler production pyramid. *Appl. Environ. Microbiol.* **85**, 1–12 (2019).
62. Anjum, M. *et al.* Fate of CMY-2-encoding plasmids introduced into the human fecal microbiota by exogenous *Escherichia coli*. *Antimicrob. Agents Chemother.* **63**, 1–13 (2019).
63. Alikhan, N.-F., Zhou, Z., Sergeant, M. J. & Achtman, M. A genomic overview of the population structure of *Salmonella*. *PLOS Genet.* **14**, e1007261 (2018).
64. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
65. Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).
66. Wirth, T. *et al.* Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.* **60**, 1136–1151 (2006).
67. Johnson, M. *et al.* NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
68. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**, D687–D692 (2019).
69. Assefa, S., Keane, T. M., Otto, T. D., Newbold, C. & Berriman, M. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* **25**, 1968–1969 (2009).
70. Lanza, V. F. *et al.* Plasmid flux in *Escherichia coli* ST131 sublineages, analyzed by Plasmid Constellation Network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. *PLoS Genet.* **10**, e1004766 (2014).
71. Sullivan, M. J., Petty, N. K. & Beatson, S. A. Easyfig: a genome comparison visualizer. *Bioinformatics* **27**, 1009–1010 (2011).
72. Bimboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513–1523 (1979).
73. Smith, H. E., Wisselink, H. J., Vecht, U., Gielkens, A. L. J. & Smits, M. A. High-efficiency transformation and gene inactivation in *Streptococcus suis* type 2. *Microbiology* **141**, 181–188 (1995).
74. Villa, L., García-Fernández, A., Fortini, D. & Carattoli, A. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *J. Antimicrob. Chemother.* **65**, 2518–2529 (2010).
75. Carattoli, A. *et al.* Identification of plasmids by PCR-based replicon typing. *J. Microbiol. Methods* **63**, 219–228 (2005).
76. Ondov, B. D. *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 132 (2016).
77. Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M. & Lund, O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS ONE* **9**, e104984 (2014).
78. Cheng, L., Connor, T. R., Siren, J., Aanensen, D. M. & Corander, J. Hierarchical and spatially explicit clustering of DNA Sequences with BAPS Software. *Mol. Biol. Evol.* **30**, 1224–1228 (2013).

## Acknowledgements

This work was supported by the University of Padua (Fondi Investimento Strategico di Dipartimento (SID), Anno 2016—prot. BIRD167540). The authors would like to thank Lisa Carraro and Raffaella Franch for their assistance in the preparation of genomic libraries.

## Author contributions

A.P. and I.A. designed the study. C.F., I.E. and I.A. acquired the data. I.A., C.F. and N.P. performed the data analysis. I.A. prepared the first draft of the manuscript. I.A., C.F., N.P., I.E., S.S., L.F. and A.P. interpreted the data, read, contributed to, and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-68036-9>.

**Correspondence** and requests for materials should be addressed to A.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020