




# One Health Genomic Surveillance of *Escherichia coli* Demonstrates Distinct Lineages and Mobile Genetic Elements in Isolates from Humans versus Livestock

Catherine Ludden,<sup>a,b</sup> Kathy E. Raven,<sup>c</sup> Dorota Jamrozy,<sup>b</sup> Theodore Gouliouris,<sup>c,d,e</sup> Beth Blane,<sup>c</sup> Francesc Coll,<sup>a,b</sup> Marcus de Goffau,<sup>b</sup> Plamena Naydenova,<sup>c</sup> Carolyne Horner,<sup>f</sup> Juan Hernandez-Garcia,<sup>g</sup> Paul Wood,<sup>h</sup> Nazreen Hadjirin,<sup>g</sup> Milorad Radakovic,<sup>g</sup> Nicholas M. Brown,<sup>d,e,f</sup> Mark Holmes,<sup>g</sup>  Julian Parkhill,<sup>b</sup> Sharon J. Peacock<sup>a,b,c</sup>

<sup>a</sup>London School of Hygiene & Tropical Medicine, London, United Kingdom

<sup>b</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom

<sup>c</sup>Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, United Kingdom

<sup>d</sup>Clinical Microbiology and Public Health Laboratory, Public Health England, Cambridge, United Kingdom

<sup>e</sup>Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom

<sup>f</sup>British Society for Antimicrobial Chemotherapy, Birmingham, United Kingdom

<sup>g</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

<sup>h</sup>Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, United Kingdom

**ABSTRACT** Livestock have been proposed as a reservoir for drug-resistant *Escherichia coli* that infect humans. We isolated and sequenced 431 *E. coli* isolates (including 155 extended-spectrum  $\beta$ -lactamase [ESBL]-producing isolates) from cross-sectional surveys of livestock farms and retail meat in the East of England. These were compared with the genomes of 1,517 *E. coli* bacteria associated with bloodstream infection in the United Kingdom. Phylogenetic core genome comparisons demonstrated that livestock and patient isolates were genetically distinct, suggesting that *E. coli* causing serious human infection had not directly originated from livestock. In contrast, we observed highly related isolates from the same animal species on different farms. Screening all 1,948 isolates for accessory genes encoding antibiotic resistance revealed 41 different genes present in variable proportions in human and livestock isolates. Overall, we identified a low prevalence of shared antimicrobial resistance genes between livestock and humans based on analysis of mobile genetic elements and long-read sequencing. We conclude that within the confines of our sampling framework, there was limited evidence that antimicrobial-resistant pathogens associated with serious human infection had originated from livestock in our region.

**IMPORTANCE** The increasing prevalence of *E. coli* bloodstream infections is a serious public health problem. We used genomic epidemiology in a One Health study conducted in the East of England to examine putative sources of *E. coli* associated with serious human disease. *E. coli* from 1,517 patients with bloodstream infections were compared with 431 isolates from livestock farms and meat. Livestock-associated and bloodstream isolates were genetically distinct populations based on core genome and accessory genome analyses. Identical antimicrobial resistance genes were found in livestock and human isolates, but there was limited overlap in the mobile elements carrying these genes. Within the limitations of sampling, our findings do not support the idea that *E. coli* causing invasive disease or their resistance genes are commonly acquired from livestock in our region.

**KEYWORDS** ESBL, *Escherichia coli*, antimicrobial resistance, genomics, livestock

**Citation** Ludden C, Raven KE, Jamrozy D, Gouliouris T, Blane B, Coll F, de Goffau M, Naydenova P, Horner C, Hernandez-Garcia J, Wood P, Hadjirin N, Radakovic M, Brown NM, Holmes M, Parkhill J, Peacock SJ. 2019. One Health genomic surveillance of *Escherichia coli* demonstrates distinct lineages and mobile genetic elements in isolates from humans versus livestock. mBio 10:e02693-18. <https://doi.org/10.1128/mBio.02693-18>.

**Editor** Philippe J. Sansonetti, Pasteur Institute

**Copyright** © 2019 Ludden et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Catherine Ludden, [catherine.ludden@lshtm.ac.uk](mailto:catherine.ludden@lshtm.ac.uk).

**Received** 5 December 2018

**Accepted** 11 December 2018

**Published** 22 January 2019

*Escherichia coli* is a leading cause of infection in hospitals and the community (1, 2). The prevalence of *E. coli* bloodstream infections has shown a marked increase in Europe and the United States since the early 2000s (3–6). This has been associated with the emergence and global dissemination of *E. coli* that produce extended-spectrum  $\beta$ -lactamases (ESBL-*E. coli*). Such isolates are resistant to many penicillin and cephalosporin antibiotics (3, 5) and are associated with excess morbidity, mortality, longer hospital stay, and higher health care costs compared with infections caused by *E. coli* that are not ESBL producers (7–9). Successful therapy has been further challenged by the emergence of multidrug-resistant (MDR) *E. coli* with acquired resistance to the carbapenem drugs and more recently to colistin, a drug of last resort for multidrug-resistant infections (10, 11).

Tackling the rising trends in prevalence of MDR *E. coli* infections in humans requires an understanding of reservoirs and sources for human acquisition. Food-producing animals have been proposed as a source of ESBL-*E. coli* in humans based on comparison of bacterial genotypes using multilocus sequence typing (MLST) (12–14). This method lacks sufficient discrimination to generate robust phylogenetic comparisons of population genetics and does not capture information on accessory genome composition such as genes encoding drug resistance. Whole-genome sequencing overcomes both of these limitations, but there are limited published data on the use of this technique to address the transmission of antibiotic-resistant *E. coli* between livestock and humans (15). Here, we report the findings of a genomic epidemiological investigation of *E. coli* sourced from livestock, meat, and patients with bloodstream infections within a tightly defined geographical location.

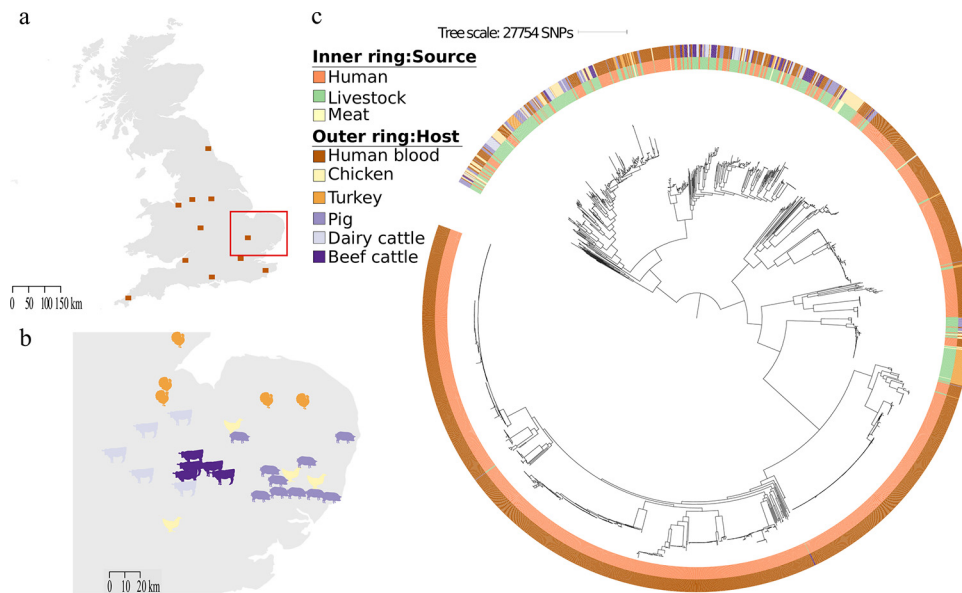
## RESULTS

**Isolation of *E. coli* from livestock farms and retail meat.** A cross-sectional survey was performed between 2014 and 2015 to isolate ESBL-*E. coli* and non-ESBL-producing *E. coli* from livestock at 29 farms in the East of England, United Kingdom (UK) (10 cattle [5 beef cattle and 5 dairy cattle], 10 pig, and 9 poultry [4 chickens and 5 turkeys]) (Fig. 1b). A total of 136 pooled fecal samples (34 from cattle, 53 from pigs, and 49 from poultry) were collected and cultured, with an average of five pooled samples taken per farm. Numerous colonies were picked and identified from each sample to capture different lineages in the same sample. *E. coli* were isolated from all 29 farms, and ESBL-*E. coli* were isolated from 16 (55%) of these farms (see Table S1 in the supplemental material). The highest prevalence of ESBL-*E. coli* occurred in poultry farms (8/9), followed by pig farms (5/10) and cattle farms (3/10).

A cross-sectional survey was performed in April 2015 to isolate ESBL-*E. coli* from 97 prepackaged fresh meat products purchased in 11 major supermarkets in Cambridge, East of England (5 to 16 products per supermarket) (16). These originated from 11 different countries, although the majority (69/97 [71%]) were from the UK. Nineteen ESBL-*E. coli* isolates were cultured from 17/97 (18%) products (2 isolates sequenced from 2 samples representing different antibiograms), of which 16 isolates were from chicken originating from the UK ( $n = 12$ ), Ireland ( $n = 2$ ), Hungary (packaged in Ireland;  $n = 1$ ), or multiple origins (Brazil, Thailand, and Poland;  $n = 1$ ), and 3 were from turkey ( $n = 2$ ) and pork ( $n = 1$ ) from the UK.

We sequenced a total of 431 *E. coli* isolates from livestock ( $n = 411$ ) and meat ( $n = 20$ ), of which 155 were ESBL-*E. coli* (136 from livestock and 19 from meat).

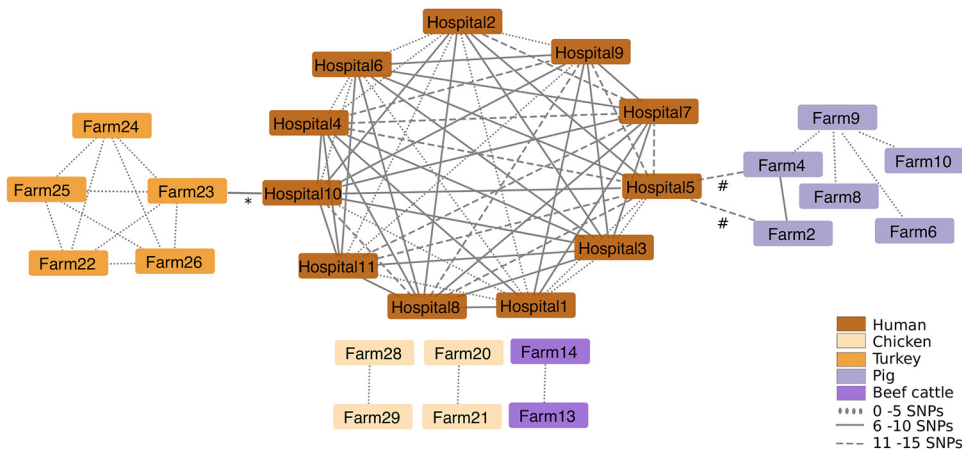
**Evaluation of an *E. coli* collection from patients with bloodstream infection.** A key study objective was to determine whether livestock and retail meat represented potential sources of *E. coli* associated with serious invasive disease in humans. In light of this, the human isolates used in the comparison with livestock isolates had caused bloodstream infection. A total of 1,517 open access *E. coli* genomes (142 ESBL and 1,375 non-ESBL) associated with bloodstream infections were retrieved (17, 18). Bloodstream isolates obtained from patients admitted to the Cambridge University Hospitals NHS Foundation Trust in the East of England between 2006 and 2012 ( $n = 424$ ) (17, 18) were combined with bloodstream isolates submitted to the British Society of Antimicrobial



**FIG 1** (a) Map of the United Kingdom showing the locations for the human clinical isolates, with the East of England highlighted by a red box. (b) Map of the East of England showing the locations of farms (images indicate livestock species). (c) Maximum likelihood tree based on SNPs in the core genes of 1,948 *E. coli* isolates cultured from livestock farms, retail meat, and patients with bloodstream infection.

Chemotherapy from 11 hospitals across England ( $n = 1,093$ ) between 2001 and 2011 (locations shown in Fig. 1a, and full isolate listing in Table S1) (17, 18). A potential limitation of this human isolate collection is that they might overrepresent hospital-acquired isolates, while a comparison of *E. coli* from livestock would require a comparison of community-acquired bacteria. Two analyses were undertaken to evaluate this possibility. First, we defined where the bloodstream infection was acquired for 1,303 cases for whom we had this information. This demonstrated that 886/1,303 (66%) cases were community associated. We then constructed a maximum likelihood tree of the invasive disease genomes to compare the phylogeny of isolates associated with community- versus health care-associated disease (Fig. S1). This demonstrated that genomes from the two categories were intermixed and distributed across the phylogeny, with no evidence of clustering by origin of infection. We concluded that our invasive collection was likely to include strong representation of *E. coli* carried by people in the community.

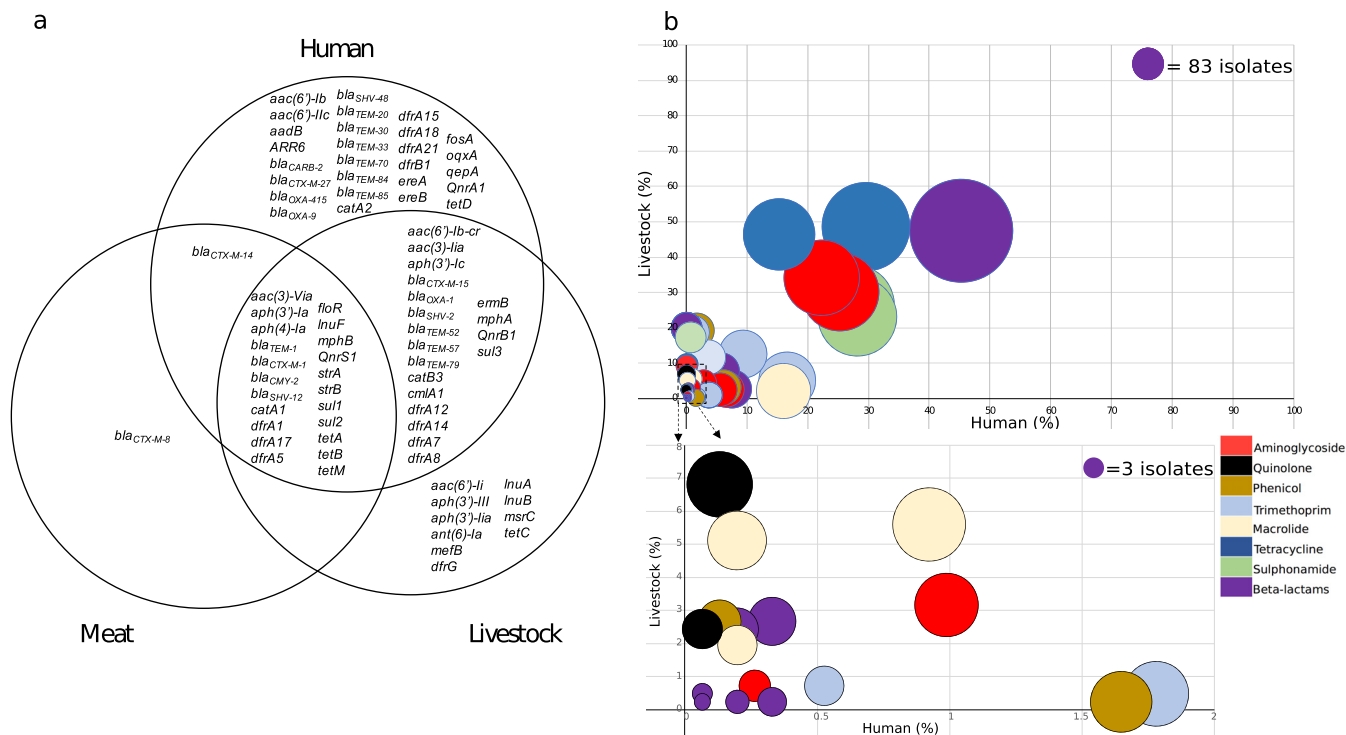
**Genomic comparison of *E. coli* from patients with bloodstream infection, livestock, and retail meat.** We combined and compared the 1,517 human invasive *E. coli* genomes with the 431 livestock-associated *E. coli* genomes. Analysis of the 1,948 genomes identified 331 multilocus sequence types (STs), 44 clonal complexes (CCs), and 149 singletons (STs that did not share alleles at six out of seven loci with any other ST). Most STs were only found in one source type, with 192 human-specific STs (1,261/1,517 isolates [83%]), 98 livestock-specific STs (225/411 isolates [55%]), and 4 meat-specific STs (4/20 isolates [20%]). Thirty-five STs contained isolates from both humans and livestock/meat ( $n = 431$ ), while 2 STs were found only in isolates from livestock and meat ( $n = 27$ ) (Table S1). The three most common STs associated with bloodstream infection were ST73, ST131, and ST95, while the three most common STs associated with livestock were ST10, ST117, and ST602 (Table S1), although the distribution of STs varied depending on the livestock host (Table S2). The greatest overlap in STs between the two reservoirs occurred in ST10 (3%, 16%, and 0% of human, livestock, and meat isolates, respectively), and ST117 (1%, 8%, and 20% of human, livestock, and meat isolates, respectively). Phylogroups varied depending on the source, with B2 predominating in the human invasive samples (1,026/1,517 [68%]), but B2 was identified in only 4 livestock samples (pig = 2, turkey = 1, beef cattle = 1) (Table S2).



**FIG 2** Network analysis of *E. coli* isolates cultured from livestock farms and patients with bloodstream infection. The results shown are limited to those isolate pairs identified in a pairwise comparison that differed by  $\leq 15$  or less SNPs in the core genome. The place of origin for each isolate pair are connected by lines, and the style of the line reflects the SNP distance. The asterisk indicates one ST69 human isolate from hospital 10 linked to two ST69 turkey isolates from farm 23 that differed by 10 and 12 SNPs, respectively. The number or hash sign indicates one ST1081 human isolate from hospital 5 linked to one ST1081 pig isolate from farm 4 (differed by 10 SNPs) and 2 ST1081 (probably duplicate) pig isolates from farm 2 that differed by 14 SNPs.

Phylogenies based on single nucleotide polymorphisms (SNPs) in the core (conserved) genomes of isolates representing CC10 ( $n = 149$ ) and CC117 ( $n = 64$ ) demonstrated that human isolates were intermixed with livestock isolates in CC10, but were generally distinct from livestock isolates in CC117 (Fig. S2 and S3). Pairwise SNP analysis demonstrated that the most closely related human/livestock isolate pairs were 85 and 96 SNPs different for CC10 and CC117, respectively. The estimated mutation rate for *E. coli* is one SNP/core genome/year (19, 20), and so CC10 and CC117 isolates in humans and livestock were not associated with recent transmission between the two groups. Combining the study CC117 isolates with 7 publicly available ST117 genomes (NCBI SRA accession numbers [ERR769196](https://www.ncbi.nlm.nih.gov/submitter/study/ERR769196), [ERR769195](https://www.ncbi.nlm.nih.gov/submitter/study/ERR769195), [ERR769183](https://www.ncbi.nlm.nih.gov/submitter/study/ERR769183), [ERR769169](https://www.ncbi.nlm.nih.gov/submitter/study/ERR769169), [SRR1314275](https://www.ncbi.nlm.nih.gov/submitter/study/SRR1314275), [SRR3410778](https://www.ncbi.nlm.nih.gov/submitter/study/SRR3410778), and [SRR3438297](https://www.ncbi.nlm.nih.gov/submitter/study/SRR3438297)) in a Bayesian phylogenetic analysis provided further evidence for the lack of recent transmission between human and livestock hosts in our study. The dated phylogeny revealed a UK cluster of 47 CC117 isolates (containing 44 turkey, 1 chicken, and 2 human isolates), for which the estimated time of most recent common ancestor (TMRCA) was 1989 (95% highest posterior density interval [HPD], 1979 to 1996), coinciding with the first global report of *bla*<sub>CTX-M-1</sub> (21). Of the 47 isolates, 36 (77%) carried *bla*<sub>CTX-M-1</sub>, which was uncommon in the rest of the bacterial population. All 36 isolates were from turkeys, representing a *bla*<sub>CTX-M-1</sub> poultry-associated lineage, for which the TMRCA was 2011 (95% HPD, 2010 to 2013) (Fig. S4), suggesting acquisition of *bla*<sub>CTX-M-1</sub> by this lineage between 1989 and 2011.

We then compared the genetic relatedness of the 431 livestock/meat *E. coli* isolates with the 1,517 *E. coli* isolates associated with human bloodstream infections. A maximum likelihood phylogenetic tree of the 1,948 genomes based on 277,533 core gene SNPs demonstrated high genetic diversity overall, with limited phylogenetic intermixing between isolates from humans and livestock (Fig. 1c). Pairwise SNP analysis between human- and livestock/meat-associated isolates demonstrated a median SNP distance of 41,658 (range, 10 to 47,819; interquartile range [IQR], 34,730 to 42,348), with 5 and 1 human isolates falling within 50 SNPs of livestock and meat, respectively (Fig. S5). Network analysis based on a range of SNP cutoffs captured just 2 (0.1%) human isolates (from hospitals in the South East and North West) that were within 15 SNPs of livestock isolates (2 pig isolates and 1 turkey isolate from three different farms [Fig. 2]). In contrast, we observed highly related isolates (0 to 5 SNPs) from the same animal species on different farms (Fig. 2).

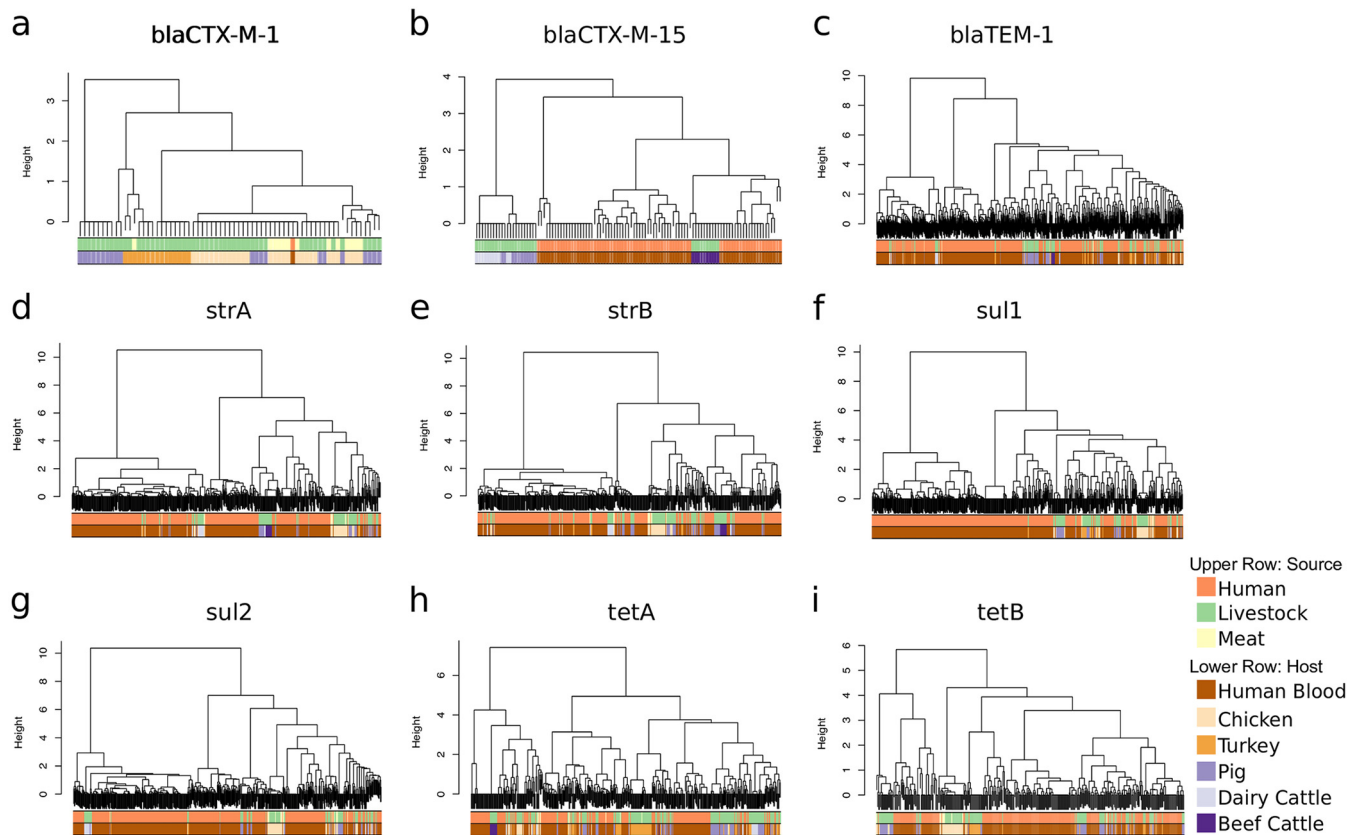


**FIG 3** (a) Venn diagram displaying antibiotic resistance genes identified in 1,948 *E. coli* isolates cultured from livestock, meat, and patients with bloodstream infections. (b) Bubble graph showing the proportion of genes shared between *E. coli* from humans and livestock. The bottom graph shows an expanded view of very low prevalence genes that are clustered in the lower left-hand corner of the graph. The size of each bubble represents the number of isolates that the gene was identified in. Bubbles are colored by antibiotic class.

The *E. coli* isolated during this study are likely to be an underrepresentation of the diversity of *E. coli* in the wider UK livestock population and the meat sold in supermarkets, which could reduce our power to detect a transmission event between livestock and humans or vice versa. To explore this further, we undertook additional analyses using UK livestock isolates in the public domain; specifically, the published livestock genomes held in Enterobase (<http://enterobase.warwick.ac.uk>), which comprised 51 genomes of isolates cultured between 1999 and 2013. The 24 STs in this collection were compared with the STs assigned to our invasive isolate collection from across the UK. A single ST in this new data set was also present in our bloodstream collection (ST398), but this had already been identified in our livestock collection. This provides further support that sharing of invasive *E. coli* lineages between humans and livestock in the UK is uncommon.

We evaluated and compared the accessory (non-conserved) genome of the 1,948 study isolates using principal-component analysis (PCA). Principal components 1 (PC1) and 2 (PC2), which accounted for 50.5% and 8.3% of the variation within the data, respectively, separated the collection into two main clusters (referred to as group 1 or group 2, respectively). Group 1 predominantly contained human isolates, and group 2 contained a mixture of human and livestock isolates (Fig. S6a). PCA also showed that isolates from the same STs clustered together and formed distinct subclusters within groups 1 and 2 (Fig. S6b). Table S3 lists the top 100 genes from PC1 and PC2 that were most strongly associated with group 1 or 2.

**Genetic analysis of antimicrobial resistance genes and associated mobile genetic elements.** Screening of the 1,948 isolates for accessory genes encoding antibiotic resistance revealed that 41 different resistance genes were present in isolates from both humans and livestock (Fig. 3a). The prevalence of resistance genes in the two groups varied considerably, with some predominating in the human or livestock reservoir only, while others were common in both (Fig. 3b). The seven most frequently shared genes



**FIG 4** Dendrograms of mobile genetic element clusters identified for *bla*<sub>CTX-M-1</sub> (a), *bla*<sub>CTX-M-15</sub> (b), *bla*<sub>TEM-1</sub> (c), *strA* (d), *strB* (e), *sul1* (f), *sul2* (g), *tetA* (h), and *tetB* (i) in livestock, humans, and retail meat.

(each present in >300 isolates) conferred resistance to beta-lactams (*bla*<sub>TEM-1</sub> = 882), sulfonamides (*sul2* = 530, *sul1* = 522), aminoglycosides (*strA* = 509, *strB* = 478), and tetracyclines (*tetA* = 423, *tetB* = 335). The predominant genes conferring resistance to extended-spectrum cephalosporins were *bla*<sub>CTX-M-15</sub> (human = 87, livestock = 32) and *bla*<sub>CTX-M-1</sub> (human = 1, livestock = 82, meat = 13). No carbapenemase or colistin resistance genes were detected.

To better understand whether genes encoding resistance in isolates from livestock and humans were carried by the same or different mobile genetic elements, contigs containing the specific gene of interest were extracted, and isolates were clustered using hierarchical cluster analysis based on contig presence/absence. This was performed for each of the seven most common resistance genes (*bla*<sub>TEM-1</sub>, *sul2*, *sul1*, *strA*, *strB*, *tetA*, and *tetB*) and the two most prevalent ESBL genes (*bla*<sub>CTX-M-15</sub> and *bla*<sub>CTX-M-1</sub>). Using a height cutoff value of 0 (corresponding to an identical mobile genetic element [MGE] contig carriage profile), clusters were screened for the presence of isolates derived from both human and livestock/meat. The majority of human isolates did not reside in clusters containing animal samples, the exception being *bla*<sub>CTX-M-1</sub> where a single human isolate carrying this gene clustered with 23 animal isolates (Fig. 4). For the nine resistance genes analyzed, between 0.6% and 9.8% of human isolates carrying a resistance gene shared a cluster with livestock isolates. The lowest frequency of relatedness was observed for *bla*<sub>TEM-1</sub> (3 clusters, involving isolates from humans = 4, livestock = 3), followed by *sul1* (3 clusters, humans = 3, livestock = 8), *strA* (6 clusters, humans = 11, livestock = 22), *tetA* (6 clusters, humans = 7, livestock = 19), *tetB* (5 clusters, humans = 11, livestock = 9), *bla*<sub>CTX-M-15</sub> (1 cluster, humans = 5, livestock = 10), *sul2* (9 clusters, humans = 41, livestock = 38, meat = 1), *strB* (8 clusters, humans = 33, livestock = 29), and *bla*<sub>CTX-M-1</sub> (1 cluster, humans = 1, livestock = 17,

meat = 6) (Table S4). Individual clusters often contained different STs (Table S4), which is indicative of horizontal transfer of mobile genetic elements between lineages.

Further characterization of *bla*<sub>CTX-M</sub> plasmids was undertaken using long-read sequencing. Two livestock-human isolate pairs positive for *bla*<sub>CTX-M-1</sub> or *bla*<sub>CTX-M-15</sub> were selected for sequencing using the PacBio RSII instrument. Illumina reads for the entire study collection were then mapped to the complete plasmid assemblies of these four isolates. The single *bla*<sub>CTX-M-1</sub>-positive human isolate contained an IncI1 *bla*<sub>CTX-M-1</sub> plasmid that was highly similar (>99% identity and ≥98% coverage) to 28 livestock isolates (chicken = 18, chicken meat = 8, pig = 2) belonging to four different STs. In contrast, the *bla*<sub>CTX-M-15</sub> plasmid in the livestock (E01) and human (D01) isolate pair were dissimilar (17% sequence shared at 99% identity [ID]) and had different replicon types (E01 = IncHI2, D01 = IncFIA and FII fusion). The human *bla*<sub>CTX-M-15</sub> plasmid (D01) was not identified in any other isolate (human or livestock), while the livestock *bla*<sub>CTX-M-15</sub> plasmid (E01) was found in other livestock isolates from the same pooled fecal sample from 1 beef farm.

We then investigated whether *bla*<sub>CTX-M-15</sub> could be shared on a smaller transposable element. A 7,926-bp region encoding *bla*<sub>CTX-M-15</sub> that was identical to a Tn3 transposon previously identified from *E. coli* plasmid [GU371928](#) (22) was detected in 22/32 (69%) livestock isolates (pig = 11, dairy cattle = 11) from 4 farms, and 3/87 (3%) of *bla*<sub>CTX-M-15</sub>-positive human isolates from 2 hospitals, one of which was located in the East of England. This 7,926-bp region was flanked by 5-bp direct repeats of TTTTA, indicating its potential for transfer between isolates.

## DISCUSSION

We investigated the prevalence and genetic relatedness of *E. coli* from livestock, meat, and humans in the East of England using a “One Health” approach. ESBL-*E. coli* was isolated from 55% of livestock farms, with a frequency of ESBL-*E. coli* in different livestock species that was consistent with previous findings (23). In addition, ESBL-*E. coli* bacteria were found in 18% of prepackaged fresh meat products. The high prevalence of ESBL-*E. coli* in chicken meat (16/30 [53%]) is similar to previous studies conducted in the UK and the Netherlands (12, 24, 25). However, *E. coli* from livestock were not closely related to isolates causing human disease in our region, suggesting that livestock are not a direct source of infecting isolates and that human invasive *E. coli* are not being shared with livestock. *E. coli* phylogroup B2 was most frequently associated with human invasive samples (68%) as previously reported (26), but was rarely identified in livestock (1%), providing further evidence for distinct populations associated with invasive human disease and livestock. In contrast, highly related isolates were identified between the same livestock species on different farms. Previous studies in the Netherlands that compared isolates from clinical and livestock sources using MLST indicated that the same ST could be isolated from humans and livestock (12–14, 27). We replicated this finding for CC10 and CC117, but using the more discriminatory sequence-based analysis identified that isolates from the two reservoirs were genetically distinct. A study of cephalosporin-resistant *E. coli* in the Netherlands (15) reported genetic heterogeneity between human and poultry-associated isolates but closely related isolates from farmers and their pigs. Here, we included ESBL-positive and non-ESBL *E. coli*, an important feature of the study since the majority of *E. coli* human infections in the UK are due to non-ESBL *E. coli* (17).

Screening of *E. coli* isolates from livestock, meat, and humans with serious infections revealed the frequency of antimicrobial-resistant genes in each reservoir and confirmed the presence of similar antimicrobial resistance genes in both livestock and humans, including *bla*<sub>TEM-1</sub>, *sul2*, *sul1*, *strA*, *strB*, *tetA*, *tetB*, *bla*<sub>CTX-M-15</sub>, and *bla*<sub>CTX-M-1</sub>. These genes confer resistance to four antibiotic classes, all of which are used in both livestock and humans (28). This confirms their ubiquitous distribution but does not provide evidence for recent transfer of genes between the two reservoirs. To address this, we hypothesized that recent sharing would be associated with transmission via the same or highly related mobile genetic elements (MGEs), as previously suggested for ESBL genes (15).

Previous studies have highlighted the challenge in reconstructing plasmids and other mobile elements encoding resistance genes from whole-genome sequencing (29, 30), hindering our understanding of the transmission dynamics of resistance genes. We developed an approach to detect and genetically compare mobile elements across our large study collection, with validation of findings for ESBLs using long-read sequencing. The findings from this were consistent with predominantly distinct mobile elements between livestock and humans, with an estimated 69/1,517 (5%) human isolates potentially sharing closely related antimicrobial resistance-associated mobile elements with those found in livestock.

Our study has several limitations. We acknowledge that the *E. coli* from humans predated the surveys of farms and retail meat but took account of this by identifying relatedness based on a 0 to 15 SNP cutoff given the estimated *E. coli* mutation rate of 1 SNP/core genome/year (19, 20). We did not include all possible sources of *E. coli* for humans (for example, vegetables, fruits, and pets), although a recent study found no *E. coli* with *bla*<sub>CTX-M-15</sub> (the dominant human ESBL type) in retail meat, fruit, and vegetables in five UK regions (24). Additional studies are required to understand whether our findings will be reproduced in other geographical areas, to determine other sources of invasive lineages such as wastewater or recreational waters, to better understand within-host diversity in livestock, to differentiate between historical and recent transmission events by collecting data over a longer time period, and to identify whether livestock are a source for other types of infection in humans such as urinary tract infections.

In conclusion, this study has not generated evidence to indicate that *E. coli* causing severe human infections in our region were derived recently from livestock, with host-specific *E. coli* lineages identified from hospitals versus farms. We identified limited sharing of antimicrobial resistance genes between livestock and humans based on long-read sequencing and analysis of mobile genetic elements. Further investigations are required to pursue the identification of the source of *E. coli* and resistance genes in isolates associated with severe human disease.

## MATERIALS AND METHODS

**Sampling of livestock feces and retail meat.** A cross-sectional survey was performed between August 2014 and April 2015 to isolate *E. coli* at 20 livestock farms (10 cattle and 10 pig) in the East of England. A pooled sample of approximately 50 g of freshly passed fecal material was collected from each major area in a given farm (such as different pens) using a sterile scoop (Sterilin X400; Thermo Fisher Scientific, Loughborough, United Kingdom). Each pooled sample was placed into a dry sterile 150-ml container (Sterilin polystyrene containers; Fisher Scientific). A median of 4 samples (range, 1 to 5) were taken from each cattle farm, and a median of 4.5 samples (range, 3 to 9) were taken from each pig farm, resulting in a total of 85 pooled samples (34 cattle and 51 pig). In addition, cecal contents were collected from 2 deceased pigs at the time of necropsy.

Poultry reared at nine farms (4 chicken and 5 turkey) in the East of England were sampled at two abattoirs between February and April 2015. Two sample types were taken for each farm: (i) pooled feces with a total weight of approximately 50 g from 10 to 20 transportation crates immediately after the livestock were removed; (ii) pools of cecal material from up to 10 birds after slaughter. Each sample was taken using a sterile scoop, and a sterile surgical scalpel was used for each cecal dissection. A median of 4 (range, 2 to 4) cecal pools and 4 (range, 3 to 4) fecal pools were collected from livestock from each chicken farm, and a median of 1.5 (range, 1 to 2) cecal pools and 2.5 (range, 2 to 3) fecal pools were collected from livestock from each turkey farm. This resulted in a total of 49 pooled samples (29 chicken and 20 turkey). All samples were immediately refrigerated at 4°C upon return to the laboratory and processed on the same day.

In April 2015, 97 retail meat samples (beef [15], chicken [30], pork [42], turkey [7], venison [1], veal [1], mixed minced pork and beef [1]) were purchased from 11 supermarkets in Cambridge, UK, with 5 to 16 meat products collected from each supermarket that were selected to capture diversity in the products available. The country of origin for each meat product was recorded, and where multiple countries/regions were stated on the packaging, all names were recorded.

**Microbiology.** Pooled fecal samples were diluted 1:1 with sterile phosphate-buffered saline and mixed vigorously, and 100- $\mu$ l aliquots were plated onto Chromocult coliform agar (VWR, Leuven, Belgium) and *Brilliance* ESBL agar (Oxoid, Basingstoke, UK), which are selective chromogenic agars that support the growth of coliforms and ESBL-producing organisms, respectively. Agar plates were incubated at 37°C for 48 h in air prior to inspection. Enrichment cultures were also used to detect ESBL-producing *E. coli* by adding 1 ml of fecal preparation to 9 ml of tryptic soy broth containing 20  $\mu$ g cefpodoxime and incubating for 24 h in a shaking incubator (150 rpm) at 37°C in air, before 100  $\mu$ l was plated onto *Brilliance* ESBL agar and incubated for 48 h in air. Numerous *E. coli* colonies were picked from primary



cultures of positive pooled stool samples based on diversity in colonial morphology. Up to 32 colonies of presumptive *E. coli* based on colony morphology were picked from samples taken on each farm.

Preparation and culture of meat samples followed the European standard ISO 6887–2:2003. All exterior packaging was disinfected with alcohol prior to removal of meat. A 5-g sample of meat was aseptically removed, added to 45 ml peptone broth, and homogenized using a Stomacher paddle blender (Stomacher80 Laboratory System, Seward Ltd., UK) for 2 min. Samples were transferred into 50-ml Falcon tubes and incubated in a shaking incubator for 24 h at 150 rpm at 37°C. After incubation, all samples were plated onto *Brilliance* ESBL agar and incubated at 37°C for 48 h. In addition, swabs were obtained from whole chicken carcasses and incubated in 3 ml brain heart infusion (BHI) broth (FIO-QSwabs; Copan Italia spa, Brescia, Italy) in a shaking incubator for 24 h at 150 rpm at 37°C. Following incubation, 100  $\mu$ l was plated onto *Brilliance* ESBL agar and incubated as described before. One colony of presumptive ESBL *E. coli* was picked from each *Brilliance* ESBL agar for further evaluation, with the exception of two meat samples, where two colonies were picked to represent each of two distinct colony morphologies.

All bacterial colonies suspected to be *E. coli* were identified to the species level using matrix-assisted laser desorption ionization–time of flight mass spectrometry (Bruker Daltonik, Bremen, Germany). Antimicrobial susceptibility was defined for each *E. coli* colony using the Vitek2 system (bioMérieux, Marcy l’Etoile, France) with the AST-N206 card and calibrated against EUCAST breakpoints ([http://www.eucast.org/clinical\\_breakpoints/](http://www.eucast.org/clinical_breakpoints/)).

**DNA sequencing.** Bacterial genomic DNA was extracted using the QIAextractor (Qiagen, Valencia, CA, USA) according to the manufacturer’s instructions. Library preparation was conducted according to the Illumina protocol and sequenced on an Illumina HiSeq2000 (Illumina, San Diego, CA, USA) with 100-cycle paired-end runs. Sequence data were retrieved for a further 1,517 open access *E. coli* isolates associated with bloodstream infections (17, 18). Of these, 424 were isolated between January 2006 and December 2012 at the Cambridge University Hospitals NHS Foundation Trust, and 1,093 were submitted to the British Society for Antimicrobial Chemotherapy Bacteraemia Resistance Surveillance Project by 11 UK hospitals between 2001 and 2011 (for details, see [www.bsacsurv.org](http://www.bsacsurv.org) and Table S1 in the supplemental material) (31). Previous description and analysis of these genomes (17, 18) did not include comparisons with isolates from livestock or meat.

**Genome assembly, annotation, and multilocus sequence typing.** Taxonomic identity was assigned using Kraken (32). One isolate (VREC0294) was found to not be *E. coli* and was excluded from further analysis. Multiple assemblies were created using VelvetOptimiser v2.2.5 (33) and Velvet v1.2 (34). An assembly improvement step was applied to the assembly with the best N50, contigs were scaffolded using SSPACE (35), and sequence gaps were filled using GapFiller (36). Assemblies were annotated using Prokka v1.5 (37), and genus-specific databases from RefSeq (38). Multilocus sequence types (MLSTs) were identified using the MLST sequence archive (<https://enterobase.warwick.ac.uk>). Sequence types (STs) were classified into clonal complexes using the eBURST V3 algorithm (<http://eburst.mlst.net/>). *In silico* *E. coli* phylotyping was performed using ClermonTyping (39).

**Pan-genome analysis.** The pan-genome was calculated for all 1,948 isolates using Roary (40), with a 90% ID cutoff and genes classified as “core” if they were present in at least 99% of isolates. A maximum likelihood tree was created using RAxML (41) based on single nucleotide polymorphisms (SNPs) in the core genes. Principal-component analysis was performed across the 1,948 isolates based on the accessory genes from Roary using R. A Spearman rho correlation analysis was performed on the principal components, the gene absence/presence data, and the ST and source of isolation metadata.

**Phylogeny-based analysis of individual lineages.** Lineage-specific analyses were performed by mapping the sequence reads for isolates belonging to clonal complex 10 (CC10) and CC117 to an *E. coli* reference genome from the same clonal complex using SMALT 0.7.4 (<http://www.sanger.ac.uk/resources/software/smalt/>). *E. coli* MG1655 K-12 (ENA accession number U00096.2) was used as the reference genome for CC10, and a *de novo* assembly of the CC117 study isolate (ENA accession number ERR1204146) with the lowest number of contigs was used as the reference genome for CC117 as no reference genomes were available. To create a “core” genome, mobile genetic elements (MGEs) were identified using gene annotation, PHAST ([phast.wishartlab.com](http://phast.wishartlab.com)), and BLAST (<https://blast.ncbi.nlm.nih.gov>) and removed, together with contigs less than 500 bp in length. Recombination was removed using Gubbins (42). A maximum likelihood phylogeny was created using RAxML (41) with 100 bootstraps and a midpoint root. Genetic diversity was calculated based on pairwise differences in SNPs in the core genomes using an in-house script. Visualization of phylogenetic trees was performed using iTOL (<http://itol.embl.de>) (43) and FigTree v 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

**Bayesian Evolutionary Analysis Sampling Trees (BEAST).** All published genomes for CC117 in the *Enterobase* online database (<https://enterobase.warwick.ac.uk>, accessed 10 June 2016) that had been generated on an Illumina instrument and had the country, year, and source of isolation available were identified (ERR769196, ERR769195, ERR769183, ERR769169, SRR1314275, SRR3410778, and SRR3438297). These were mapped to the CC117 reference using SMALT, combined with the CC117 study isolates, and mobile elements and recombination were removed as before. Dating of this lineage was completed using BEAST v.1-8 (44). BEAST v.1-8 was run using the Hasegawa, Kishino, and Yano (HKY) and gamma substitution model. We compared combinations of three population size change models (constant, exponential, and Bayesian skyline plot) and three molecular clock models (strict, exponential, and uncorrelated lognormal). A Bayesian skyline population model and an uncorrelated lognormal molecular clock were selected based on Bayes factors calculated from path sampling and stepping stone sampling (44, 45).

**Detection of antimicrobial resistance and mobile elements.** Acquired genes encoding antibiotic resistance were identified using Antibiotic Resistance Identification By Assembly (ARIBA), comparing the study genomes against an in-house curated version of the Resfinder database (46–48) consisting of 2,015 known resistance gene variants. Genes were classified as present using an identity of 90% nucleotide similarity. Genes reported as fragmented, partial, or interrupted were excluded.

For all isolates positive for the *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>TEM-1</sub>, *sul1*, *strA*, *strB*, *sul2*, *tetA*, and *tetB* genes, whole-genome assemblies were screened to identify the contig carrying the antimicrobial resistance (AMR) gene using the blastn application (49) with the AMR gene sequence as the query sequence. The identified contigs were then aligned against a previously curated database of complete *Enterobacteriaceae* plasmids (50) in order to filter out sequences representing *E. coli* chromosome fragments. For each AMR gene, a database containing unique AMR-carrying contigs was created using cd-hit-est (51) based on 90% identity cutoff. To determine contig carriage, for each of the eight AMR genes, all isolates positive for that gene were mapped against the respective gene-specific database of contigs using short-read sequencing typing (SRST2) using a minimum 90% coverage cutoff. The contig presence/absence data were converted into a distance matrix, and hierarchical clustering was performed using R function *hclust* and the *ward.D2* method.

To examine plasmids carrying the ESBL genes *bla*<sub>CTX-M-15</sub> and *bla*<sub>CTX-M-1</sub>, two pairs of livestock and human isolates were selected for sequencing on the PacBio RSII instrument (Pacific Biosciences, Menlo Park, CA, USA) (*n* = 4), and *in silico* PCR was used to perform plasmid incompatibility group/replicon typing (52). These two genes were selected, as they were the most prevalent ESBL genes found in the 1,948 isolates. The pair of *bla*<sub>CTX-M-15</sub>-positive isolates was selected, as they contained 5 identical genes encoding antimicrobial resistance. The single human *bla*<sub>CTX-M-1</sub>-positive isolate in the collection was selected, and a livestock isolate with the most similar resistance gene profile was selected, with both isolates containing 3 identical genes encoding antimicrobial resistance. DNA was extracted using the phenol-chloroform method (53) and sequenced using the PacBio RS II instrument. Sequence reads were assembled *de novo* with HGAP v3 (54) within the SMRT Analysis version 2.3.0 software (<https://www.pacb.com/products-and-services/analytical-software/smrt-analysis/>), circularized using Circlator v1.1.3 (55) and Minimus 2 (56), and polished using the PacBio RS\_Resequencing protocol and Quiver v1 (54). Fully assembled plasmids were compared using WebACT (<http://www.webact.org>) and BLASTn (<https://blast.ncbi.nlm.nih.gov>).

**Ethical approval.** The study protocol was approved by the Cambridge University Hospitals NHS Foundation Trust Research and Development Department (reference A093285) and the National Research Ethics Service East of England Ethics Committee (reference 12/EE/0439 and 14/EE/1123).

**Data availability.** Sequence data for all isolates have been submitted to the European Nucleotide Archive ([www.ebi.ac.uk/ena](http://www.ebi.ac.uk/ena)) under study accession number PRJEB4681 (all human *E. coli*), PRJEB8774 (non-ESBL-producing *E. coli* from livestock), and PRJEB8776 (ESBL-producing *E. coli* from livestock and meat), with the accession numbers for individual isolates listed in Table S1 in the supplemental material.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.02693-18>.

**FIG S1**, TIF file, 1.2 MB.

**FIG S2**, TIF file, 0.6 MB.

**FIG S3**, TIF file, 0.4 MB.

**FIG S4**, TIF file, 0.4 MB.

**FIG S5**, TIF file, 0.5 MB.

**FIG S6**, TIF file, 0.6 MB.

**TABLE S1**, XLSX file, 0.1 MB.

**TABLE S2**, PDF file, 0.1 MB.

**TABLE S3**, XLSX file, 0.01 MB.

**TABLE S4**, XLSX file, 0.1 MB.

## ACKNOWLEDGMENTS

We thank the Wellcome Sanger Institute core library construction, sequence and informatics teams, and the Pathogen Informatics team. We thank the staff at farms and abattoirs for assistance in sample collection and Elizabeth Lay for laboratory support during the meat survey. We thank Olivier Restif for statistical advice. The flocked swabs used in the meat survey were donated by Copan Italia spa.

This publication presents independent research supported by the Health Innovation Challenge Fund (WT098600, HICF-T5-342), a parallel funding partnership between the Department of Health and Wellcome Trust. This project was also funded by a grant awarded to the Wellcome Trust Sanger Institute (098051). T.G. is a Wellcome Trust Research Training Fellow (103387/Z/13/Z). C.L. is a Wellcome Trust Sir Henry Postdoctoral Fellow (110243/Z/15/Z). F.C. is a Wellcome Trust Sir Henry Postdoctoral Fellow

(201344/Z/16/Z). D.J. and J.P. are funded by Wellcome Trust grant 098051. M.D.G. is funded by the Medical Research Council (MR/K021133/1).

The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health or Wellcome Trust.

C.L., T.G., and S.J.P. designed the study. C.L. and T.G. developed study protocols. C.L. and T.G. collected farm samples. N.H., T.G., and M.H. found sources of retail meat and performed microbiological culture of retail meat samples. J.H.-G., P.W., M.R., and M.H. obtained access to farms and abattoirs. C.L., B.B., and P.N. performed bacterial identification, susceptibility testing, and phenotypic testing. C.L., D.J., K.E.R., F.C., and M.D.G. performed bioinformatics analyses and interpretation. C.L., D.J., F.C., and M.D.G. produced the figures. N.M.B. and C.H. contributed sequence data. C.L. and S.J.P. wrote the manuscript. J.P. and S.J.P. were responsible for supervision and management of the study. All authors read and approved the final manuscript.

## REFERENCES

- European Centre for Disease Prevention and Control. 2013. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2011–2012. ECDC Surveillance Report. European Centre for Disease Prevention and Control, Stockholm, Sweden. <http://ecdc.europa.eu/en/publications/Publications/healthcare-associated-infections-antimicrobial-use-PPS.pdf>.
- Pitout JDD. 2012. Extraintestinal pathogenic *Escherichia coli*: a combination of virulence with antibiotic resistance. *Front Microbiol* 3:9. <https://doi.org/10.3389/fmicb.2012.00009>.
- European Centre for Disease Prevention and Control. 2017. Antimicrobial resistance surveillance in Europe 2015. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). European Centre for Disease Prevention and Control, Stockholm, Sweden.
- Gagliotti C, Balode A, Baquero F, Degener J, Grundmann H, Gür D, Jarlier V, Kahlmeter G, Monen J, Monnet DL, Rossolini GM, Suetens C, Weist K, Heuer O. EARS-Net Participants (Disease Specific Contact Points for AMR). 2011. *Escherichia coli* and *Staphylococcus aureus*: bad news and good news from the European Antimicrobial Resistance Surveillance Network (EARS-Net, formerly EARSS), 2002 to 2009. *Euro Surveill* 16:19819.
- Thaden JT, Fowler VG, Sexton DJ, Anderson DJ. 2016. Increasing incidence of extended-spectrum  $\beta$ -lactamase-producing *Escherichia coli* in community hospitals throughout the Southeastern United States. *Infect Control Hosp Epidemiol* 37:49–54. <https://doi.org/10.1017/ice.2015.239>.
- Poolman JT, Wacker M. 2016. Extraintestinal pathogenic *Escherichia coli*, a common human pathogen: challenges for vaccine development and progress in the field. *J Infect Dis* 213:6–13. <https://doi.org/10.1093/infdis/jiv429>.
- Schwaber MJ, Carmeli Y. 2007. Mortality and delay in effective therapy associated with extended-spectrum  $\beta$ -lactamase production in *Enterobacteriaceae* bacteraemia: a systematic review and meta-analysis. *J Antimicrob Chemother* 60:913–920. <https://doi.org/10.1093/jac/dkm318>.
- Tumbarello M, Sanguinetti M, Montuori E, Trecarichi EM, Posteraro B, Fiori B, Citton R, D'Inzeo T, Fadda G, Cauda R, Spanu T. 2007. Predictors of mortality in patients with bloodstream infections caused by extended-spectrum- $\beta$ -lactamase-producing *Enterobacteriaceae*: importance of inadequate initial antimicrobial treatment. *Antimicrob Agents Chemother* 51:1987–1994. <https://doi.org/10.1128/AAC.01509-06>.
- Rottier WC, Ammerlaan HSM, Bonten MJM. 2012. Effects of confounders and intermediates on the association of bacteraemia caused by extended-spectrum  $\beta$ -lactamase-producing *Enterobacteriaceae* and patient outcome: a meta-analysis. *J Antimicrob Chemother* 67:1311–1320. <https://doi.org/10.1093/jac/dks065>.
- Mediavilla JR, Patrawalla A, Chen L, Chavda KD, Mathema B, Vinnard C, Dever LL, Kreiswirth BN. 2016. Colistin- and carbapenem-resistant *Escherichia coli* harboring mcr-1 and blaNDM-5, causing a complicated urinary tract infection in a patient from the United States. *mBio* 7:e01191–16. <https://doi.org/10.1128/mBio.01191-16>.
- Albiger B, Glasner C, Struelens MJ, Grundmann H, Monnet DL, European Survey of Carbapenemase-Producing *Enterobacteriaceae* (EuSCAPE) working group. 2015. Carbapenemase-producing *Enterobacteriaceae* in Europe: assessment by national experts from 38 countries, May 2015. *Euro Surveill* 20:30062. <https://doi.org/10.2807/1560-7917.ES.2015.20.45.30062>.
- Leverstein-van Hall MA, Dierikx CM, Stuart JC, Voets GM, van den Munckhof MP, van Essen-Zandbergen A, Platteel T, Fluit AC, van de Sande-Bruinsma N, Scharinga J, Bonten MJM, Mevius DJ. 2011. Dutch patients, retail chicken meat and poultry share the same ESBL genes, plasmids and strains. *Clin Microbiol Infect* 17:873–880. <https://doi.org/10.1111/j.1469-0691.2011.03497.x>.
- Kluytmans JAJW, Overvest ITMA, Willemsen I, Kluytmans-van den Bergh MFQ, van der Zwaluw K, Heck M, Rijnsburger M, Vandenbroucke-Grauls CMJE, Savelkoul PHM, Johnston BD, Gordon D, Johnson JR. 2013. Extended-spectrum  $\beta$ -lactamase-producing *Escherichia coli* from retail chicken meat and humans: comparison of strains, plasmids, resistance genes, and virulence factors. *Clin Infect Dis* 56:478–487. <https://doi.org/10.1093/cid/cis929>.
- Overvest I, Willemsen I, Rijnsburger M, Eustace A, Xu L, Hawkey P, Heck M, Savelkoul P, Vandenbroucke-Grauls C, van der Zwaluw K, Huijsdens X, Kluytmans J. 2011. Extended-spectrum  $\beta$ -lactamase genes of *Escherichia coli* in chicken meat and humans, the Netherlands. *Emerg Infect Dis* 17:1216–1222. <https://doi.org/10.3201/eid1707.110209>.
- de Been M, Lanza VF, de Toro M, Scharringa J, Dohmen W, Du Y, Hu J, Lei Y, Li N, Tooming-Klunderud A, Heederik DJJ, Fluit AC, Bonten MJM, Willemsen RJL, de la Cruz F, van Schaik W. 2014. Dissemination of cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by specific plasmid lineages. *PLoS Genet* 10:e1004776. <https://doi.org/10.1371/journal.pgen.1004776>.
- Gouliouris T, Raven K, Ludden C, Blane B, Corander J, Horner C, Hernandez-Garcia J, Wood P, Hadjirin N, Radakovic M, Holmes M, de Goffau M, Brown N, Parkhill J, Peacock S. 2018. Genomic surveillance of *Enterococcus faecium* reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. *mBio* 9:e01780-18. <https://doi.org/10.1128/mBio.01780-18>.
- Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, Martin V, Peacock SJ, Parkhill J. 2017. Systematic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res* 27:1437–1449. <https://doi.org/10.1101/gr.216606.116>.
- Brodrick HJ, Raven KE, Kallonen T, Jamrozny D, Blane B, Brown NM, Martin V, Török ME, Parkhill J, Peacock SJ. 2017. Longitudinal genomic surveillance of multidrug-resistant *Escherichia coli* carriage in a long-term care facility in the United Kingdom. *Genome Med* 9:70. <https://doi.org/10.1186/s13073-017-0457-6>.
- Reeves PR, Liu B, Zhou Z, Li D, Guo D, Ren Y, Clabots C, Lan R, Johnson JR, Wang L. 2011. Rates of mutation and host transmission for an *Escherichia coli* clone over 3 years. *PLoS One* 6:e26907. <https://doi.org/10.1371/journal.pone.0026907>.
- Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, Sebra R, Turner P, Anson LW, Kasarskis A, Batty EM, Kos V, Wilson DJ, Phetsouvanh R, Wyllie D, Sokurenko E, Manges AR, Johnson TJ, Price LB, Peto TEA, Johnson JR, Didelot X, Walker AS, Crook DW. Modernizing Medical Microbiology Informatics Group (MMMIG). 2016. Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *mBio* 7:e02162. <https://doi.org/10.1128/mBio.02162-15>.
- Bauernfeind A, Schweighart S, Grimm H. 1990. A new plasmidic cefo-

- taximase in a clinical isolate of *Escherichia coli*. *Infection* 18:294–298. <https://doi.org/10.1007/BF01647010>.
22. Smet A, Van Nieuwerburgh F, Vandekerckhove TTM, Martel A, Deforce D, Butaye P, Haesebrouck F. 2010. Complete nucleotide sequence of CTX-M-15-plasmids from clinical *Escherichia coli* isolates: insertional events of transposons and insertion sequences. *PLoS One* 5:e11202. <https://doi.org/10.1371/journal.pone.0011202>.
  23. Horton RA, Randall LP, Snary EL, Cockrem H, Lotz S, Wearing H, Duncan D, Rabie A, McLaren I, Watson E, La Ragione RM, Coldham NG. 2011. Fecal carriage and shedding density of CTX-M extended-spectrum  $\beta$ -lactamase-producing *Escherichia coli* in cattle, chickens, and pigs: implications for environmental contamination and food production. *Appl Environ Microbiol* 77:3715–3719. <https://doi.org/10.1128/AEM.02831-10>.
  24. Randall LP, Lodge MP, Elviss NC, Lemma FL, Hopkins KL, Teale CJ, Woodford N. 2017. Evaluation of meat, fruit and vegetables from retail stores in five United Kingdom regions as sources of extended-spectrum beta-lactamase (ESBL)-producing and carbapenem-resistant *Escherichia coli*. *Int J Food Microbiol* 241:283–290. <https://doi.org/10.1016/j.ijfoodmicro.2016.10.036>.
  25. Belmar Campos C, Fenner I, Wiese N, Lensing C, Christner M, Rohde H, Aepfelbacher M, Fenner T, Hentschke M. 2014. Prevalence and genotypes of extended spectrum beta-lactamases in *Enterobacteriaceae* isolated from human stool and chicken meat in Hamburg, Germany. *Int J Med Microbiol* 304:678–684. <https://doi.org/10.1016/j.ijmm.2014.04.012>.
  26. Nash JH, Villegas A, Kropinski AM, Aguilar-Valenzuela R, Konczyk P, Mascarenhas M, Ziebell K, Torres AG, Karmali MA, Coombes BK. 2010. Genome sequence of adherent-invasive *Escherichia coli* and comparative genomic analysis with other *E. coli* pathotypes. *BMC Genomics* 11:667. <https://doi.org/10.1186/1471-2164-11-667>.
  27. Davis GS, Waits K, Nordstrom L, Weaver B, Aziz M, Gauld L, Grande H, Bigler R, Horwinski J, Porter S, Stegger M, Johnson JR, Liu CM, Price LB. 2015. Intermingled *Klebsiella pneumoniae* populations between retail meats and human urinary tract infections. *Clin Infect Dis* 61:892–899. <https://doi.org/10.1093/cid/civ428>.
  28. O'Neill J. 2015. Antimicrobials in agriculture and the environment: reducing unnecessary use and waste. The Review on Antimicrobial Resistance, London, United Kingdom.
  29. Orlek A, Stoesser N, Anjum MF, Doumith M, Ellington MJ, Peto T, Crook D, Woodford N, Walker AS, Phan H, Sheppard AE. 2017. Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front Microbiol* 8:182. <https://doi.org/10.3389/fmicb.2017.00182>.
  30. Arredondo-Alonso S, van Schaik W, Willems RJ, Schurch AC. 2017. On the (im)possibility to reconstruct plasmids from whole genome short-read sequencing data. *bioRxiv* <https://doi.org/10.1101/086744>.
  31. Reynolds R, Hope R, Williams L. 2008. Survey, laboratory and statistical methods for the BSAC Resistance Surveillance Programmes. *J Antimicrob Chemother* 62:ii15–ii28. <https://doi.org/10.1093/jac/dkn349>.
  32. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
  33. Gladman S, Seemann T, Victorian Bioinformatics Consortium. 2008. Velvet Optimiser: for automatically optimising the primary parameter options for the Velvet de novo sequence assembler. Victorian Bioinformatics Consortium, Monash University, Melbourne, Australia.
  34. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
  35. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
  36. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol* 13:R56. <https://doi.org/10.1186/gb-2012-13-6-r56>.
  37. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
  38. Pruitt KD, Tatusova T, Brown GR, Maglott DR. 2012. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* 40:D130–D135. <https://doi.org/10.1093/nar/gkr1079>.
  39. Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. 2018. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb Genom* 4:e000192. <https://doi.org/10.1099/mgen.0.000192>.
  40. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
  41. Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
  42. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 43:e15. <https://doi.org/10.1093/nar/gku1196>.
  43. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* <https://doi.org/10.1093/nar/gkw290>.
  44. Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214. <https://doi.org/10.1186/1471-2148-7-214>.
  45. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol* 29:2157–2167. <https://doi.org/10.1093/molbev/mss084>.
  46. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
  47. Reuter S, Ellington MJ, Cartwright EJP, Köser CU, Török ME, Gouliouris T, Harris SR, Brown NM, Holden MTG, Quail M, Parkhill J, Smith GP, Bentley SD, Peacock SJ. 2013. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med* 173:1397–1404. <https://doi.org/10.1001/jamainternmed.2013.7734>.
  48. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, Harris SR. 2017. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *bioRxiv* <https://doi.org/10.1101/118000>.
  49. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
  50. Orlek A, Phan H, Sheppard AE, Doumith M, Ellington M, Peto T, Crook D, Walker AS, Woodford N, Anjum MF, Stoesser N. 2017. A curated dataset of complete *Enterobacteriaceae* plasmids compiled from the NCBI nucleotide database. *Data Brief* 12:423–426. <https://doi.org/10.1016/j.dib.2017.04.024>.
  51. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
  52. Carattoli A, Bertini A, Villa L, Falbo V, Hopkins KL, Threlfall EJ. 2005. Identification of plasmids by PCR-based replicon typing. *J Microbiol Methods* 63:219–228. <https://doi.org/10.1016/j.mimet.2005.03.018>.
  53. Hull RA, Gill RE, Hsu P, Minshew BH, Falkow S. 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. *Infect Immun* 33:933–938.
  54. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
  55. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
  56. Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64. <https://doi.org/10.1186/1471-2105-8-64>.