

# Dynamics of extended-spectrum cephalosporin resistance genes in *Escherichia coli* from Europe and North America

Received: 18 February 2022

Accepted: 11 November 2022

Published online: 12 December 2022

 Check for updates

Roxana Zamudio <sup>1</sup>, Patrick Boerlin<sup>2</sup>, Racha Beyrouthy<sup>3,4</sup>, Jean-Yves Madec<sup>5</sup>, Stefan Schwarz <sup>6,7</sup>, Michael R. Mulvey<sup>8</sup>, George G. Zhanel<sup>9</sup>, Ashley Cormier<sup>2</sup>, Gabhan Chalmers <sup>2</sup>, Richard Bonnet<sup>3,4</sup>, Marisa Haenni<sup>5</sup>, Inga Eichhorn<sup>6,7</sup>, Heike Kaspar<sup>10</sup>, Raquel Garcia-Fierro<sup>5</sup>, James L. N. Wood <sup>11</sup> & Alison E. Mather <sup>1,12</sup> 

Extended-spectrum cephalosporins (ESCs) are critically important antimicrobial agents for human and veterinary medicine. ESC resistance (ESC-R) genes have spread worldwide through plasmids and clonal expansion, yet the distribution and dynamics of ESC-R genes in different ecological compartments are poorly understood. Here we use whole genome sequence data of Enterobacterales isolates of human and animal origin from Europe and North America and identify contrasting temporal dynamics. AmpC  $\beta$ -lactamases were initially more dominant in North America in humans and farm animals, only later emerging in Europe. In contrast, specific extended-spectrum  $\beta$ -lactamases (ESBLs) were initially common in animals from Europe and later emerged in North America. This study identifies differences in the relative importance of plasmids and clonal expansion across different compartments for the spread of different ESC-R genes. Understanding the mechanisms of transmission will be critical in the design of interventions to reduce the spread of antimicrobial resistance.

Third- and fourth-generation cephalosporins are broad-spectrum  $\beta$ -lactam antimicrobial agents classified as critically important for human medicine by the World Health Organization due to their use in the management and treatment of serious infections caused by multidrug

resistant (MDR) Gram-negative and Gram-positive bacteria<sup>1,2</sup>. Third-generation cephalosporins were introduced for medical use in the early 1980s, with fourth-generation cephalosporins introduced in 1994<sup>3</sup>; since then, resistance to these drugs has increased<sup>4</sup>. Resistance

<sup>1</sup>Quadram Institute Bioscience, Norwich Research Park, Norwich NR4 7UQ, UK. <sup>2</sup>Department of Pathobiology, University of Guelph, Guelph N1G 2W1, Canada. <sup>3</sup>Microbes Intestin Inflammation et Susceptibilité de l'Hôte (M2ISH), Faculté de Médecine, Université Clermont Auvergne, Clermont-Ferrand 63001, France. <sup>4</sup>Centre National de Référence de la résistance aux antibiotiques, Centre Hospitalier Universitaire de Clermont-Ferrand, Clermont-Ferrand 63000, France. <sup>5</sup>Unité Antibiorésistance et Virulence Bactériennes, Anses Laboratoire de Lyon, Université de Lyon, Lyon 69007, France. <sup>6</sup>Institute of Microbiology and Epizootics, Department of Veterinary Medicine, Freie Universität Berlin, Berlin 14163, Germany. <sup>7</sup>Veterinary Centre for Resistance Research (TZR), Department of Veterinary Medicine, Freie Universität Berlin, Berlin 14163, Germany. <sup>8</sup>National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba R3E 3R2, Canada. <sup>9</sup>Department of Medical Microbiology and Infectious Diseases, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba R3E 0J9, Canada. <sup>10</sup>Department Method Standardisation, Resistance to Antibiotics Unit Monitoring of Resistance to Antibiotics, Federal Office of Consumer Protection and Food Safety, Berlin 12277, Germany. <sup>11</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK. <sup>12</sup>University of East Anglia, Norwich NR4 7TJ, UK. ✉e-mail: [alison.mather@quadram.ac.uk](mailto:alison.mather@quadram.ac.uk)

to these extended-spectrum cephalosporins (ESCs) is widespread worldwide in Enterobacterales from both humans and animals, and is mediated predominantly by two main groups of inactivating enzymes: extended-spectrum beta-lactamases (ESBLs) and plasmid-mediated AmpC  $\beta$ -lactamases (AmpCs)<sup>5,6</sup>.

There are differences in the distribution of ESC resistance (ESC-R) genes between host species and geographical origins<sup>6</sup>. ESBLs were initially detected in Europe; in 1983, the first ESBL—a SHV variant—was identified in Germany in *Klebsiella pneumoniae* and *Serratia marcescens* isolates<sup>7</sup>. Until the end of the 1990s, most of the ESBLs detected were variants of SHV and TEM<sup>8</sup>. Since then, the main types in Europe have changed, with CTX-M  $\beta$ -lactamases becoming the dominant ESC-R enzymes in Enterobacterales in humans and farm animals<sup>9</sup>. For example, the ESBL gene *bla*<sub>CTX-M-1</sub> is broadly disseminated among *Escherichia coli* from animals in Europe<sup>6</sup>. In human patients, the frequency of ESBL-producing *E. coli* and *K. pneumoniae* isolates was higher in Europe compared to North America between 2004 and 2006<sup>8,10</sup>. In Canada, a national study reported the AmpC gene *bla*<sub>CMY-2</sub> in *E. coli* isolates from patients of tertiary-care hospitals in 2000<sup>11</sup>. Subsequently, this variant was also reported in *E. coli* isolates from intensive care units (ICUs) between 2005 and 2006<sup>12</sup>. The first reports of *bla*<sub>CMY-2</sub> in Enterobacterales from food-producing animals in Canada were between 1995 and 1999<sup>13</sup>, and the gene has been commonly found in *E. coli* and *Salmonella enterica* from food-producing animals and food since then<sup>14–16</sup>. ESBLs in Canada were first observed in 2000 in *E. coli* and *Klebsiella* spp. from human clinical samples<sup>17</sup>. More recently, there have been reports of increasing occurrence of CTX-M-type  $\beta$ -lactamases in Enterobacterales in farm and companion animals from Canada<sup>18,19</sup>. In contrast, *bla*<sub>CMY-2</sub> was rarely reported in *E. coli* from humans in Europe<sup>20,21</sup>, but *bla*<sub>CMY-2</sub> isolates from livestock production have been reported<sup>6</sup>. A recent study from Germany found *bla*<sub>CMY-2</sub> in isolates from humans, livestock and meat products<sup>22</sup>, and it has also been reported in *E. coli* from companion animals in France<sup>23</sup>. International travel and trade have likely influenced the global dissemination of ESC-R genes<sup>24</sup>. Comparison among ESBL and AmpC-producing *E. coli* isolates of human and animal origins from Germany, the Netherlands and the United Kingdom showed that human isolates were more closely related to each other than to isolates from animals<sup>25</sup>. In another recent study in The Netherlands, the majority of community-acquired ESBL and AmpC carriage among *E. coli* was found to be due to human-to-human transmission, with much less attributed to food, animal and environmental sources<sup>26</sup>.

ESBL and AmpC genes are typically located on mobile genetic elements<sup>27</sup> and their spread has been linked to horizontal gene transfer via plasmids<sup>28–30</sup>. However, the expansion of bacterial clones carrying ESC-R genes also plays a role<sup>31</sup>. For example, the global spread of *bla*<sub>CTX-M-15</sub> in humans is associated with the *E. coli* epidemic lineage sequence type (ST) 131<sup>31,32</sup>. Although numerous studies have been conducted examining the types and putative sources of ESC-R genes<sup>9–16,19,21,27</sup>, these mostly investigated isolates from single host species or geographical regions. To gain a broader understanding of the emergence and spread of resistance to these important drugs, diverse sources must be examined. In addition, the relative importance of horizontal gene transfer and clonal expansion in the spread of ESC-R between ecological niches and biological compartments is unclear but may shed light on where to target interventions to reduce the burden of antimicrobial resistance (AMR). In this study, we used genomic and evolutionary approaches with an extensive collection of Enterobacterales isolates to determine the dynamics and evolution of ESBL and AmpC over time in different ecological compartments (country and source) and identify potential factors contributing to the spread and persistence of ESC resistance. Thus we (i) identified similarities and differences in ESC-R genes between multiple Enterobacterales species, host species and geographical regions; (ii) analysed clonal lineages and putative plasmids able to spread across

compartments; and (iii) examined the temporal trend of ESC-R genes in different compartments.

## Results

### Bacterial isolates and epidemiological data

A total of 1930 genomes were used in this study (Supplementary Table S1) for which extensive epidemiological metadata were available. These included *E. coli*, *K. pneumoniae*, *S. enterica*, *Enterobacter hormaechei*, *Klebsiella oxytoca*, *Proteus mirabilis*, *Klebsiella michiganensis*, *Citrobacter europaeus*, *Citrobacter freundii* and *Raoultella ornithinolytica* (Supplementary Table S2). Of the 1818 *E. coli*, 1617 were ESC-R isolates collected between 2008 and 2016, 79 ESC-R isolates before 2008 or after 2016 and 122 ESC-susceptible isolates mainly from 2008 to 2016. The set also included 112 non-*E. coli*, comprising 108 ESC-R genomes from 2008 to 2016, and four ESC-susceptible non-*E. coli* (Supplementary Figs. S1–S3).

The isolates derived from diverse sources, which were classified into eleven categories: human, cattle, chicken, food (porcine, bovine and avian meat), pig, turkey, dog, horse, wastewater (including wastewater from human activity as well as from animal and meat processing plants), cat and other avian sources (bird and duck) (Supplementary Table S3 and Source Data file). Overall, the majority of isolates were of human ( $n = 817$ , 42.3%), cattle ( $n = 372$ , 19.3%) and chicken ( $n = 263$ , 13.6%) origins (Supplementary Fig. S4). Assessment of the potential inclusion of outbreak-associated isolates, which may introduce bias into the analyses, identified only two ST10 and six ST177 genomes sharing  $\leq 10$  single nucleotide polymorphisms (SNPs) in their core genome. For ST131, there were only 36 genomes (mostly with *bla*<sub>CTX-M-15</sub>) sharing  $\leq 10$  SNPs, but these were from different years (from 2008 to 2016) and countries. Overall, this suggests outbreak isolates were not included in this study.

### MLST and plasmid replicon identification

Assembly summary statistics are presented in the Source Data file. There were 285 known STs and 13 unknown STs identified among the 1818 *E. coli* isolates, the most common being ST131 at 19.7% ( $n = 359$ ), ST10 at 6.9% ( $n = 125$ ) and ST117 at 5.6% ( $n = 102$ ) (Supplementary Fig. S5A). These three common STs were identified in all four countries (Canada, France, Germany and the UK) (Supplementary Fig. S5B) and multiple host species, with ST131 mainly associated with humans, ST10 mainly with cattle and ST117 mainly with poultry (Supplementary Fig. S5C–F). The distribution of isolates showed relatively even coverage of the main study time period between 2008 and 2016 for Canada, France and Germany. Relatively higher numbers of isolates were included in specific years for the UK (2015) and Germany (2012), but this was due to the sample selection in the previously published studies<sup>22,32,33</sup> (Supplementary Figs. S6A, B, S7A–D). Regarding identification of putative plasmids, 60 replicons were identified in the *E. coli* isolates, representing 19 plasmid incompatibility (Inc) types (Supplementary Table S4, Supplementary Figs. S8, S9).

### Genes conferring resistance to third generation cephalosporins

By study design, the majority of isolates selected for inclusion were phenotypically resistant to ESCs and so do not represent the prevalence of ESC-R in the general populations. Among the 1696 ESC-resistant *E. coli* isolates, 39 different ESC-R genes were identified. There were 55 combinations, or profiles, of ESC-R genes, some of which comprised more than one ESC-R gene. The most frequent types and profiles of ESC-R genes were: *bla*<sub>CTX-M-1</sub> ( $n = 414$ ) and *bla*<sub>CTX-M-15</sub> ( $n = 418$ ) with 24.4% and 24.6%, respectively, followed by *bla*<sub>CMY-2</sub> 19.2% ( $n = 326$ ), *bla*<sub>CTX-M-14</sub> 10.8% ( $n = 183$ ), *bla*<sub>SHV-12</sub> 8.3% ( $n = 140$ ), *bla*<sub>CTX-M-27</sub> 3.5% ( $n = 60$ ), *bla*<sub>CTX-M-55</sub> 1.2% ( $n = 20$ ), *bla*<sub>CTX-M-2</sub> 0.9% ( $n = 16$ ), and *bla*<sub>CMY-2</sub> + *bla*<sub>SHV-12</sub> 0.7% ( $n = 12$ ). The remaining 46 ESC-R profiles were rarely observed, each with 10 or fewer isolates. The non-*E. coli* isolates shared 11 ESC-R profiles with *E. coli*, while there were 44

ESC-R profiles unique to *E. coli* and 16 ESC-R profiles unique to the non-*E. coli* isolates (Supplementary Fig. S10A, B).

The major ESC-R genes in *E. coli*, *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-14</sub> and *bla*<sub>SHV-12</sub>, were globally distributed, though differences in frequency were found between countries, sources and compartments (Fig. 1a, b, Supplementary Fig. S11 and Supplementary Tables S5, S6). To note, this analysis did not include the isolates from Pietsch et al.<sup>22</sup> ( $n = 148$  ESC-R genomes) as the design of that study focused specifically on German *bla*<sub>CMY-2</sub> isolates; after excluding these 148 genomes, this resulted in 1548 ESC-R *E. coli* genomes which were used for further comparative analysis. In this collection of 1548 ESC-R *E. coli* genomes, the *bla*<sub>CTX-M-1</sub> gene was more frequent in European countries (France, Germany and the UK), while *bla*<sub>CTX-M-15</sub> and *bla*<sub>CMY-2</sub> were more common in North America (Canada) (Fisher's Exact test, adjusted  $p$ -value  $< 0.0001$ ; Fig. 1a). The *bla*<sub>CTX-M-1</sub> gene was more frequent in isolates of animal (cattle, chicken, pig, turkey, dog, horse and cat) than of human origin, whereas *bla*<sub>CTX-M-15</sub> was more frequent in human isolates (Fisher's Exact test, adjusted  $p$ -value  $< 0.0001$ ). The frequency of *bla*<sub>CMY-2</sub> was similar between isolates from humans and animals (Fisher's Exact test, adjusted  $p$ -value = 0.1) (Fig. 1b).

Fifteen ESC-R profiles were observed at least in two countries (Fig. 1d). Six profiles were found only in Canada, in human ( $n = 5$ ), chicken ( $n = 1$ ) and cattle ( $n = 1$ ) isolates. Three profiles were unique to UK isolates from human ( $n = 1$ ), food ( $n = 1$ ) and chicken ( $n = 10$ ) sources (Fig. 1e). Twenty-three ESC-R profiles were unique to France, mainly in human isolates (Fig. 1e). Highly diverse ESC-R profiles were observed in isolates of French human origin (Fig. 1c), which was not the case for the ST and plasmid replicon profiles (Supplementary Fig. S12A, B).

### Location of ESC-R genes in plasmid and chromosome contigs

Ascertainment of ESC-R gene locations in plasmids or chromosomes was obtained through MOB-suite<sup>34</sup> and RFPlasmid<sup>35</sup> (Supplementary Table S7). Initial use of MOB-suite alone identified an unexpectedly high proportion of ESC-R genes on chromosome-classified contigs, particularly for the *bla*<sub>CTX-M-15</sub> and *bla*<sub>SHV-12</sub> genes. Plasmid-classified contigs with ESC-R genes were further categorised into typeable plasmids (presence of replicon sequences in the same contig) and non-typeable plasmids (no identified replicon sequences in the same contig). In contrast, application of the RFPlasmid tool classified the majority of all ESC-R genes examined as located on plasmids, the majority of which were non-typeable plasmids. To resolve these disagreements and validate the location of ESC-R genes, we randomly selected 20 isolates—five each with the four most common ESC-R genes identified, *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>CMY-2</sub> and *bla*<sub>CTX-M-14</sub>—where prediction of the genomic location of the ESC-R gene differed between MOB-suite and RFPlasmid and performed long-read sequencing to obtain complete contiguated genome sequences. In all 20 long-read genome sequences, the RFPlasmid predictions were correct, suggesting MOB-Suite over-predicted the number of chromosome-located ESC-R genes. Using the RFPlasmid results (Supplementary Table S7), the majority of each of the five main ESC-R genes were located on plasmids, but to different degrees. The vast majority of *bla*<sub>CMY-2</sub> genes were predicted to be plasmid-borne (95.7%), in contrast to *bla*<sub>CTX-M-14</sub> (78.1%) and *bla*<sub>CTX-M-15</sub> (79.9%). Overall, 313 typeable plasmids linked with ESC-R genes (hereafter ESC-R plasmids) were recovered from the short-read data (*bla*<sub>CTX-M-1</sub>  $n = 136$ , *bla*<sub>CTX-M-15</sub>  $n = 23$ , *bla*<sub>CMY-2</sub>  $n = 88$ , *bla*<sub>CTX-M-14</sub>  $n = 40$  and *bla*<sub>SHV-12</sub>  $n = 26$ ), representing 27.6% of the predicted plasmids associated with ESC-R genes. Of these, IncII and IncF were the most frequently identified plasmid Inc types associated with ESC-R genes. *bla*<sub>CTX-M-1</sub> and *bla*<sub>CMY-2</sub> were predominantly associated with IncII plasmids, *bla*<sub>CTX-M-15</sub> and *bla*<sub>SHV-12</sub> with IncF plasmids, and *bla*<sub>CTX-M-14</sub> with both IncII and IncF plasmids (Supplementary Table S8). Analysis of the entire genetic content (plasmidome) of the 313 typeable and a subset of 79 non-typeable ESC-R-carrying plasmid

contigs showed clustering by ESC-R genes, plasmid replicon types and subtypes (Supplementary Fig. S13).

### Inferring the mechanisms of transmission and fitness of main ESC-R genes in *E. coli*

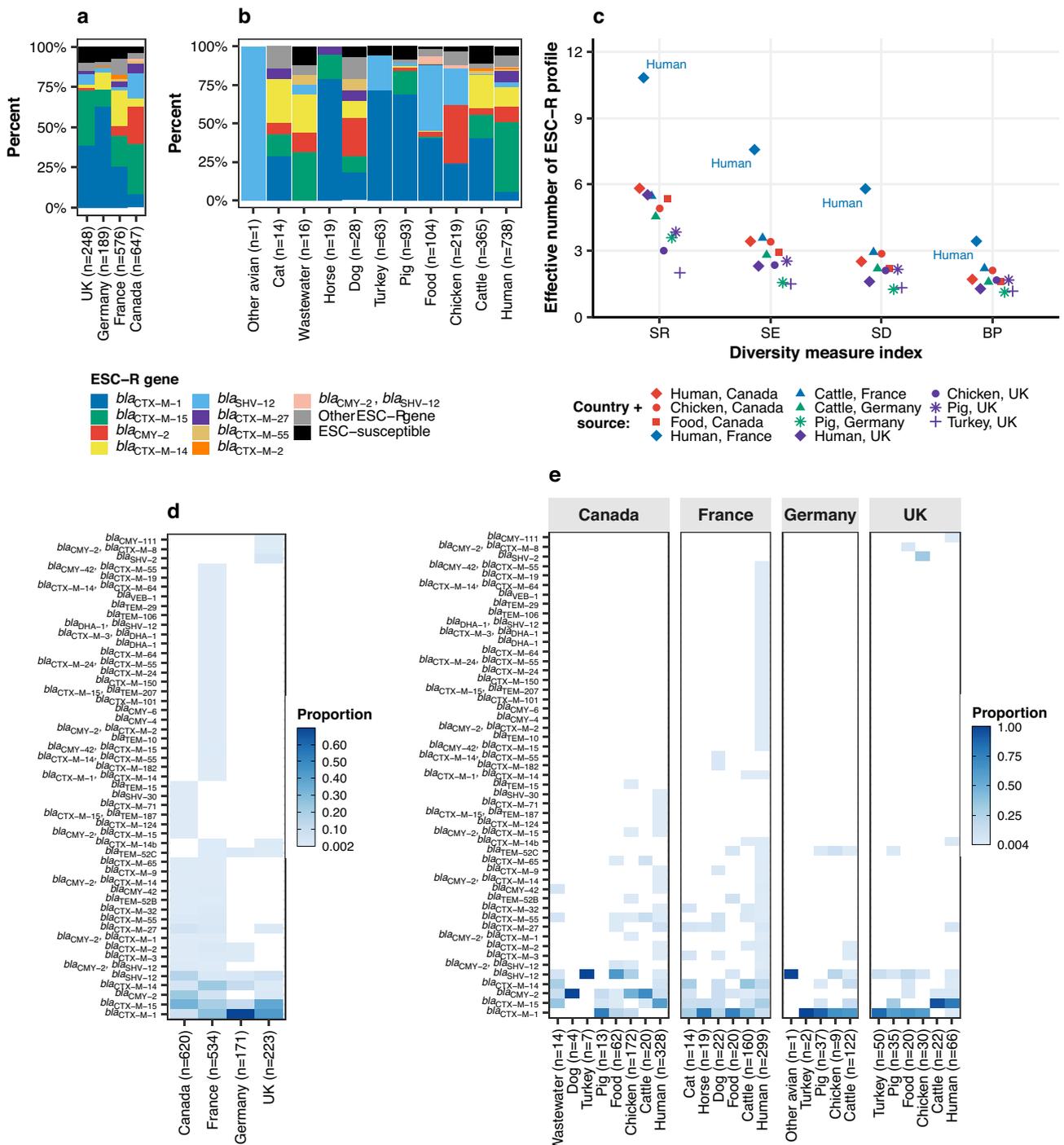
Simpson's diversity indices of STs and plasmid replicon profiles in the subset ( $n = 1343$ ) of the genomes carrying at least one of the five major ESC-R genes were calculated and compared; we refer to this as the 'main ESC-R gene' dataset hereafter. This analysis was also performed on the subset of 313 typeable ESC-R plasmid contigs as a validation step (Table 1). The ST diversity indices for most ESC-R genes were similar, ranging from 17.94 to 33.90 in the main ESC-R gene dataset and from 10.46 to 23.79 in the typeable plasmid subset. The exception was for genomes with *bla*<sub>CTX-M-15</sub>, where the ST diversity was 3.24 in the main ESC-R gene dataset and 3.66 in the typeable plasmid subset. The plasmid replicon profile diversity for genomes carrying most ESC-R genes were also similar in the main ESC-R gene dataset, with indices ranging from 72.88 to 81.88; the exception was genomes with *bla*<sub>CMY-2</sub>, where the index was 145.71. In the typeable plasmid subset, the range of plasmid replicon profile diversity was lower, between 1.20 and 5.83. To infer the relative importance of clonal expansion versus putative plasmids for dissemination of ESC-R genes in *E. coli*, the ratio of plasmid-to-ST diversities for each major ESC-R gene was calculated. For *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-14</sub> and *bla*<sub>SHV-12</sub> in the main ESC-R gene dataset, the ratio was slightly  $< 3$  (range 2.42–2.88); for *bla*<sub>CMY-2</sub> the ratio was 8.12, and for *bla*<sub>CTX-M-15</sub> the ratio was 22.49. In the typeable plasmid subset, the ratio was 0.05 and 0.10 for *bla*<sub>CTX-M-1</sub> and *bla*<sub>CTX-M-14</sub>, respectively, 0.17 for *bla*<sub>CMY-2</sub>, 0.56 for *bla*<sub>SHV-12</sub> and 1.59 for *bla*<sub>CTX-M-15</sub>. Plasmid replicon profile and ST Simpson's diversity indices were also calculated for each major ESC-R gene by individual compartment in the full dataset (Supplementary Fig. S14).

### Trends over time of ESC-R genes by compartment in *E. coli*

In the two countries from which both human and animal data were available, the dynamics of AmpCs and ESBLs changed over the study period (Fig. 2a, b). In farm animals (mainly chickens) in Canada, the main ESC-R gene type shifted from AmpC to ESBL (Wald z-statistic,  $p < 0.0001$ ). Conversely, in humans from France, the proportion of ESBLs declined over time while that of AmpCs increased (Wald z-statistic,  $p < 0.0001$ ); nevertheless, ESBLs remained the most common gene type (Supplementary Table S9). There were no significant changes in AmpC and ESBL frequency over time in the other examined compartments, humans from Canada (Wald z-statistic,  $p = 0.2$ ), nor in farm and companion animals from France (Wald z-statistic,  $p = 1.0$ ) (Supplementary Table S9). When evaluated by individual ESC-R genes, contrasting trends within individual compartments were observed. For example, over the study period *bla*<sub>CMY-2</sub> increased in human isolates from France, while this ESC-R gene decreased in chickens from Canada and was replaced by *bla*<sub>CTX-M-1</sub> (Fig. 2c, Supplementary Table S9).

### Resistance to other antimicrobial classes

In addition to the ESC-R genes, an additional 137 acquired AMR genes conferring resistance to other antimicrobial agents were identified, with up to 22 AMR genes and up to 12 plasmid replicons in a single genome (Supplementary Figs. S15, S16). There was variable presence of determinants (genes or SNPs) conferring resistance to different antimicrobial classes between compartments (Fig. 3a, Supplementary Tables S10, S11). Logistic regression analysis identified that determinants conferring resistance to aminoglycosides, sulfonamides, tetracycline, beta-lactams, diaminopyrimidines and quinolones-fluoroquinolones (due to mutations in *gyrA*, *parC* and *parE*; Supplementary Fig. S17 and Table S12) decreased over time in human isolates from France (Wald z-statistic,  $p$ -value  $< 0.00125$ ), while in chickens from Canada the presence of genes conferring resistance to sulfonamides and



**Fig. 1 | Frequency and diversity of ESC-R genes in *E. coli* by country and source.** **a** Percentage of ESC-R gene by country and **b** source (total  $n = 1,660$  ESC-R + ESC-susceptible genomes\*). The bars are coloured by ESC-R gene, as shown in the inset legend. **c** Diversity indices (SR: species richness; SE Shannon entropy, SD Simpson diversity, and reciprocal BP Berger-Parker) of ESC-R profile by compartment (country + source), where only compartments with minimum sample size of  $n = 30$  were included in the analysis (total  $n = 1361$  ESC-R genomes\*). The shape of the dots

is linked with the source and the colours with the country. **d** Common and unique ESC-R genes among countries and **e** compartments (total  $n = 1548$  ESC-R genomes\*). Heatmaps are coloured as per inset scale bars. \*Genomes from Pietsch et al.<sup>22</sup> ( $n = 158$ ; 148 ESC-R genomes + 10 ESC-susceptible genomes) were excluded. Source data are available in the Source Data file. ESC-R: extended-spectrum cephalosporin resistance.

tetracycline increased over time (Wald z-statistic,  $p$ -value < 0.00125) (Fig. 3b). Diversity indices showed highly diverse AMR profiles in isolates of French human origin and animals from France and Germany (Supplementary Fig. S18).

**Core-gene SNP-based phylogeny for *E. coli***

A maximum likelihood core-gene tree was constructed for 1818 *E. coli* genomes ( $n = 1696$  ESC-R and  $n = 122$  ESC-susceptible). The length of the core-gene alignment was 2300,993 bp (with 365,784 SNPs), which

**Table 1 | Characteristics and Simpson's diversity (SD) of sequence types (STs) and plasmid replicon profiles of *E. coli* (n = 1,343) and *E. coli* with typeable ESC-R gene-carrying plasmids (n = 313) harbouring the five major ESC-R genes**

Main ESC-R gene	Main ESC-R gene genome dataset			Typeable ESC-R plasmids subset				Inference of transmission <sup>b</sup>	
	# of isolates (% of total 1,343) <sup>a</sup>	SD of STs (95% CI) <sup>a</sup>	SD of plasmid replicon profiles (95% CI) <sup>a</sup>	Ratio of plasmid/ST diversity <sup>a</sup>	# of typeable plasmids <sup>b</sup>	SD of STs (95% CI) <sup>b</sup>	SD of plasmid replicon profiles (95% CI) <sup>b</sup>		Ratio of plasmid/ST diversity <sup>b</sup>
<i>bla</i> <sub>CTX-M-1</sub>	414 (30.8%)	29.35 (27.53–34.39)	79.04 (66.64–99.11)	2.69	136	21.86 (19.07–29.03)	1.20 (1.20–1.32)	0.05	Dominant plasmid [Inc11 pST3]
<i>bla</i> <sub>CTX-M-15</sub>	418 (31.1%)	3.24 (3.23–3.74)	72.88 (62.31–94.27)	22.49	23	3.66 (3.37–6.68)	5.81 (4.94–8.31)	1.59	Dominant ST [ST131]
<i>bla</i> <sub>CMY-2</sub>	188 (14.0%)	17.94 (16.53–23.90)	145.71 (93.26–198.16)	8.12	88	18.12 (15.3–26.14)	3.10 (3.05–3.74)	0.17	Diverse plasmids
<i>bla</i> <sub>CTX-M-14</sub>	183 (13.6%)	25.56 (22.64–30.89)	73.61 (52.91–106.33)	2.88	40	23.79 (15.39–35.63)	2.37 (2.32–2.81)	0.10	Dominant plasmid [IncF F2:A-B-]
<i>bla</i> <sub>SHV-12</sub>	140 (10.4%)	33.90 (27.61–43.16)	81.88 (52.13–112.92)	2.42	26	10.46 (7.86–22.18)	5.83 (5.05–8.91)	0.56	Diverse plasmids in a common ST [ST117]

<sup>a</sup>Numbers and diversity analysis obtained from 1343 *E. coli* genomes carrying the main five ESC-R genes from short-read data.

<sup>b</sup>Numbers and diversity analysis obtained from 313 typeable ESC-R plasmids.

represents 45% of the reference genome EC958 (Refseq: NZ\_HG941718.1). In the phylogenetic tree, there were eight major clusters largely concordant with phylogroup identification<sup>36</sup> (Fig. 4). Within phylogroup A, CC/ST10 was predominant, in phylogroup G the main ST was ST117 and in phylogroup B2 it was CC/ST131 (Fig. 4). These results were consistent with the pangenome network analysis, where clusters largely matched the population structure of the core-gene tree (Supplementary Fig. S19). There were some exceptions, particularly with phylogroup D genomes, that were phylogenetically clustered in the core-gene tree but fell into distinct clusters in the gene content network (Supplementary Fig. S19A).

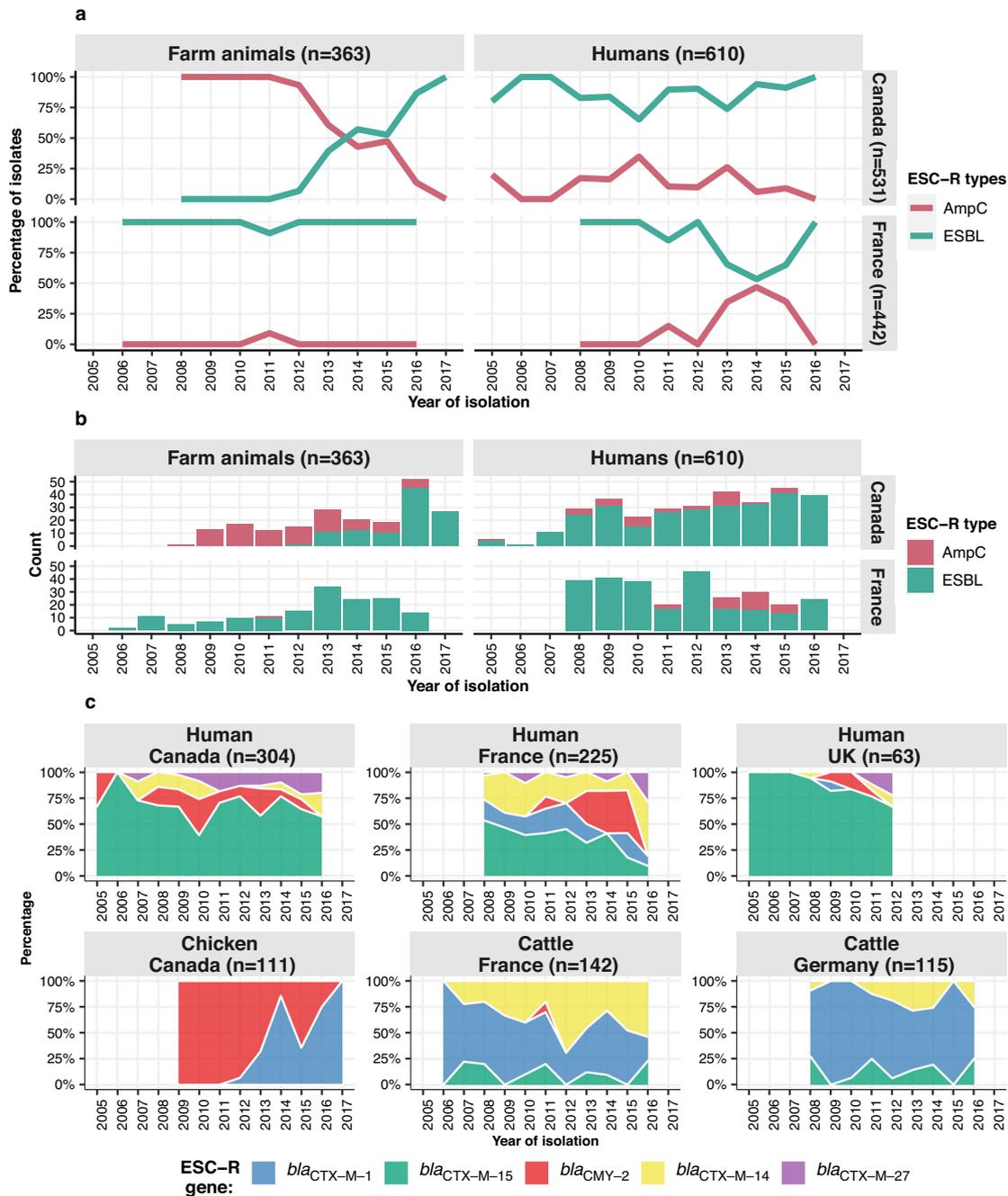
Genomes from different countries and sources were distributed across the different phylogroups, with ESC-susceptible genomes placed in diverse clusters. Previously identified host associations were observed, such as ST131 with humans and ST117 with chickens. Similarly, the ESC-R genes were widely distributed in the different clusters, with the exception of the previously identified association of *bla*<sub>CTX-M-15</sub> with ST131. Separate phylogenetic trees were constructed for genomes carrying *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-14</sub>, *bla*<sub>SHV-12</sub>, *bla*<sub>CTX-M-27</sub>, *bla*<sub>CTX-M-55</sub> or *bla*<sub>CTX-M-2</sub> (Supplementary Figs. S20–S22). Overall, the main clusters displayed in each tree correlated with phylogroups.

The 360 CC/ST131 genomes were classified in five clades based on marker alleles for the type 1 fimbriae *fimH* (H) adhesin: clade A (H41, n = 13) with half of the isolates carrying *bla*<sub>CTX-M-14</sub>, the remaining carrying other ESC-R genes; clade B (H22, n = 47; H30, n = 2; H54, n = 1) mostly carrying *bla*<sub>CMY-2</sub>; clade B0 (H27, n = 1) with a single *bla*<sub>SHV-12</sub> isolate; clade C1, also known as H30-R (H30, n = 70), with more than half of isolates carrying *bla*<sub>CTX-M-27</sub>; and clade C2, also known as H30-Rx (H30, n = 220; H35, n = 6), mostly harbouring *bla*<sub>CTX-M-15</sub>. Similar results were previously reported<sup>22,32,37–41</sup>. All genomes in Clades C1 and C2 also had double mutations in the *gyrA* and *parC* genes, which are associated with high-level resistance to fluoroquinolones<sup>42</sup>.

## Discussion

Resistance to ESC drugs is a global public health concern due to its widespread nature and the use of these drugs for difficult to treat infections. Detailed within-country investigations have been conducted, which have identified patterns of transmission between host species. For example, in The Netherlands ESBL/AmpC gene profile distribution from livestock and food-associated *E. coli* were different from those in *E. coli* from humans<sup>43</sup>, and a further study revealed humans as the main source of community-acquired ESBL and AmpC carriage in *E. coli* from humans based on distinguishable patterns of ESBL and AmpC gene occurrence in different sources<sup>26</sup>. Similarly, in the UK, whole genome sequencing (WGS) demonstrated distinct populations of *E. coli*, including those with ESBL and AmpC genes, in livestock and retail meat compared with those isolated from bloodstream infections in humans<sup>33</sup>. However, given the global nature of food supplies, animal movement and human travel, examining patterns from multiple countries and sources can provide additional insight and comparison of patterns and transmission of AMR. In particular, the use of WGS applied in this context provides valuable information for monitoring trends of specific ESBL and AmpC genes, as well as other AMR genotypes and potential linkage to particular plasmids, as was recently demonstrated in a collection of pig isolates from the UK<sup>44</sup> and in calf isolates from France<sup>45</sup>. Here, a genomic and evolutionary approach was used to determine the distribution and dynamics over time of ESBL and AmpC in *E. coli* from different geographical origins and sources and to infer the mechanisms underlying the transmission of the major ESC-R genes.

In comparing Canada and France, countries for which both human and animal data were available, two trends are clear. First, in farm animals in Canada (represented predominantly by chickens in this dataset), the majority of isolates carried AmpC genes at the beginning



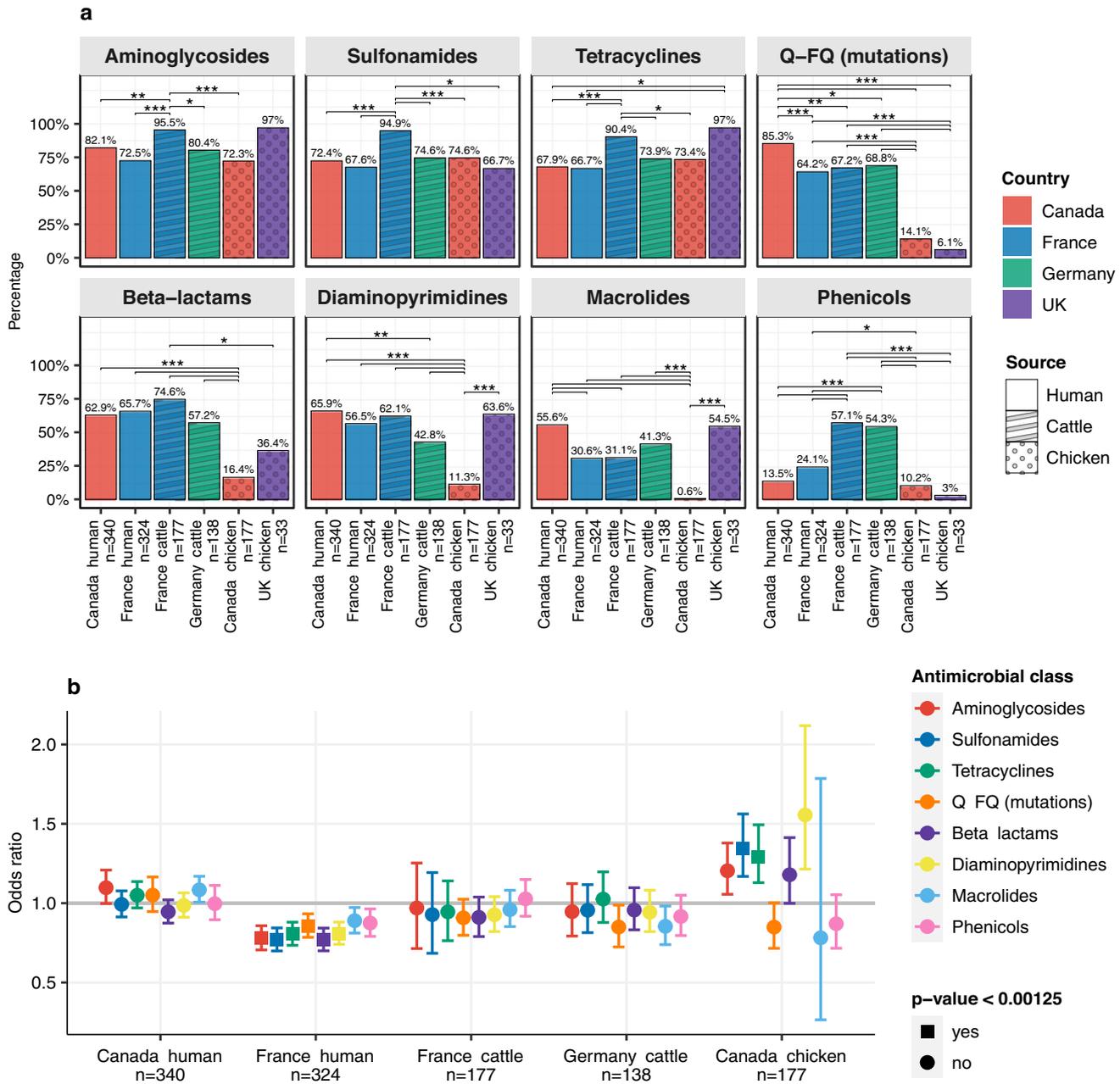
**Fig. 2 | Distribution and trend over time of ESC-R genes in *E. coli* by compartment.** **a** Trends in percentage and **b** count for ESC-R types over time in Canada and France across humans and farm animals (cattle, chickens, pigs and turkeys). The lines and bars are coloured by ESC-R types. **c** Trends over time of the

percentage of the major ESC-R gene in main compartments (colour key is in inset legend). Source data are available in the Source Data file. ESC-R extended-spectrum cephalosporin resistance, ESBL extended-spectrum  $\beta$ -lactamases.

of the study period, but by the end, the majority of ESC-R was due to ESBL genes. In contrast, the main determinant of ESC-R in French farm animals (represented predominantly by cattle in this dataset) consistently was ESBL genes. Second, in humans from Canada both AmpC and ESBL genes were found over the entire study period, with ESBL genes consistently found at a higher prevalence. In humans from France, ESC-R was initially dominated by ESBL genes, and then part way through the study period the AmpC gene *bla*<sub>CMY-2</sub> was identified and increased over time, although ESBLs remained as the main resistance mechanism. These observations agree with previous work describing the emergence and increase of ESBLs in Europe<sup>8,9,46</sup>, where *bla*<sub>CTX-M</sub> variants have become major determinants in humans and

animals in different enteric bacterial species<sup>9</sup>, with later emergence of *bla*<sub>CMY-2</sub><sup>22</sup>. The results presented are also consistent with previous work from North America, where the AmpC gene *bla*<sub>CMY-2</sub> was identified in isolates from humans, animals and foods in the early to mid 2000s<sup>12,14</sup>, while more recent studies have detected the presence of *bla*<sub>CTX-M</sub>-type variants in animals from Canada<sup>19</sup>.

The major ESC-R genes in *E. coli* identified in this study were *bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-14</sub> and *bla*<sub>SHV-12</sub>. These genes are also found in other enteric bacterial species, like those included in this study, and across the globe<sup>6</sup>. Despite their ubiquitous nature, there was variation in their distribution and frequency between geographical origins and sources. In these data, *bla*<sub>CTX-M-1</sub> was more frequent in



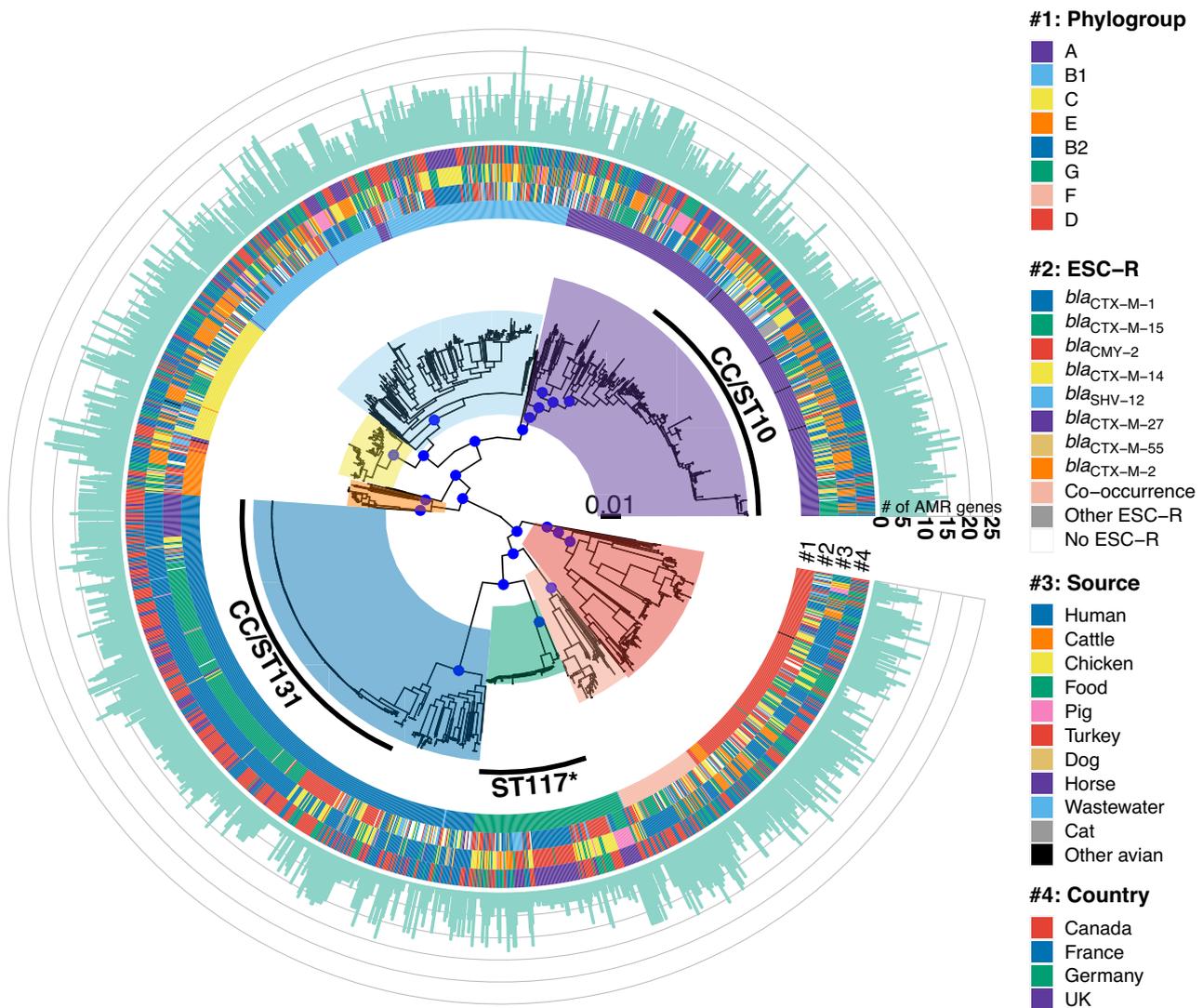
**Fig. 3 | Presence and dynamics of AMR determinants in *E. coli*.** **a** Presence of resistance determinants by antimicrobial class and compartment (country + source) for  $n = 1189$  *E. coli* isolates. Two-sided Fisher’s Exact Test was used to compare the proportion of AMR class between compartments, and the significance levels are indicated with asterisks: \* $p < 0.01$ , \*\* $p < 0.001$  and \*\*\* $p < 0.0001$ . The bars

are coloured by country and the pattern by source. **b** Odds ratios and 95% confidence intervals (coloured as per inset legend) demonstrating change in AMR determinant presence over the study period by compartment; significance cut-off is 0.00125 after Bonferroni adjustment for multiple hypothesis testing. Source data are available in the Source Data file. Q-FQ quinolone-fluoroquinolone.

Europe compared to North America, where  $bla_{CMY-2}$  and  $bla_{CTX-M-15}$  were more common. With respect to source (host species),  $bla_{CTX-M-1}$  was more frequent in animals,  $bla_{CTX-M-15}$  in humans, and  $bla_{CMY-2}$  was frequent in both humans and animals. The link between specific sources and these ESC-R genes has been previously reported<sup>6,22,32,33</sup>. ESC-susceptible genomes were distributed throughout different clusters and compartments (country + source), thereby indicating that, in general and with the exceptions mentioned, ESC-R genes were not associated with particular chromosomal backgrounds or ecological compartments.

The first report of  $bla_{CTX-M-15}$  identified its location on large plasmids isolated from *E. coli*, *K. pneumoniae* and *E. aerogenes* isolates from

urinary and pulmonary specimens of patients hospitalized in New Delhi, India, in 1999<sup>47</sup>. From August 2000 to April 2002 in Canada, a major outbreak of multidrug-resistant and ESBL-producing *E. coli* isolates was identified in long-term care facilities<sup>48</sup>. A following study identified a plasmid with the  $bla_{CTX-M-15}$  gene associated with the outbreak strain, and this plasmid was genetically very similar to the one isolated in India<sup>49</sup>. In the UK,  $bla_{CTX-M-15}$  isolates were initially isolated from clinical samples in 2001<sup>50</sup>; subsequently, the spread of  $bla_{CTX-M-15}$  was associated with the expansion of the human-associated clone ST131<sup>32</sup>. In this study, a higher proportion of  $bla_{CTX-M-15}$  was identified in Canada and the UK compared to France, related to the higher proportion of ST131 in these countries. In France, more diverse STs were



**Fig. 4 | Core-gene phylogenetic tree for 1818 *E. coli* isolates.** Maximum likelihood tree with eight main clusters linked with phylogroup. Bootstrap support values > 95 are represented by blue dots for main clusters. Metadata are represented next to the tree by four rings: phylogroup, ESC-R gene, source and country (coloured as per

inset legend). The number of AMR genes per genome is plotted on the outside as a bar plot (light green colour). \*ST117 cluster contains mainly ST117 but also other closely related STs. Source data are available in the Source Data file. ESC-R extended-spectrum cephalosporin resistance.

associated with *bla*<sub>CTX-M-15</sub>, suggesting plasmids—particularly IncF plasmids—may play a role in the dissemination of this gene in that setting. Although ST131 and, therefore, *bla*<sub>CTX-M-15</sub> was strongly associated with human-derived isolates, it was also observed in isolates of dog, pig and wastewater origin. Transmission between humans and pets has been observed previously<sup>6,51</sup>. In addition to the *bla*<sub>CTX-M-15</sub>/ST131/human association, *bla*<sub>CTX-M-27</sub> was also observed to occur mainly in ST131 clade C1 human isolates, again with multiple IncF replicons. In 2006, ST131 C1/H30R with *bla*<sub>CTX-M-27</sub> was detected in Japan<sup>52</sup>, followed by a reported rise of *bla*<sub>CTX-M-27</sub><sup>46</sup>. The expansion of the ST131 C1/H30R clade was found to be responsible for the increase of this ESBL gene<sup>40</sup>, and was subsequently observed in France<sup>53</sup> and Germany<sup>54</sup>. Moreover, our study also found *bla*<sub>CMY-2</sub> predominantly in ST131 clade B in humans and chickens from Canada, an association previously reported elsewhere<sup>22,41</sup>.

The AmpC gene *bla*<sub>CMY-2</sub> has been reported frequently in North America and to be increasing in Europe<sup>22</sup>. Here we confirm those results in our temporal analysis of genomes from France, from which both human- and animal-derived genomes are available. We also revealed that

this significant increase was observed only among isolates from humans and not from farm animals (here, predominantly cattle). We observed similar frequencies between humans from Canada and France, but temporal trend analysis showed the frequency increased over time in French human isolates. In contrast to the relatively constant presence of the gene in Canadian human-derived genomes, *bla*<sub>CMY-2</sub> decreased over time in Canadian chickens, and was replaced by *bla*<sub>CTX-M-1</sub>. IncA/C plasmids have previously been associated with *bla*<sub>CMY-2</sub> in isolates from cattle faeces and wastewater in Canada<sup>55</sup>; in this study, IncA/C, IncB/O/K/Z and IncI replicons were observed frequently in *bla*<sub>CMY-2</sub>-carrying *E. coli* from chickens in Canada. In Germany, the spread of *bla*<sub>CMY-2</sub> was attributed to transmission of IncII and IncK plasmids in isolates of human origin<sup>22</sup>. Here we observed that human French *bla*<sub>CMY-2</sub>-carrying isolates frequently harboured IncII replicons, which is consistent with a previous French study of isolates from companion animals<sup>23</sup>.

Recovery of fully contiguated plasmid sequences from short-read data is very challenging due to the frequent occurrence of repeat elements that impede assembly of complete plasmid or chromosome sequences<sup>56</sup>. Despite this limitation, diverse bioinformatic tools are

available for plasmid detection, classification and reconstruction<sup>34,57,58</sup>, with a recent study identifying MOB-suite as the only tool to reconstruct correctly the majority of *E. coli* plasmids in a test set<sup>56</sup>. Using MOB-suite, we were able to recover typeable ESC-R plasmids, but it over-represented the proportion of ESC-R genes carried on the chromosome. This was particularly true for *bla*<sub>CTX-M-15</sub> and *bla*<sub>SHV-12</sub> genes, which may be because these genes are associated with IncF plasmids. IncF plasmids are highly diverse with multiple copies of insertion sequences (ISs)<sup>59</sup>, which will limit successful plasmid assembly. The chromosome overrepresentation was resolved with RFPlasmid<sup>35</sup>, which identified a much larger proportion of ESC-R genes carried on plasmids and was later confirmed by long-read sequence data of a subset of isolates. Through this plasmid prediction approach, we were able to infer the proportions of ESC-R genes located on plasmids and the chromosome: *bla*<sub>CMY-2</sub>, *bla*<sub>SHV-12</sub> and *bla*<sub>CTX-M-1</sub> were frequently located on plasmids with 95.7%, 93.6% and 85.4%, respectively, while the proportion of plasmid-borne *bla*<sub>CTX-M-14</sub> and *bla*<sub>CTX-M-15</sub> was 78.1% and 79.9%, respectively. A total of 313 typeable ESC-R plasmids were recovered by MOB-suite and RFPlasmid, representing 27.6% of all predicted plasmid contigs associated with ESC-R genes. Although relatively low, this percentage is comparable to that obtained in a recent study<sup>56</sup>; the remaining plasmid contigs were not typeable as they did not contain replicon sequences.

As part of this study, the predominant mode of transmission of each of the five main ESC-R genes was assessed using the diversities of bacterial backgrounds (STs), plasmid replicons and their ratio in two datasets: (1) in the main ESC-R gene dataset of 1343 *E. coli* ESC-R genomes carrying at least one of the five main ESC-R genes, and (2) in the subset of 313 genomes containing typeable ESC-R plasmids (Table 1). From the analysis of the main ESC-R gene dataset, two main trends were apparent. First, in four of the five ESC-R genes, ST diversity was similar; the exception was *bla*<sub>CTX-M-15</sub>, which had lower ST diversity. This was to be expected, given the predominance of ST131 representing the chromosomal backgrounds of isolates with the *bla*<sub>CTX-M-15</sub> gene. In this study, 55% of genomes with *bla*<sub>CTX-M-15</sub> were ST131; for the other ESC-R genes, the most common ST (either ST10, ST117 or ST131) represented only 9–17% of genomes. Second, the diversity of plasmid profiles, representing unique combinations of replicons, was similar for four out of the five main ESC-R genes. In this case, the exception was *bla*<sub>CMY-2</sub>, with a Simpson's diversity index of 145.7 compared to the 72.9–81.9 observed with the other ESC-R genes. However, when the diversity of plasmid replicons in genomes carrying *bla*<sub>CMY-2</sub> was assessed in individual compartments (Supplementary Fig. S14), the diversities within each compartment appeared to be similar to those for other ESC-R genes. The reason for this apparent contradiction is that the Simpson's diversity index evaluates the number and abundance of unique combinations of plasmid replicons; it does not assess whether or not those plasmid replicons are the same. Here, each compartment had approximately the same diversity of plasmid replicon profiles with the *bla*<sub>CMY-2</sub> gene, but these profiles were different between compartments, explaining the higher diversity when all compartments were considered together, as in Table 1. Given the difficulty in reconstructing full plasmid sequences with short-read genome data and, therefore, definitively linking ESC-R genes with plasmid types, we repeated this analysis on a subset of 313 genomes containing typeable plasmids, where the plasmid contig contained both the ESC-R gene and a replicon sequence. Similar to the main ESC-R gene dataset of 1343 genomes, this approach identified *bla*<sub>CTX-M-15</sub> with the lowest ST diversity and highest ratio of plasmid replicon to ST diversity. In contrast to the main ESC-R gene dataset where *bla*<sub>CMY-2</sub> genomes had the highest plasmid replicon diversity, in the typeable plasmid subset *bla*<sub>CTX-M-15</sub> and *bla*<sub>SHV-12</sub> had the highest plasmid replicon diversity. This discrepancy is likely because of the differential recovery of typeable plasmids; due to the multiple ISs in IncF plasmids, IncF plasmids represented a lower proportion of plasmids in the typeable

plasmid subset (26.8%) than in the main ESC-R gene dataset (36.7%) (Supplementary Table S13). As IncF plasmids are highly diverse<sup>60</sup>, we used the replicon sequence typing (RST) scheme for IncF plasmids<sup>61</sup> in these analyses to subdivide them beyond IncF. Therefore, the lower proportion of IncF plasmids in the typeable plasmid subset would have an even greater effect on plasmid replicon diversity compared to the main ESC-R gene dataset. Conversely, IncII plasmids were over-represented in the typeable plasmid subset (62.9%) vs the main ESC-R gene dataset (15.3%). Furthermore, as specific ESC-R genes have associations with specific plasmid Inc types, such as *bla*<sub>CTX-M-15</sub> with IncF (Supplementary Table S8), the typeable plasmid subset is not representative of the main ESC-R gene dataset.

The comparative use of ST and plasmid replicon diversities for different ESC-R genes can shed light into their evolutionary history and epidemiology. Using *bla*<sub>CMY-2</sub> from the main ESC-R gene dataset as an example, the ST diversity was much higher in genomes derived from French humans, where this ESC-R emerged more recently than in Canada and, therefore, may not have yet undergone the same process of purifying selection. The ratios of ST to plasmid diversity for each ESC-R gene also identified differences in how these genes spread. A higher ratio could be due to two factors: (1) the ST diversity is relatively low, as was the case for *bla*<sub>CTX-M-15</sub> in both the main ESC-R gene dataset and typeable plasmid subset, reflecting clonal expansion of lineages with *bla*<sub>CTX-M-15</sub> carried on either plasmids or the chromosome as the predominant mechanism of spread, or (2) the plasmid replicon diversity is relatively high, as was the case for *bla*<sub>CMY-2</sub> in the main ESC-R gene dataset, suggesting the gene is highly mobile between plasmids. While acknowledging the typeable plasmid subset is not representative of the main ESC-R gene dataset, examining those plasmid contigs with both ESC-R genes and plasmid replicons provides further insight into the mechanisms of ESC-R spread. This revealed dominant plasmid types associated with *bla*<sub>CTX-M-1</sub> and *bla*<sub>CTX-M-14</sub>–IncII plasmid sequence type (pST) 3 plasmids with *bla*<sub>CTX-M-1</sub> and IncF (subtype F2:A-B-) plasmids with *bla*<sub>CTX-M-14</sub>—suggesting the spread of these genes is mainly due to dominant plasmids across different genetic backgrounds. The higher ST:plasmid diversity ratio of *bla*<sub>SHV-12</sub> in the typeable plasmid subset is a result of relatively low ST and high plasmid diversity, which suggest that both ST clonal expansion (here, mainly ST117) and plasmids contribute to its spread. In the 26 genomes with *bla*<sub>SHV-12</sub> typeable plasmids, ST117 comprised 30.8% of the STs harbouring *bla*<sub>SHV-12</sub> IncF plasmids, while other STs carried IncF and IncII plasmids. When comparing the plasmid subtypes between *E. coli* and non-*E. coli*, we found three common plasmid subtypes (IncN pST1 *bla*<sub>CTX-M-1</sub>, IncII pST12 *bla*<sub>CMY-2</sub> and IncII pST26 *bla*<sub>SHV-12</sub>), which suggest that some plasmid subtypes are crossing species borders and mobilising the ESC-R genes between bacterial species (Supplementary Table S14). This result supports previous observations where plasmid replicons<sup>62</sup> and ESC-R genes were detected in diverse Enterobacterales species<sup>14,18</sup>.

It is important to consider the limitations of these data when interpreting the results. The bacterial collections from which isolates were selected for sequencing were generated for a variety of purposes, including surveillance and research projects, and as such utilised different inclusion criteria and methods as outlined in the Supplementary Material. The human isolates from France and Canada and the animal isolates from Germany and France were from cases of clinical disease, whereas animal isolates from Canada were from clinical cases and asymptomatic carriage. To reduce potential bias, isolates collected over the same time period were randomly selected from the different collections proportionately by year and from a diverse assortment of host species. Importantly, the identification of ESC-R was based on phenotypic testing, thus allowing comparison of the identity and diversity of ESC-R genes between compartments and was thus the focus of this study. To ensure the most robust inference from this dataset, we also restricted statistical comparisons to those ecological

compartments with good representation, such as farm animals and humans from Canada and France when analysing the temporal trends in ESC-R genes. For the same reason, in the diversity analyses we took multiple random subsets of ecological compartments with larger sample sizes to comparable sizes of ecological compartments with smaller sample sizes. One striking result was the comparatively high diversity of ESC-R genes in genomes derived from French humans; these genomes were collected over the entire study period and from various parts of the country. Although the sampling strategy cannot be ruled out as a causative factor in the results, the underlying causes for this greater diversity of ESC-R in *E. coli* from humans in France warrants further investigation. Through the plasmidome analysis, we showed that the included 79 non-typeable plasmids were closely related to the typeable plasmids, demonstrating their likely Inc type but also the limitations of assembly with short-read data. Thus, large scale long-read sequence data are necessary to identify reliably the location of ESC-R genes to the chromosome or to specific plasmids, as we demonstrated with the 20 genomes for which we used long-read sequence data to validate the ESC-R gene location. With such data, a greater resolution of the plasmidome network and transmission can be obtained.

The differing trends and types of ESC-R in the countries examined could be due in part to different amounts of antimicrobial use (AMU). In Europe, the use of antimicrobials agents as growth promoters in animals has been stopped since 2006, and overall sales of veterinary antimicrobial agents have decreased, including in France<sup>63</sup>, Germany<sup>64</sup> and the UK<sup>64</sup>. In comparison, Canada stopped the use of antimicrobials for growth promotion in 2016 and had higher sales for use in animals, with the predominant antimicrobial agents sold being tetracyclines, beta-lactams, macrolides and diaminopyrimidines-sulfonamides<sup>65</sup>. This contrast is perhaps reflected in the results of this study identifying an increase in resistance to sulfonamides and tetracyclines over the study period in chicken isolates from Canada. A recent study on AMU and AMR in broiler chickens in Canada during 2013-2019 showed that resistance to some antimicrobial classes decreased due to an overall reduction in AMU, but other antimicrobial classes, such as aminoglycosides, increased; this could be associated with the increased use of aminoglycosides in water to treat disease<sup>66</sup>. There can be differences in AMU between humans and animals by antimicrobial class, and between different species of animals<sup>63,67</sup>, which may also influence the prevalence of resistance to specific antimicrobials. However, there are limited data available on antimicrobial consumption and there are other factors, such as genetic linkage, which make understanding the causal relationships of use and resistance complex. Combining the high-resolution data conferred by WGS to understand genetic linkage of AMR genes along with greater available of AMU data will facilitate greater understanding of the relative contributions of different drivers of AMR.

In conclusion, this study confirmed the changes over time of the relative importance of ESBL and AmpC genes causing resistance to ESC drugs in Europe and North America. There was variation in the major ESC-R gene frequency between compartments (country + source), and in the main mechanism, clonality and/or plasmids, responsible for the dissemination of these genes. In addition, the co-occurrence of ESC-R genes with other AMR determinants was higher in some compartments in different countries. The changing patterns of antimicrobial use may be associated with the changing trends of resistance, but more data are required; given the demonstrated differences between compartments, the collection/addition of AMU data to WGS in future studies will enable greater understanding of this link. Utilising WGS to obtain high-resolution data not only to identify the type and diversity of ESC-R genes present, but also linking this information to bacterial host and plasmid diversity, offers deeper insight into the epidemiology of resistance to these critical drugs. Further in-depth study of these

plasmids and the role of other mobile genetic elements, such as transposons, in the transmission and global epidemiology of ESC-R is required, and will be facilitated by the increasing feasibility of large scale long-read genome sequencing to generate fully contiguated genome sequences. In line with the conclusions of other studies<sup>26</sup>, these results highlight the need for long-term internationally harmonised monitoring in order to understand the dynamics and relative importance of different sources to the burden of AMR over time.

## Methods

### Bacterial isolates

All ESC-R isolates included in this study were selected independent of the genetic mechanism causing the resistance. The majority (84%) of isolates in this study were selected to include representatives from different ecological compartments (human, animal, food/meat) of ESC-R *E. coli* isolated from 2008 to 2016. These isolates were randomly selected from available culture collections across three countries: France, Germany and Canada. The remaining 16% of isolates were selected to provide context, namely: (1) ESC-R *E. coli* isolates obtained prior to 2008 or after 2016 (4%); (2) ESC-R isolates of bacterial species other than *E. coli* from 2008 to 2016 (5.6%); and (3) ESC-susceptible *E. coli* (6.5%). With this strategy, 1524 isolates were selected: 718 isolates from Canada, 607 from France and 199 from Germany (Table S1). Bacterial collection and isolation methods for each subset are described in the Supplemental methods, section 1.

In addition, genomes from previous studies were included. This included 158 from Germany<sup>22</sup> and 248 from the UK<sup>32,33</sup>; these genomes were selected following a similar approach of sample selection as described above, where all genomes available between 2008 and 2016 were included and a small number of genomes collected prior to 2008 and/or genomes of ESC-susceptible isolates were selected.

### Whole genome data, MLST and typing of *fimH* allele and phylogenies

Paired-end short-read whole genome sequence data were generated on the Illumina platform for the 1524 isolates selected above, and fastq data available from the previous studies for the German (study accession PRJEB23663)<sup>22</sup> and UK genomes (study accessions PRJEB4681, PRJEB8774 and PRJEB8776)<sup>32,33</sup> were downloaded from the European Nucleotide Archive (<http://www.ebi.ac.uk/ena>). The methods for DNA extraction and WGS are detailed for each subset in the Supplemental method, section 1. Raw reads were processed with Trimmomatic v0.33<sup>68</sup> to remove adapters and poor quality bases. The trimmed reads were used to assemble draft genomes using Spades v3.11.1<sup>69</sup>; summary statistics of the assemblies were obtained using Quast v4.5<sup>70</sup>, and assemblies with a high number of contigs (>900 contigs) were excluded. Genome quality was also assessed using CheckM v1.1.2<sup>71</sup> where genomes with completeness below 95.0% and/or contamination in the assembly with more than 15.0% were excluded. Centrifuge 1.0.3-beta<sup>72</sup> and Kleborate v2.0.4<sup>73</sup> were used for the species assignment. In cases where species assignment differed between Centrifuge and Kleborate, FastANI v1.32<sup>74</sup> was applied and compared with NCBI genome references. BWA v0.7.17<sup>75</sup>, Samtools v1.9<sup>76</sup> and Bcftools v1.8<sup>77</sup> were used to map reads against their corresponding assemblies for the largest four contigs to estimate sequencing mean chromosomal depth coverage.

The Achtman sequence typing scheme<sup>78</sup> was applied using ARIBA v2.13.5<sup>79</sup> to assign the multilocus sequence type (MLST) for each genome. The sequence type (ST) was allocated to clonal complexes (CC) using the *E. coli* database available from the Enterobase website<sup>80</sup>. The *fimH* alleles for *E. coli* were downloaded from the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org>) and used with ARIBA for typing the alleles. Phylotyping for *E. coli* was performed using the ClermonTyping v1.4.1<sup>36</sup> tool.

## Assessment of potential inclusion of outbreak isolates

To determine whether or not highly similar outbreak isolates were included in the dataset, which could affect the proportion of specific ESC-R genes and temporal analysis, pairwise SNP distances were analysed within the core genome for each main ST: ST131, ST10 and ST117. To obtain the core genome, reads were mapped to the relevant reference genome using Snippy v4.3.6<sup>81</sup>. The reference genomes were: EC958 (NZ\_HG941718.1; [https://www.ncbi.nlm.nih.gov/nucore/NZ\\_HG941718.1](https://www.ncbi.nlm.nih.gov/nucore/NZ_HG941718.1)) for ST131, K-12 MG1655 (NC\_000913.3; [https://www.ncbi.nlm.nih.gov/nucore/NC\\_000913.3](https://www.ncbi.nlm.nih.gov/nucore/NC_000913.3)) for ST10 and MDR\_56 (NZ\_CP019903.1; [https://www.ncbi.nlm.nih.gov/nucore/NZ\\_CP019903.1](https://www.ncbi.nlm.nih.gov/nucore/NZ_CP019903.1)) for ST117. Pairwise SNP differences were obtained using snp-dists v0.7.0 (<https://github.com/tseemann/snp-dists>). High similarity between pairs of isolates was considered to be differences  $\leq 10$  SNPs.

## In silico identification of AMR determinants and plasmid replicons

ARIBA v2.13.5<sup>79</sup> was used to identify acquired AMR genes and plasmid replicons with the ResFinder<sup>82</sup> and PlasmidFinder<sup>61</sup> databases, respectively. ESBL and AmpC genes were identified from these databases. Mutations in gyrase genes (*gyrA* and *gyrB*) and topoisomerase IV genes (*parC* and *parE*) associated with quinolone resistance were also identified using ARIBA v2.13.5<sup>79</sup> with the wild-type sequences as the input databases. The frequency and distribution of resistance determinants to ESC and other antimicrobial classes were analysed by country, by source (i.e., host species) and by compartment, which is defined as the unique combination of country and source. A similar approach was used for plasmid replicons and mutations in topoisomerase genes. The numbers of acquired AMR genes and plasmid replicons for each sample were compiled and compared by analysing the distribution of their frequencies among groups (country + ESC-R genes) for each source.

## Prediction of ESC-R gene location to plasmid or chromosome contig in de novo assembly

The prediction of contigs as plasmid or chromosomal origin was first achieved using MOB-suite v2.0<sup>34</sup> with the mob\_recon option with default setting. In addition, RFPlasmid<sup>35</sup> was also used; a cut-off of  $\geq 0.6$  was used to identify the plasmid contigs. Resfinder<sup>82</sup> and Plasmidfinder<sup>61</sup> in Abricate v0.9.8 (<https://github.com/tseemann/abricate>) were used to identify ESC-R genes and plasmid replicon sequences in contigs. The plasmid subtyping was done using Blast+ v2.9.0<sup>83</sup> and the plasmid MLST (pMLST) scheme ([https://bitbucket.org/genomicpidemiology/pmlst\\_db/](https://bitbucket.org/genomicpidemiology/pmlst_db/)). Where the prediction of chromosome vs plasmid origin differed between MOB-suite and RFPlasmid, a subset of five genomes for the four most common ESC-R genes, for a total of 20 isolates, was randomly selected for long-read sequencing (see Supplemental Method, section 2, for further information on long-read data), to confirm the localization of the ESC-R gene either to a plasmid or the chromosome. The plasmidome analysis using PANINI<sup>84</sup> was performed in the set of 313 typeable plasmids plus a subset of 79 non-typeable plasmids (identified with MOB-suite) carrying ESC-R genes recovered from short-read data (See Supplemental Method, Section 3).

## Diversity analysis of ST, ESC-R, AMR and plasmid replicon profiles

The diversities of STs, ESC-R genes, AMR profiles and plasmid replicon profiles were estimated as previously described for *Salmonella* Typhimurium<sup>85</sup> using the vegan v2.5-6<sup>86</sup> R package. Four diversity measures were calculated, covering a range of weightings of richness and abundance: SR (species richness), SE (Shannon entropy), SD (Simpson's diversity) and reciprocal BP (Berger-Parker)<sup>85,87</sup>. The ESC-R profile was defined as the combination of unique ESC-R genes within a single genome. The AMR profile comprised the presence of all AMR

determinants, including ESC-R genes and mutations in the topoisomerase genes. The diversity analysis was performed by compartment (country + source). As sample size varied, compartments with larger sample sizes were randomly subsampled to the size of the smaller compartment. The subsampling consisted of 10,000 permutations without replacement and the mean with 95% confidence intervals of diversity indices were calculated. The calculated values for SE, SD and BP were converted into the effective number of profiles prior to plotting the results<sup>85</sup>.

In addition, Simpson's diversity (SD) for ST and plasmid replicon profiles were further investigated in two datasets: (1) in the main ESC-R gene dataset of 1343 *E. coli* genomes with the five main ESC-R genes, where we investigated diversity (i) within genomes carrying each major ESC-R gene, and (ii) between main compartments within each major ESC-R gene; (2) in the subset of 313 typeable ESC-R plasmids, where SD assessment of plasmid replicon profiles and their respective bacterial ST across ESC-R genes were performed; in this analysis, contigs predicted to be plasmids (identified through both MOB-suite and RFPlasmid) that were carrying both a main ESC-R gene and a replicon sequence were used. The diversity estimates were computed through the sample-size-based rarefaction and extrapolation method implemented in the iNEXT v2.0.20<sup>88</sup> R<sup>89</sup> package.

## Phylogenetic analysis

The draft assemblies were annotated using Prokka v1.13<sup>90</sup> and the pan-genome was analysed with Roary v3.12.0<sup>91</sup>; core genes were identified with 90% sequence similarity and presence in 95% of isolates. SNPs were extracted from the core gene alignment using snp-sites v2.5.1<sup>92</sup> and used to construct a maximum likelihood tree using the HKY and discrete GAMMA nucleotide model including constant sites information in IQ-tree v1.6.11<sup>93</sup>. A phylogenetic tree was constructed for the entire *E. coli* ( $n = 1839$ ) genome collection; due to the large sample size, the ultrafast bootstrap approximation<sup>94</sup> -B 1000 was used. Separate phylogenies were also constructed for genomes containing each main ESC-R gene (*bla*<sub>CTX-M-1</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>CMY-2</sub>, *bla*<sub>CTX-M-14</sub> and *bla*<sub>SHV-12</sub>) using subsets of the same core gene alignment. For the separate phylogenies, the standard nonparametric bootstrap<sup>95</sup> -b 100 was applied. Metadata were plotted against the trees using the ggtree v2.0.2<sup>96</sup> R<sup>89</sup> package. The pangenome network analysis was also performed using PANINI<sup>84</sup>; details are provided in the Supplemental Method, Section 4.

## Statistical analysis

Fisher's Exact Test from the rstatix's v0.5.0<sup>97</sup> R<sup>89</sup> package was used to compare the proportions of the major ESC-R genes between host species (human vs animal) and continents (Europe vs North America), and similarly, between main compartments within each major ESC-R gene and for overall AMR class. A Bonferroni correction was applied to account for multiple hypothesis testing.

Kruskal-Wallis tests, also available in rstatix's v0.5.0<sup>97</sup>, were used to compare the numbers of acquired AMR genes between groups constituted by country and ESC-R gene; groups with similar (statistically not significant) distributions were combined. The medians of the number of acquired AMR genes between the groups were compared through Mann-Whitney *U* tests.

To test for temporal trends in antimicrobial resistance, a logistic regression was applied, where the binary outcome variable was the presence or absence of a specific AMR determinant and the predictor variable was the year of isolation. To note, this analysis did not include a number of genomes from humans in France and Canada: 20 genomes that had both AmpC and ESBL genes and four AmpC-carrying genomes, as they were additional samples that were not originally selected with the same study selection strategy. Logistic regression analysis was done individually by compartment (country + source) for each ESC-R gene type (all AmpC genes, including *bla*<sub>CMY</sub> variants and *bla*<sub>DHA-1</sub>; and all ESBL genes, including variants of *bla*<sub>CTX-M</sub>, *bla*<sub>TEM</sub>, *bla*<sub>SHV</sub> and

*bla*<sub>VEB-1</sub>), main ESC-R genes and other antimicrobial classes (aminoglycosides, sulfonamides, tetracyclines, other beta-lactams [non-ESBL and non-AmpC genes], diaminopyrimidines, macrolides, phenicols and quinolones/fluoroquinolones due to mutations in *gyrA*, *parC* and *parE*). The glm(y - x, family = “binomial”) function available in R<sup>89</sup> was used to obtain the estimates, p-values based on z-scores (Wald statistic), odds ratio (OR) along with the 95% confidence intervals. A Bonferroni correction was applied to account for multiple hypothesis testing, as described above.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Illumina sequence read data generated in this study have been deposited in the ENA under BioProject [PRJEB38235](#), [PRJEB42322](#), [PRJNA523640](#), [PRJNA556083](#), [PRJNA740259](#) and [PRJEB50837](#) and individual accession numbers are available in Source Data file. Long-read data for the subset of 20 genomes are deposited in the ENA under BioProject [PRJEB54884](#); the accession numbers for each genome are placed in the Source Data file. Thus, the source data (Source Data file) provided with this paper lists accession numbers for each genome analysed in this study, alongside with the isolate collection metadata generated in this study. Source data are provided with this paper.

### References

- Arumugham, V. B. & Cascella, M. *Third Generation Cephalosporins. StatPearls* (2021).
- Leone, S., Cascella, M., Pezone, I. & Fiore, M. New antibiotics for the treatment of serious infections in intensive care unit patients. *Curr. Med. Res. Opin.* **35**, 1331–1334 (2019).
- Endimiani, A., Perez, F. & Bonomo, R. A. Cefepime: a reappraisal in an era of increasing antimicrobial resistance. *Expert Rev. Anti. Infect. Ther.* **6**, 805–824 (2008).
- Paterson, D. L. & Bonomo, R. A. Extended-spectrum beta-lactamases: a clinical update. *Clin. Microbiol. Rev.* **18**, 657–686 (2005).
- Harris, P. Clinical management of infections caused by *Enterobacteriaceae* that express extended-spectrum  $\beta$ -lactamase and AmpC enzymes. *Semin. Respir. Crit. Care Med.* **36**, 056–073 (2015).
- Ewers, C., Bethe, A., Semmler, T., Guenther, S. & Wieler, L. H. Extended-spectrum  $\beta$ -lactamase-producing and AmpC-producing *Escherichia coli* from livestock and companion animals, and their putative impact on public health: a global perspective. *Clin. Microbiol. Infect.* **18**, 646–655 (2012).
- Knothe, H., Shah, P., Krcmery, V., Antal, M. & Mitsuhashi, S. Transferable resistance to cefotaxime, ceftiofur, cefamandole and cefuroxime in clinical isolates of *Klebsiella pneumoniae* and *Serratia marcescens*. *Infection* **11**, 315–317 (1983).
- Coque, T. M., Baquero, F. & Canton, R. Increasing prevalence of ESBL-producing *Enterobacteriaceae* in Europe. *Euro Surveill.* **13**, 19044 (2008).
- Cantón, R. et al. Prevalence and spread of extended-spectrum beta-lactamase-producing *Enterobacteriaceae* in Europe. *Clin. Microbiol. Infect.* **14**, 144–153 (2008).
- Reinert, R. R. et al. Antimicrobial susceptibility among organisms from the Asia/Pacific Rim, Europe and Latin and North America collected as part of TEST and the in vitro activity of tigecycline. *J. Antimicrob. Chemother.* **60**, 1018–1029 (2007).
- Mulvey, M. R. et al. Molecular characterization of ceftiofur-resistant *Escherichia coli* from Canadian hospitals. *Antimicrob. Agents Chemother.* **49**, 358–365 (2005).
- Baudry, P. J., Mataseje, L., Zhanel, G. G., Hoban, D. J. & Mulvey, M. R. Characterization of plasmids encoding CMY-2 AmpC  $\beta$ -lactamases from *Escherichia coli* in Canadian intensive care units. *Diagn. Microbiol. Infect. Dis.* **65**, 379–383 (2009).
- Allen, K. J. & Poppe, C. Occurrence and characterization of resistance to extended-spectrum cephalosporins mediated by beta-lactamase CMY-2 in *Salmonella* isolated from food-producing animals in Canada. *Can. J. Vet. Res.* **66**, 137–144 (2002).
- Martin, L. C., Weir, E. K., Poppe, C., Reid-Smith, R. J. & Boerlin, P. Characterization of *bla*<sub>CMY-2</sub> plasmids in *Salmonella* and *Escherichia coli* isolates from food animals in Canada. *Appl. Environ. Microbiol.* **78**, 1285–1287 (2012).
- Chalmers, G. et al. Determinants of virulence and of resistance to ceftiofur, gentamicin, and spectinomycin in clinical *Escherichia coli* from broiler chickens in Québec, Canada. *Vet. Microbiol.* **203**, 149–157 (2017).
- Vounba, P., Arsenault, J., Bada-Alambédji, R. & Fairbrother, J. M. Antimicrobial resistance and potential pathogenicity of *Escherichia coli* isolates from healthy broilers in Québec, Canada. *Microb. Drug Resist.* **25**, 1111–1121 (2019).
- Mulvey, M. R. et al. Ambler class A extended-spectrum beta-lactamase-producing *Escherichia coli* and *Klebsiella* spp. in Canadian hospitals. *Antimicrob. Agents Chemother.* **48**, 1204–1214 (2004).
- Zhang, P. L. C. et al. Prevalence and mechanisms of extended-spectrum cephalosporin resistance in clinical and fecal *Enterobacteriaceae* isolates from dogs in Ontario, Canada. *Vet. Microbiol.* **213**, 82–88 (2018).
- Cormier, A. et al. Diversity of CTX-M-positive *Escherichia coli* recovered from animals in Canada. *Vet. Microbiol.* **231**, 71–75 (2019).
- Jørgensen, R. L., Nielsen, J. B., Friis-Møller, A., Fjeldsøe-Nielsen, H. & Schønning, K. Prevalence and molecular characterization of clinical isolates of *Escherichia coli* expressing an AmpC phenotype. *J. Antimicrob. Chemother.* **65**, 460–464 (2010).
- Valenza, G. et al. Extended-spectrum- $\beta$ -lactamase-producing *Escherichia coli* as intestinal colonizers in the German community. *Antimicrob. Agents Chemother.* **58**, 1228–1230 (2014).
- 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).
- Haenni, M., Saras, E., Métayer, V., Médaille, C. & Madec, J.-Y. High prevalence of *bla*<sub>CTX-M-1</sub>/Inc11/ST3 and *bla*<sub>CMY-2</sub>/Inc11/ST2 plasmids in healthy urban dogs in France. *Antimicrob. Agents Chemother.* **58**, 5358–5362 (2014).
- Hawkey, P. M. Multidrug-resistant Gram-negative bacteria: a product of globalization. *J. Hosp. Infect.* **89**, 241–247 (2015).
- Wu, G. et al. Comparative analysis of ESBL-positive *Escherichia coli* isolates from animals and humans from the UK, the Netherlands and Germany. *PLoS ONE* **8**, e75392 (2013).
- Mughini-Gras, L. et al. Attributable sources of community-acquired carriage of *Escherichia coli* containing  $\beta$ -lactam antibiotic resistance genes: a population-based modelling study. *Lancet Planet. Heal.* **3**, e357–e369 (2019).
- Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.* **31**, 1–61 (2018).
- Cantón, R., González-Alba, J. M. & Galán, J. C. CTX-M enzymes: origin and diffusion. *Front. Microbiol.* **3**, 1–19 (2012).
- Day, M. J. et al. Diversity of STs, plasmids and ESBL genes among *Escherichia coli* from humans, animals and food in Germany, the Netherlands and the UK. *J. Antimicrob. Chemother.* **71**, 1178–1182 (2016).
- 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).
- Dunn, S. J., Connor, C. & McNally, A. The evolution and transmission of multi-drug resistant *Escherichia coli* and *Klebsiella pneumoniae*:

- the complexity of clones and plasmids. *Curr. Opin. Microbiol.* **51**, 51–56 (2019).
32. Kallonen, T. et al. 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 (2017).
  33. Ludden, C. et al. One Health genomic surveillance of *Escherichia coli* demonstrates distinct lineages and mobile genetic elements in isolates from humans versus livestock. *MBio* **10**, 1–12 (2019).
  34. Robertson, J. & Nash, J. H. E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb. Genom.* **4**, 1–7 (2018).
  35. van der Graaf-van Bloois, L., Wagenaar, J. A. & Zomer, A. L. RFPPlasmid: predicting plasmid sequences from short-read assembly data using machine learning. *Microb. Genom.* **7**, 000683 (2021).
  36. Beghain, J., Bridier-Nahmias, A., Le Nagard, H., Denamur, E. & Clermont, O. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb. Genom.* **4**, e000192 (2018).
  37. Ben Zakour, N. L. et al. Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. *MBio* **7**, 1–11 (2016).
  38. Ludden, C. et al. Genomic surveillance of *Escherichia coli* ST131 identifies local expansion and serial replacement of subclones. *Microb. Genom.* **6**, e000352 (2020).
  39. Bonnet, R. et al. Host colonization as a major evolutionary force favoring the diversity and the emergence of the worldwide multidrug-resistant *Escherichia coli* ST131. *MBio* **12**, e0145121 (2021).
  40. Matsumura, Y. et al. Global *Escherichia coli* sequence type 131 clade with *bla*<sub>CTX-M-27</sub> gene. *Emerg. Infect. Dis.* **22**, 1900–1907 (2016).
  41. Roer, L. et al. ST131 fimH 22 *Escherichia coli* isolate with a *bla*<sub>CMY-2/Inc11/ST12</sub> plasmid obtained from a patient with bloodstream infection: highly similar to *E. coli* isolates of broiler origin. *J. Antimicrob. Chemother.* **74**, 557–560 (2019).
  42. Price, L. B. et al. The epidemic of extended-spectrum-β-lactamase-producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, H30-Rx. *MBio* **4**, 1–10 (2013).
  43. 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).
  44. Duggett, N. et al. The importance of using whole genome sequencing and extended spectrum beta-lactamase selective media when monitoring antimicrobial resistance. *Sci. Rep.* **10**, 19880 (2020).
  45. Massot, M. et al. Interplay between bacterial clones and plasmids in the spread of antibiotic resistance genes in the gut: lessons from a temporal study in veal calves. *Appl. Environ. Microbiol.* **87**, e0135821 (2021).
  46. Bevan, E. R., Jones, A. M. & Hawkey, P. M. Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. *J. Antimicrob. Chemother.* **72**, 2145–2155 (2017).
  47. Karim, A., Poirel, L., Nagarajan, S. & Nordmann, P. Plasmid-mediated extended-spectrum beta-lactamase (CTX-M-3 like) from India and gene association with insertion sequence ISEcp1. *FEMS Microbiol. Lett.* **201**, 237–241 (2001).
  48. Muller, M. et al. Outbreaks of multi-drug resistant *Escherichia coli* in long-term care facilities in the Durham, York and Toronto regions of Ontario, 2000–2002. *Can. Commun. Dis. Rep.* **28**, 113–118 (2002).
  49. Boyd, D. A. et al. Complete nucleotide sequence of a 92-kilobase plasmid harboring the CTX-M-15 extended-spectrum beta-lactamase involved in an outbreak in long-term-care facilities in Toronto, Canada. *Antimicrob. Agents Chemother.* **48**, 3758–3764 (2004).
  50. Mushtaq, S. Detection of CTX-M-15 extended-spectrum -lactamase in the United Kingdom. *J. Antimicrob. Chemother.* **52**, 528–529 (2003).
  51. Melo, L. C. et al. Emergence of the C1-M27 cluster in ST131 *Escherichia coli* from companion animals in France. *J. Antimicrob. Chemother.* **74**, 3111–3113 (2019).
  52. Matsumura, Y. et al. CTX-M-27- and CTX-M-14-producing, ciprofloxacin-resistant *Escherichia coli* of the H30 subclonal group within ST131 drive a Japanese regional ESBL epidemic. *J. Antimicrob. Chemother.* **70**, 1639–1649 (2015).
  53. Birgy, A. et al. CTX-M-27–Producing *Escherichia coli* of sequence type 131 and clade C1-M27, France. *Emerg. Infect. Dis.* **23**, 885–885 (2017).
  54. Ghosh, H. et al. *bla*<sub>CTX-M-27</sub> –encoding *Escherichia coli* sequence type 131 Lineage C1-M27 clone in clinical isolates, Germany. *Emerg. Infect. Dis.* **23**, 1754–1756 (2017).
  55. Tymensen, L. et al. Plasmid distribution among *Escherichia coli* from livestock and associated wastewater: unraveling factors that shape the presence of genes conferring third-generation cephalosporin resistance. *Environ. Sci. Technol.* **53**, 11666–11674 (2019).
  56. Paganini, J. A., Plantinga, N. L., Arredondo-Alonso, S., Willems, R. J. L. & Schürch, A. C. Recovering *Escherichia coli* plasmids in the absence of long-read sequencing data. *Microorganisms* **9**, 1613 (2021).
  57. Antipov, D. et al. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics* **32**, 3380–3387 (2016).
  58. Arredondo-Alonso, S. et al. mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. *Microb. Genom.* **4**, e000224 (2018).
  59. Partridge, S. R., Zong, Z. & Iredell, J. R. Recombination in IS26 and Tn2 in the evolution of multidistance regions carrying *bla*<sub>CTX-M-15</sub> on conjugative IncF Plasmids from *Escherichia coli*. *Antimicrob. Agents Chemother.* **55**, 4971–4978 (2011).
  60. Yang, Q.-E. et al. IncF plasmid diversity in multi-drug resistant *Escherichia coli* strains from animals in China. *Front. Microbiol.* **6**, 964 (2015).
  61. 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).
  62. Rozwandowicz, M. et al. Plasmids carrying antimicrobial resistance genes in *Enterobacteriaceae*. *J. Antimicrob. Chemother.* **73**, 1121–1137 (2018).
  63. French agency for food and environmental and occupational health & safety (ANSES). *Sales survey of veterinary medicinal products containing antimicrobials in France in 2020*. <https://www.anses.fr/en/system/files/ANMV-Ra-Antibiotiques2020EN.pdf> (2021).
  64. European medicines agency. *Sales of veterinary antimicrobial agents in 31 European countries in 2018*. [https://www.ema.europa.eu/en/documents/report/sales-veterinary-antimicrobial-agents-31-european-countries-2018-trends-2010-2018-tenth-esvac-report\\_en.pdf](https://www.ema.europa.eu/en/documents/report/sales-veterinary-antimicrobial-agents-31-european-countries-2018-trends-2010-2018-tenth-esvac-report_en.pdf) (2020).
  65. Government of Canada. *Canadian integrated program for antimicrobial resistance surveillance (CIPARS) 2016 Annual Report*. [https://publications.gc.ca/collections/collection\\_2018/aspc-phac/HP2-4-2016-eng.pdf](https://publications.gc.ca/collections/collection_2018/aspc-phac/HP2-4-2016-eng.pdf) (2016).
  66. Huber, L., Agunos, A., Gow, S. P., Carson, C. A. & Van Boeckel, T. P. Reduction in antimicrobial use and resistance to *Salmonella*, *Campylobacter*, and *Escherichia coli* in broiler chickens, Canada, 2013–2019. *Emerg. Infect. Dis.* **27**, 2434–2444 (2021).
  67. Moulin, G. et al. A comparison of antimicrobial usage in human and veterinary medicine in France from 1999 to 2005. *J. Antimicrob. Chemother.* **62**, 617–625 (2008).

68. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
69. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
70. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
71. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
72. Kim, D., Song, L., Breitwieser, F. P. & Salzberg, S. L. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* **26**, 1721–1729 (2016).
73. Lam, M. M. C. et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat. Commun.* **12**, 4188 (2021).
74. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
75. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
76. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
77. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Giga-science* **10**, giab008 (2021).
78. Wirth, T. et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.* **60**, 1136–1151 (2006).
79. Hunt, M. et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb. Genom.* **3**, e000131 (2017).
80. Zhou, Z., Alikhan, N.-F., Mohamed, K., Fan, Y. & Achtman, M. The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res.* **30**, 138–152 (2020).
81. Seemann, T. snippy: fast bacterial variant calling from NGS reads. <https://github.com/tseemann/snippy> (2015).
82. Zankari, E. et al. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).
83. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinforma.* **10**, 421 (2009).
84. Abudahab, K. et al. PANINI: pangenome neighbour identification for bacterial populations. *Microb. Genom.* **5**, e000220 (2019).
85. Mellor, K. C. et al. Antimicrobial resistance diversity suggestive of distinct *Salmonella Typhimurium* sources or selective pressures in food-production animals. *Front. Microbiol.* **10**, 708 (2019).
86. Oksanen, J. et al. Vegan: community ecology package. Version 2.5-6. <https://cran.r-project.org/web/packages/vegan/index.html> (2019).
87. Mather, A. E. et al. An ecological approach to assessing the epidemiology of antimicrobial resistance in animal and human populations. *Proc. R. Soc. B Biol. Sci.* **279**, 1630–1639 (2012).
88. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).
89. R Core Team. R: a language and environment for statistical computing. <https://www.r-project.org/> (2020).
90. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
91. Page, A. J. et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
92. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016).
93. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
94. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
95. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791 (1985).
96. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
97. Kassambara, A. Package 'rstatix': pipe-friendly framework for basic statistical tests (2020).

## Acknowledgements

This project was supported by the Joint Programming Initiative on Antimicrobial Resistance (JPIAMR), through the Medical Research Council (MRC, MR/R000948/1) [R.Z., J.L.N.W., A.E.M.], the Canadian Institutes of Health Research (CFC-150770) [A.C., G.C., P.B.], and the Genomics Research and Development Initiative (Government of Canada) [M.R.M.], the German Federal Ministry of Education and Research (BMBF) grant no. 01K11709 [I.E., H.K., S.S.], the French Agency for food environmental and occupational health & safety (Anses) [R.G.F., M.H., J.-Y.M.], and the French National Reference Center (CNR) for antimicrobial resistance [R. Beyrouthy, R. Bonnet]. Support was also provided by the Biotechnology and Biological Sciences Research Council (BBSRC) through the BBSRC Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent project BBS/E/F/000PR10348 (Theme 1, Epidemiology and Evolution of Pathogens in the Food Chain) [R.Z., A.E.M.]. We gratefully acknowledge use of the Quadram Institute Bioscience core facilities HPC cluster and the CLIMB-BIG-DATA-computing servers, an infrastructure supported by the UK Medical Research Council grant MR/TO30062/1. We thank Dr. George M. Savva from the core science resources at Quadram Institute Bioscience for help and guidance with the logistic regression analysis.

## Author contributions

P.B., J.-Y.M., S.S., M.R.M., R.B., H.K., J.W. and A.E.M. conceived of and designed the study. Bacterial isolation, WGS and metadata curation were done by P.B., A.C. and G.C. for human, animal, food and environmental isolates from Canada; M.R.M. and G.G.Z. for human isolates from Canada; R. Beyrouthy, and R. Bonnet for human isolates from France; M.H., J.-Y.M. and R.G.F. for animal isolates from France; and S.S., H.K. and I.E. for animal isolates from Germany. Data analysis was performed by R.Z. and A.E.M. R.Z. and A.E.M. wrote the manuscript, and all authors edited and approved the final draft of the manuscript.

## Competing interests

The authors declare no competing interests.

## Ethics statement

All bacterial genome data generated in this retrospective study were from archived extensive bacterial collections (previously approved research and surveillance studies) from researchers/co-authors of this study based at national reference and research laboratories in Canada, France and Germany. All information associated with the genome data were anonymised, thus identification of individuals is not possible and no identifiable data were used in this study. Therefore, separate ethical approval was not required.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-34970-7>.

**Correspondence** and requests for materials should be addressed to Alison E. Mather.

**Peer review information** *Nature Communications* thanks Pimlapas Leekitcharoenphon and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <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) 2022