

Comparative genomics of multi-drug resistant *Escherichia coli* ST131 from human, animal, and environmental sources in the Czech Republic

Michaela Ruzickova^{a,b,c}, Ivana Karola^a, Tomas Nohejl^{a,b}, Iva Sukkar^{a,c},
Jana Palkovicova^{a,c}, Ivo Papousek^b, Max L. Cummins^{d,e}, Steven P. Djordjevic^{d,e},
Monika Dolejska^{a,b,c,f,g,*}

^a Central European Institute of Technology, University of Veterinary Sciences Brno, Brno, Czech Republic

^b Department of Biology and Wildlife Diseases, Faculty of Veterinary Hygiene and Ecology, University of Veterinary Sciences Brno, Brno, Czech Republic

^c Department of Microbiology, Faculty of Medicine in Pilsen, Charles University, Pilsen, Czech Republic

^d Australian Institute for Microbiology and Infection, University of Technology Sydney, Australia

^e The Australian Centre for Genomic Epidemiological Microbiology, University of Technology Sydney, Australia

^f Division of Clinical Microbiology and Immunology, Department of Laboratory Medicine, University Hospital Brno, Brno, Czech Republic

^g Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Brno, Czech Republic

ARTICLE INFO

Keywords:

Escherichia coli
Antibiotic resistance
IncF plasmids
One Health
Virulence genes
Phylogeny

ABSTRACT

Globally, *Escherichia coli* is a leading cause of multiple drug-resistant (MDR) extraintestinal infections. ST131 is an MDR, pandemic sequence type that causes urinary tract and bloodstream infections in humans and companion animals, and systemic infections affecting multiple organs in poultry. ST131 has also been isolated from wastewater treatment plants and diverse wildlife sources. Studies investigating the phylogeny of ST131 from a One Health perspective are, however, limited. We performed whole-genome sequencing on 898 ST131 isolates from humans ($n = 713$), diverse environments ($n = 139$), wildlife ($n = 32$), food animals ($n = 13$), and companion animals ($n = 1$), sourced from the Czech Republic from 2005–2021, to identify evidence of clonal spread across different sources and to investigate carriage of genes encoding extended-spectrum beta-lactamases (ESBLs). ST131 belonging to clades C1 (262/898; 29 %) and C2 (528/898; 58.8 %) were dominant in our study. Irrespective of source or clade designation, most isolates (72.4 %) carried a *bla*_{CTX-M} allele, with those residing in clade C1 carrying *bla*_{CTX-M-27}, while those in clade C2 carried *bla*_{CTX-M-15}. IncF plasmid sequence types were also segregated by clades, with F29:A:B10 residing in clade A, F1:A2:B20 in clade C1, and F2:A1:B- and F31/F36:A4:B1 partitioning to clade C2. Clonal ST131 isolates (0–10 single-nucleotide polymorphisms; SNPs) were identified from humans, wild birds, and wastewater. In summary, our study showed that clonal ST131 isolates can be recovered across diverse sources in the Czech Republic, underscoring the importance of studying pandemic ExPEC lineages within a One Health context and across widespread geographic locations.

1. Introduction

Antibiotic resistance is a complex problem influencing all areas of our lives, including health of humans and animals, as well as wildlife and the environment, making it one of the most pressing global risks of the modern era. The most prevalent agents of such infections are Gram-negative bacteria, including *Escherichia coli*. One of the multidrug-resistant *E. coli* sequence types (STs) is globally diffused pandemic ST131, which is a member of the extra-intestinal pathogenic *E. coli* (ExPEC) group, causing nosocomial diseases such as urinary tract or

bloodstream infections, as well as community acquired infections, infections in poultry, wild birds, and companion animals (Pitout and Finn, 2020; Petty, Apr. 2014). Conjugative plasmids increase the possibility of antibiotic resistance genes (ARGs) and virulence-associated genes (VAGs) dissemination, resulting in this ST being a main culprit in the spread of extended-spectrum beta-lactamases (Marie-Hélène et al., Jul. 2014). The ST131 isolates frequently carry specific virulence factors, which might also provide the evolutionary advantage for this lineage to thrive over other STs, especially regarding the gastrointestinal colonisation (Mathers et al., Jul. 2015; Totsika, Oct. 2011). Along with

* Corresponding author at: Central European Institute of Technology, University of Veterinary Sciences Brno, Brno, Czech Republic.

E-mail address: monika.dolejska@gmail.com (M. Dolejska).

<https://doi.org/10.1016/j.ecoenv.2025.118320>

Received 14 February 2025; Received in revised form 9 May 2025; Accepted 10 May 2025

Available online 27 May 2025

0147-6513/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

increasing synanthropism of certain species and progressive urbanization, a cycle is also observed in which resistant strains of *E. coli* ST131, together with their plasmids and ARGs, circulate independently amongst humans, domestic animals, wildlife, and environment (White, Feb. 2024; Sato, Jun. 2024; Paulshus, Sep. 2019; Biggel et al., Nov. 2023).

E. coli ST131 is divided into five clades according to their phylogenetic relatedness and presence of various *fimH* genes for fimbriae production. Clade A is characterized by the presence of *fimH41*, and common carriage of ESBL gene *bla*_{CTX-M-27}. Clade B is defined by the presence of *fimH22* and *bla*_{CTX-M-14}. These two clades belong to the evolutionary oldest groups, having diverged from their most recent common ancestor approximately 130 years ago (Stoesser, May 2016). Evolution during their dissemination, especially under increased antimicrobial pressure, led to the emergence of a new group carrying *fimH30* gene, separating from the rest and creating clade C, around the early 1990s (Stoesser, May 2016; Ben Zakour, May 2016). This clade further split up into three clades: C0, which does not carry *bla*_{CTX-M} gene in most cases; C1 (*H30R*), with *bla*_{CTX-M-27}; and C2 (*H30Rx*), carrying *bla*_{CTX-M-15} (Stoesser, May 2016; Price, Dec. 2013). The expansion of C1 and C2 was driven by strong selective advantages conferred by fluoroquinolone resistance, encoded by chromosomal mutations in the *gyrA* and *parC* genes, which first appeared in the common ancestor of clade C. Notably, the emergence of *bla*_{CTX-M-15} in clade C2 is estimated to have occurred within a decade after the initial fluoroquinolone resistance mutations, further enhancing its global spread. Clade C2 (~50–60 %) is nowadays the most prevalent ST131 lineage worldwide, followed by clade C1 (~25–30 %), clade A (~5–10 %), and clades B and C0 (both ~1–5 %) (Pitout and Finn, 2020), (Cummins et al., 2021; Flament-Simon, Mar. 2020; Cho, Aug. 2024; Kohlenberg, Nov. 2024; Tóth, Apr. 2024).

Each clade is also linked to particular groups of IncF plasmids carrying various ARGs and VAGs (Nicolas-Chanoine, 2017; Whitmer et al., 2019). Another connection was observed between avian pathogenic *E. coli* (APEC) isolates of clade B from poultry and ColV virulence plasmids (Yang, Sep. 2015; Cummins, Jun. 2023; Reid, Feb. 2022). Similarly, uropathogenic *E. coli* (UPEC) is connected to carriage of pUTI89-like plasmid, comprising conservative virulence region *cjrABC-senB*, which might be adding to an already high virulence potential of these strains (Wyrsh, Jun. 2022), (Elankumaran et al., Jun. 2022). The IncF plasmids have been also instrumental in the rapid global spread of the ST131 clone, especially due to them harbouring high number of ARGs, VAGs and other advantageous regions, such as for the toxin-antitoxin or restriction systems. The plasticity of IncF plasmids enables them to evolve along with their hosts and causes positive selection for such strains when under selective pressure. Some combinations, e.g. clade C2 of ST131 with IncF/*bla*_{CTX-M-15} plasmid, are considered to be one of the most successful clones capable of fast dissemination throughout the globe (Phan, Apr. 2015), (Pitout and Chen, Apr. 2023).

Previous studies focused on the Czech Republic and its hospital environment and community showed that ST131 was not only one of the most common sequence types overall, but also the most prevalent CTX-M-producing clone (Kutilova, Feb. 2021), (Röderova, Jan. 2017). A genomic approach using WGS was done to assess the complex population structure of ST131 clone circulating in human and non-human sources in the Czech Republic, which has not been investigated before. Here, we aim to describe the population structure and phylogeny terms of 898 ST131 isolates from the Czech Republic, originating across various sources, and explore the association of clades with certain plasmid lineages, ARGs, and VAGs.

2. Materials and methods

2.1. Study design and sampling

ST131 isolates were picked out from a large in-house collection of 11,097 *E. coli* obtained throughout the years 2005–2021 from various

countries and sources, as a part of previous studies. The selection was performed using a PCR screening of O25 *rfb* variant and specific single-nucleotide polymorphisms (SNPs) in genes *mdh*, *gyrB* and *pabB*, targeting ST131 clone (Pitout and Finn, 2020), (Johnston and Johnson, 2017). Omission of a small number of ST131 isolates might have occurred due to filtering by the O25 variant, however, PCR screening of the SNPs increased the probability of ST131 detection. A total of 928 genomes, originating from the PCR-screened collection, as well as 98 sequences from already existing projects, were included into the analysis (Kutilova, Feb. 2021), (Röderova, Jan. 2017), (Jamborova, Jul. 2018), (Davidova-Gertzova, May 2023). Short-read sequencing was performed in order to confirm the ST, together with analysis of their genomic structure.

The complete short-read collection consisted of 1026 ST131 genomes, from which a total count of 955 (93.08 %) originated from the Czech Republic. After omitting duplicates (isolates with zero SNP difference obtained in the same location at the same time) and low-quality sequences, there was a final count of 898 short-read Czech ST131 genomes included in this study (Fig. 1). Low-quality sequences were determined by the number of contigs (threshold: 400 contigs) and the total length of the genome (threshold: 4700,000 base pairs). Out of those 898 isolates, 713 (79.40 %) were of human origin, 139 (15.48 %) came from the environment (wastewaters – 132, river – 7), 32 (3.56 %) were isolates from wild animals (rook – 18, gull – 7, raven – 6, cormorant – 1), 13 (1.45 %) were isolates from food animals (poultry – 12, swine – 1), and one isolate (0.11 %) came from a companion animal (dog). Given the predominance of isolates originating from human hosts, smaller ecological niches, such as food or companion animals, are comparatively underrepresented, potentially introducing a slight bias. The information on the source material for the entire collection of 11,097 isolates is provided in a separate sheet in the Supplementary table S1, together with the distribution rates of ST131 isolates across these sources.

The number of 619 isolates (68.93 %) were cultivated on MacConkey agar (MCA) supplemented with an antibiotic, while 279 isolates (31.07 %) were obtained by cultivation on non-supplemented MCA. Out of those 619 isolates, 406 (65.59 %) were selected on MCA with cefotaxime (2 mg/l), 201 (32.47 %) on MCA with ciprofloxacin (0.05 mg/l), 11 (1.78 %) on MCA with meropenem (0.125 mg/l), and one (0.16 %) isolate on MCA supplemented with colistin (3.5 mg/l). All metadata regarding the source, year of isolation, and antibiotic used for selective cultivation are included in the Supplementary Table S1.

2.2. Whole-genome sequencing (WGS)

Genomic DNA from all ST131 isolates was extracted using NucleoSpin® Tissue Kit (Macherey-Nagel, GmbH & Co. KG, Germany), subjected to DNA libraries preparation by Nextera® XT Library Preparation Kit (Illumina, CA, USA), and followed by sequencing using NovaSeq 6000 (Illumina) platform run in a paired-end mode. The DNA extraction was performed in our lab, and the diluted DNA was pipetted into strips and sent to the UTS Next Generation Sequencing Facility, Sydney, Australia, where the library preparation and sequencing itself was outsourced.

2.3. Data analysis and comparative genomics

Raw reads acquired by Illumina sequencing were trimmed using Trimmomatic v0.39 to remove adaptor residues and to discard low quality read regions ($Q \leq 20$) (Bolger et al., Apr. 2014), and assembled by Shovill v1.1.0 with minimum contig length set to 200 base pairs and remaining parameters set to default (<https://github.com/tseemann/shovill>). Quality check was performed with FastQC v0.11.9 (<https://github.com/s-andrews/fastqc>) for reads, and contamination check with Kraken2 v2.0.8 used for assemblies with the MiniKraken2 database v2.201904 (<https://github.com/DerrickWood/kraken2>). All data regarding the quality of the sequencing runs and assemblies are

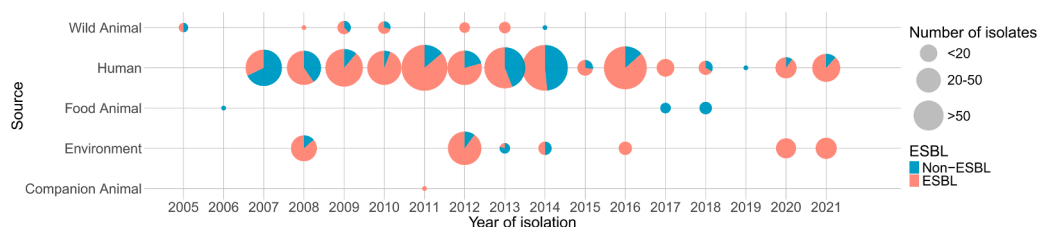


Fig. 1. Timeline of obtained isolates by source and production of ESBLs.

provided in a separate sheet in the [Supplementary table S1](#). Clade assignment was based on the presence of specific genes and point mutations (see [Supplementary Table S2](#)) and was performed using a previously published protocol (Nicolas-Chanoine, 2017).

Fast hierarchical Bayesian analysis of population structure (fastBAPS) was used for clustering the genomes into phylogenetically related groups. These groups were later compared to the clusters obtained by the clade assignment. The fastBAPS analysis was performed on two levels of hierarchy, with the second level delving deeper and dividing the clusters created by the first level into more subclusters (Tonkin-Hill et al., Jun. 2019). The abriTAMR v1.0.10 pipeline was used with the abriTAMR database v2021–09–30.1 to explore the presence of ARGs, VAGs, and point mutations encoding fluoroquinolone resistance, with species defined as *Escherichia* for the latter analysis (<https://github.com/MDU-PHL/abritamr>). The threshold for ARGs and VAGs analyses was set to 90 % coverage and 90 % identity, with isolates above this threshold being considered as carriers of respective genes. Center for Genomic Epidemiology tools FimTyper v1.0, SeroTypeFinder v2.0.1, PlasmidFinder v2.0.1, and pMLST v0.1.0 were used for the assessment of plasmid presence, plasmid sequence type, *fimH* gene variant, and serotype, respectively (<https://cge.cbs.dtu.dk/services/>). Existing methodology was adjusted and used for establishing the virotype of isolates (Marie-Hélène et al., Jul. 2014), (Dahbi, Nov. 2014), (Blanco, Oct. 2013), (Barrios-Villa, Dec. 2018). The assignment was based on the presence of following genes: *afa*, *cdtB*, *cnf1*, *hlyA*, *ibeA*, *ironK1*, *kpsMII-K2*, *kpsMII-K5*, *papGII*, *papGIII*, and *sat*; and the output was simplified to group the isolates into five virotypes A, C, D, E, and F. Combinations of genes above assigned to the respective virotypes can be seen in the [Supplementary Table S3](#).

Prokka v1.14.112 (Seemann, Jul. 2014) was used for genome annotation of isolates' assemblies, with their pangenome aligned using PIRATE v1.0.4 (Bayliss et al., Oct. 2019). Subsequently, the alignment was used to generate phylogenetic tree in IQ-TREE v1.6.12 under GTR+GAMMA model, supported by 1000 bootstraps (Nguyen et al., Jan. 2015). The alignment was run through a recombination filtering software Gubbins v3.1.0 with default settings (Croucher, Feb. 2015). Phylogenetic distances between isolates were estimated after the recombination filtering, using the pangenome alignment in snp-dists v0.6.3 (<https://github.com/tseemann/snp-dists>), considering the number of SNPs. The distance between two *E. coli* isolates lower or equal to 10 SNP was set as a threshold for evaluating such isolates as clonal (Schürch et al., Apr. 2018). The phylogenetic tree was visualized in iTOL v5.7.15 (Letunic and Bork, Apr. 2024). Statistical computing, bar charts, and SNP distance matrices were created using custom R (Team) scripts, and utilizing R packages readr v2.1.5 (Wickham et al., 2024), magrittr v2.0.3 (Bache and Wickham, 2022), dplyr v2.4.0 (Wickham et al., 2023), tidyverse v2.0.0 (Wickham, Nov. 2019), ggplot2 v3.5.0 (Wickham, 2016), pheatmap v1.0.12 (Kolde, 2019), ggtree v3.6.2 (Yu, Mar. 2020), ape v5.7–1 (Paradis and Schliep, Feb. 2019), phytools v2.1–1 (Revell, Jan. 2024), and highcharter v0.9.4 (<https://github.com/jbkunst/highcharter>). The heatmap in Fig. 7 and heatmaps in Supplementary files S6–S10 were generated using an edited publicly available script (https://github.com/maxcummins/ST95/blob/main/scripts/Supplementary_Figure_1_SNP_matrix.R) (Cummins et al., Feb. 2022). Statistical analysis used the Shapiro–Wilk test of normality, followed by

the Mann–Whitney test and T-test, performed on non-parametric and parametric datasets, respectively. Results from the statistical analysis, including the selected tests, can be found in the [Supplementary Table S4](#).

3. Results

3.1. The most widespread clades C1 and C2 also most prevalent in the collection

The phylogenetic tree based on the analysis of SNP distances showed clear distribution of 898 *E. coli* ST131 genomes into clades A, B, and C divided further into clades C0, C1, and C2 (Fig. 2). Clades were assigned using the presence of *fimH41*, *fimH22*, and *fimH30* alleles, respectively. Isolates with different *fimH* alleles were allotted to clades based on their phylogenetic closeness to the clusters of those clades. The C1 and C2 clades were then determined by the presence of any *gyrA* and *parC* chromosomal mutations, and distinguished by the presence of *bla_{CTX-M-27}* and *bla_{CTX-M-15}* genes, respectively, while no *gyrA* and *parC* chromosomal mutations were present in the C0 clade. Out of 898 isolates, clade C2 was the most prevalent one (59.58 %; *n* = 535), followed by clades C1 (28.51 %; *n* = 256), A (5.46 %; *n* = 49), B (5.01 %; *n* = 45), and C0 (1.45 %; *n* = 13). There were minor discrepancies found in eight isolates and their clade assignment, which are highlighted in the tree (Fig. 2). Two of these unusual isolates were assigned into clade C1 instead of C0 based on the phylogenetic distances, even though the *gyrA* and *parC* mutations were missing. Six of the isolates were grouped with clade C2 instead of C1 due to the tree branching, even though the IncF pMLST, *bla_{CTX-M}* gene type, and/or fastBAPS grouping suggested otherwise. The tree-building tool regards the complete genome, while the clade-defining characteristics consider only few selected genes, which might have caused the discrepancies in the clade assignment, with those eight isolates being genetically more similar to others regardless of the presence of the defining genes.

There were no apparent host associations with certain clades; however, this might be confounded by sampling bias, as a large majority (79.40 %) of the analysed genomes originated from human hosts. Such disproportionate representation might have reduced the ability to detect potential host-specific associations in underrepresented host groups. Chromosomal mutations *gyrA* and *parC*, contributing to the fluoroquinolone resistance, were present only in clades C1 (99.22 %) and C2 (100 %). All such isolates carried *gyrA* S83L (*n* = 790) followed with *gyrA* D87N (*n* = 788), *parC* S80I (*n* = 789), and *parC* E84V (*n* = 788). The majority of the genomes (98.11 %; 881/898) carried IncF plasmids, with F2:A1:B- being the most prevalent one (42.57 %; 375/881), especially in clade C2 (69.91 %; 374/535). Five different virotypes were observed across all clades, with A being the most common one (44.54 %; 400/898), followed by C (36.19 %; 325/898), E (12.03 %; 108/898), D (0.05 %; 45/898), and F (0.01 %; 4/898). There were genomes from various clades and sources with unidentifiable virotype due to unusual combination of determining genes. Strong association of clade C2 and virotypes A (66.73 %; 357/535) and E (18.32 %; 98/535) was observed. Clade C1 was then linked mostly to virotype C (96.48 %; 247/256), and clade B to virotype D (100 %; 45/45) (Fig. 3).

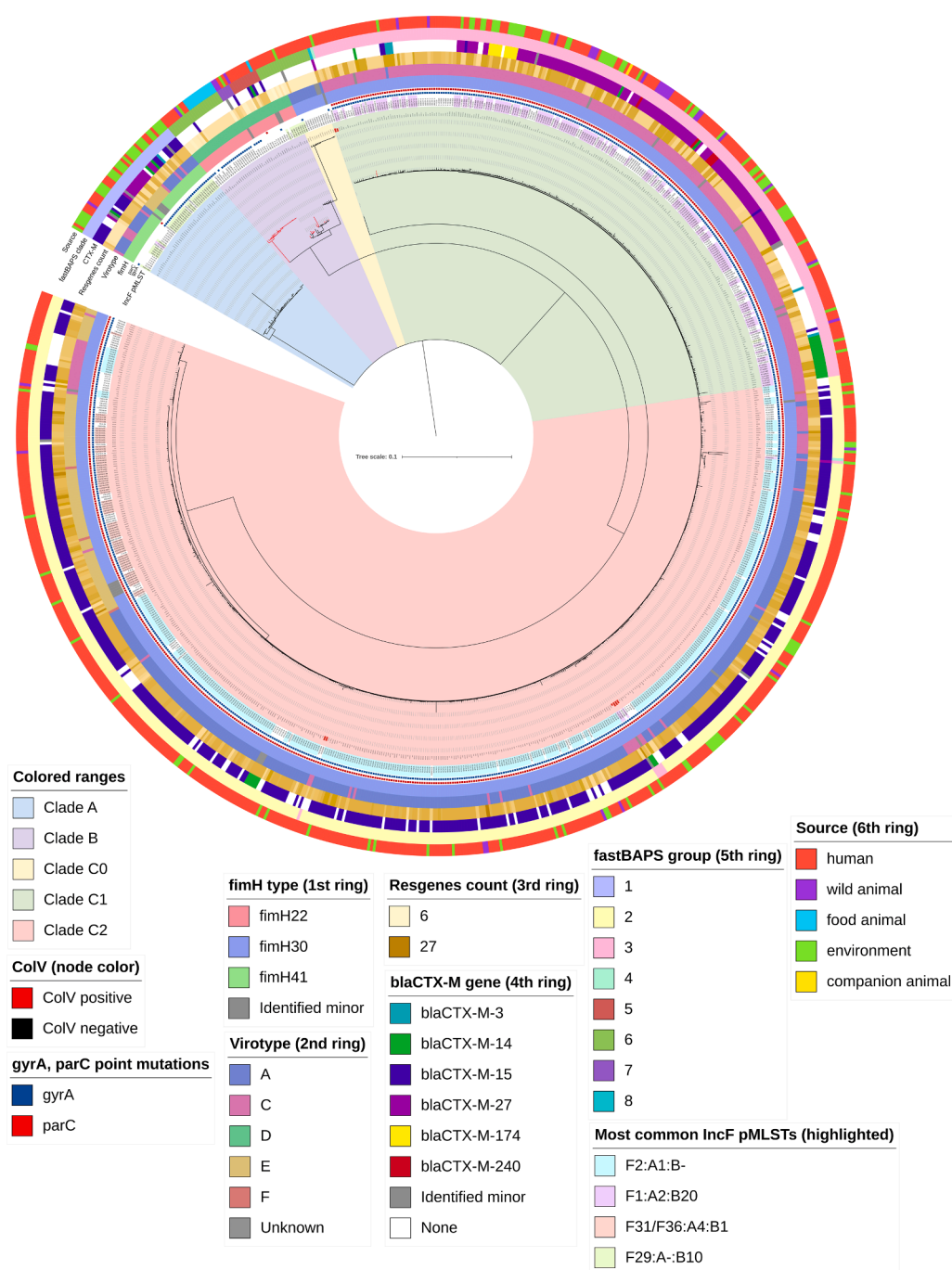


Fig. 2. SNP-based phylogenetic tree of all ST131 genomes. The tree is sectioned into five groups by clade assignment (A-blue, B-purple, C0-yellow, C1-green, C2-red). Firstly, the presence of IncF plasmids is denoted with the four most common pMLSTs highlighted in colour. Smaller squares then describe the presence of chromosomal *gyrA* and *parC* mutations. Further rings show *fimH* type, viotype, number of ARGs, *bla*_{CTX-M} gene type, fastBAPS clade assignment, and source of the isolate. Red branches display the ColV plasmid carriage. Name labels highlighted in red point out genomes whose clade assignment was not completely certain.

3.2. Higher variety in ARGs and VAGs is between clades than sources

The whole collection was screened for the presence of plasmids, IncF pMLSTs, ARGs, and VAGs. A total of 85 different ARGs were found, with clade C2 being the most resistant one (Fig. 4). It carried the highest average amount of ARGs ($n = 18.96$), with numbers ranging from 12 to 27 genes per isolate, and the genes associated with resistance to at least seven antibiotic classes in a majority of isolates. The lowest ARGs carriage was observed in clade C0 (8.54; 8–15), which also contained the highest number of isolates obtained by cultivation on media without any antibiotics. There was, however, no significant difference in the carriage

of ARGs between isolates cultivated on media with antibiotics and those without (Fig. 4A).

Considering the content of VAGs, clade B was the one with the significantly highest average amount ($n = 21.53$), with numbers ranging from 8 to 39 genes per isolate, compared to clades C0 (17.00; 17), C2 (16.99; 9–28), C1 (14.84; 6–32), and A (14.69; 8–20) (Fig. 4C). Prevalence of ColV plasmids was also significantly higher in clade B (73.33 %; 33/45), as opposed to clades A (0 %; 0/49), C0 (0 %; 0/13), C1 (0.39 %; 1/256), and C2 (0.19 %; 1/535). There was a substantial difference in carriage of VAGs among isolates selected and not selected on media with antibiotics, where the non-selected ones displayed higher

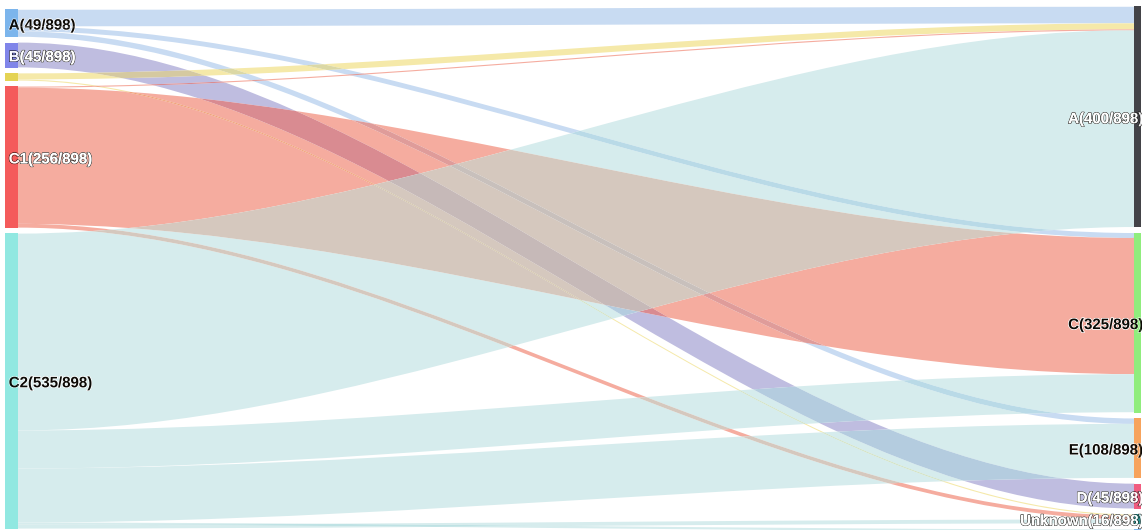


Fig. 3. Association of clades with certain virotypes. Clade on the left, virotype on the right. Interactive version also available at <https://ruzickm.github.io/ST131/Figure%203.html>.

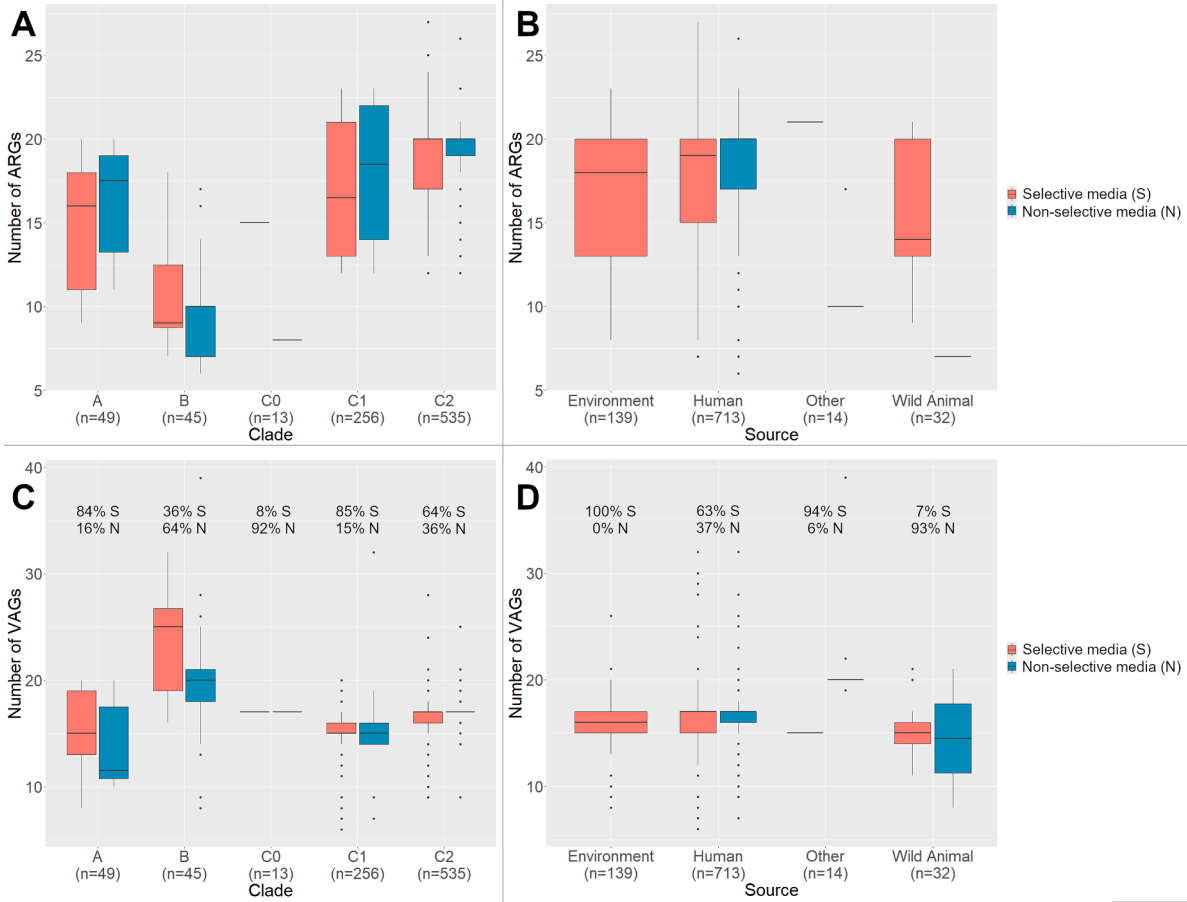


Fig. 4. Number of ARGs and VAGs across clades and sources. Fig. 4A denotes ARGs by clade, Fig. 4B ARGs by source, Fig. 4C VAGs by clade, and Fig. 4D. VAGs by source. Each part is divided into isolates that were selected on media with an antibiotic (S-red) and those that were not (N-blue). Percentage for each group is shown in the chart. The “Other” category in source groups together isolates from food animals (n = 13) and companion animals (n = 1).

carriage. This might have been caused by uneven distribution of selected and non-selected isolates between sources and clades. All p-values can be found in the [Supplementary Table S4](#).

Regarding the carriage by source, the highest number of ARGs was

found in isolates from humans (n = 17.96) ranging from 6–27 genes, as opposed to isolates from the environment (16.84; 8–21), wild animals (15.06; 7–21), and food animals (10.54; 10–17) (Fig. 4B). The highest number of VAGs then occurred in isolates from food animals, with an

average of 21.54 genes and ranging from 19–39 genes, compared to isolates from humans (16.58; 6–32), environment (15.63; 8–26), and wild animals (15.41; 8–21) (Fig. 4D). Again, the isolates sourced from food animals were solely of clade B and displayed high carriage of the ColV plasmid (12/13; 92.31 %). The isolate from a companion animal was excluded from this statistical analysis due to the small size of the group. To sum up, the largest differences were observed between livestock isolates of clade B and human isolates of clade C1 and C2 combined, with the carriage of ARGs significantly higher in isolates of human origin, and the carriage of VAGs higher in those from food animals.

3.3. Resistance to at least five groups of antibiotics present in all clades and sources

Regarding the class of ARGs, isolates originating from humans displayed the highest carriage of genes associated with resistance to aminoglycosides, macrolides, tetracyclines, and trimethoprim, as opposed to the “Other” category, sourced from food and companion animals, which mostly showed a carriage of genes linked to resistance to beta-lactams and quinolones (Fig. 5B). There was at least one ESBL gene present in 75.95 % (n = 682) of genomes, with no significant difference observed between isolates selected on media with and without antibiotics. The most common was *bla*_{CTX-M-15} (52.01 %; 467/898), found majorly in the C2 clade, and *bla*_{CTX-M-27} (15.48 %; 139/898), found mainly among C1 isolates. There was a significantly lower amount of ESBL genes in clades B and C0 (13.33 % and 15.22 %, respectively), as

opposed to clade A (84.00 %) and clade C2 (82.61 %) (Fig. 5A). The statistical analysis can be found in the [Supplementary Table S4](#).

Genes related to resistance to aminoglycosides, sulfonamides, and tetracyclines were observed throughout all clades (Fig. 5). In four clades, excluding the least represented clade C0, genes noted for resistance to macrolides, quinolones, and trimethoprim were present. Out of all eight aminoglycoside resistance genes found, the most common were *aac*(6′)-*Ib-cr*5 (70.33 %; n = 422) and *aac*(3)-*Ile* (17.67 %; n = 106). There were three genes associated with resistance to macrolides, with *mph*(A) being the most prevalent one (99.48 %; n = 573). Resistance to quinolones was suggested by four various genes, with the most common one being *aac*(6′)-*Ib-cr* (91.26 %; n = 428), and the rest represented by the *qnr* genes. Three genes linked to resistance to sulfonamides were found, with the most common *sul*1 (73.71 %; n = 569), *sul*2 (25.91 %; n = 200), and then *sul*3 (0.13 %; n = 1), together with one point mutation *folP* P64S (0.26 %; n = 2). Tetracycline resistance was indicated by two genes, with prevailing *tet*(A) (96.68 %; n = 466). Five genes associated with resistance to trimethoprim were found, with *dfr*A17 (93.04 %; n = 561) being the most common one. Genes noted for resistance to carbapenems, phenicols, and colistin were observed in a minority of isolates, and thus did not contribute to the multidrug-resistant status of any particular clade.

3.4. Clade association with specific IncF plasmids

Most of the clades showed a connection to specific IncF plasmids (Fig. 6). F29:A-B:10 plasmid, also known as pUTI89-like, was mostly

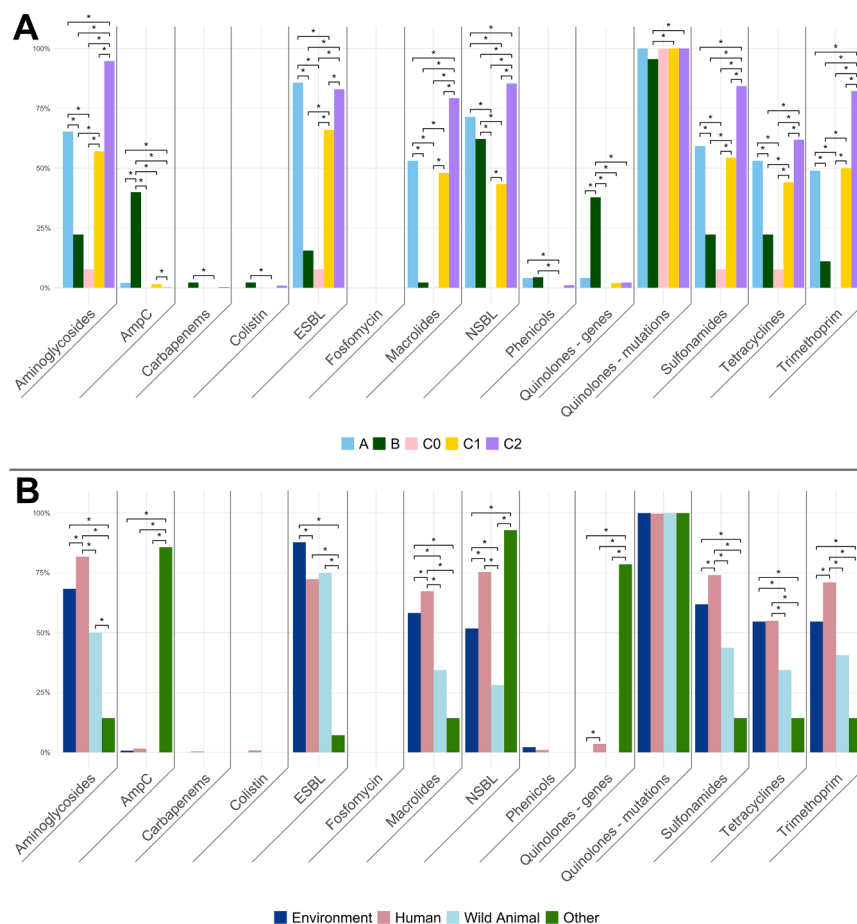


Fig. 5. Prevalence of resistance genes by antimicrobial class, clade, and source. Fig. 5A shows resistance rate in each phylogenetic clade while Fig. 5B focuses on the source of the isolate. Each antibiotic group has its own section except for quinolones which are further split by ARGs and chromosomal mutations. Resistance to beta-lactams is also divided into AmpC, ESBL and NSBL groups which stand for genes encoding the production of AmpC, extended-spectrum and narrow-spectrum beta-lactamases, respectively. Significant differences between groups are denoted by asterisks (p-value < 0.05).

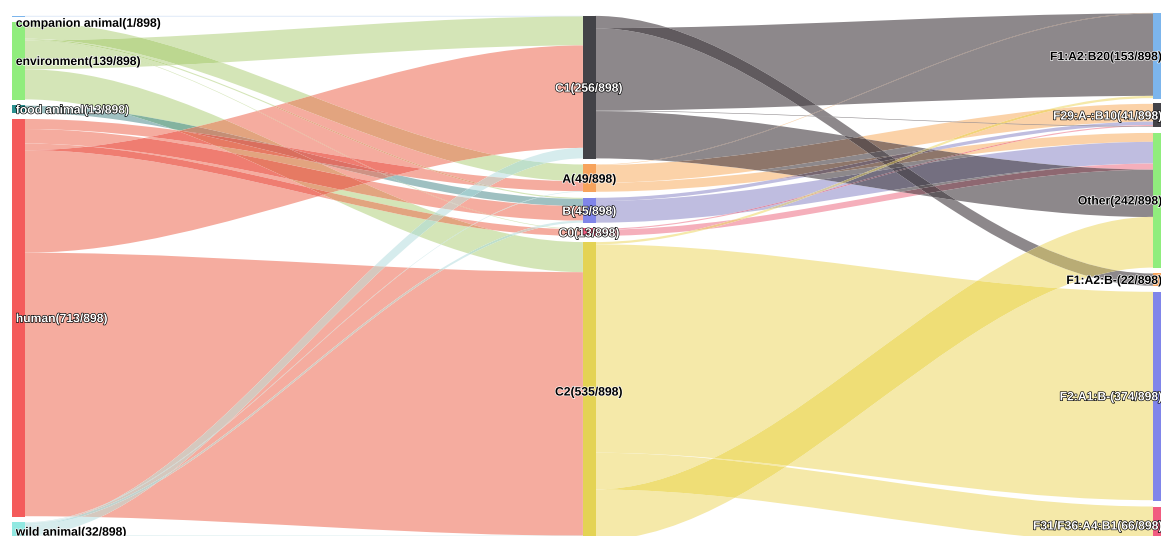


Fig. 6. Association of clades with sources and IncF plasmid sequence types. Source on the left, clade in the middle, plasmid sequence type on the right. Interactive version also available at <https://ruzickm.github.io/ST131/Figure%206.html>.

carried by clade A isolates (62.96 %; 34/54). Another strong association was observed in clade C1 with F1:A2:B20 plasmid (57.81 %; 148/256), and clade C2 with F2:A1:B- plasmid (69.91 %; 374/535). Clade C2 also displayed high carriage of F31/F36:A4:B1 plasmid (12.34 %; 66/535). Clade B isolates were linked with various sequence types of IncF plasmids, most of them falling into a subgroup of ColV plasmids, with C4/F18:A-B6 (42.42 %; 14/33) and F2:A-B1 (27.27 %; 9/33) being the most prevalent plasmid STs.

3.5. Source-diverse isolates show highest clonality in clades C1 and C2

Highest phylogenetic relatedness was observed amongst the most represented and most source-diverse clades C1 and C2, with average SNP difference after recombination filtering of 130.26 and 150.25 SNPs, respectively (Supplementary Table S5). Clade C0 was excluded from these statistics, as it only consisted of 13 isolates originating mostly from the same sampling site. Within C1 clade, highest clonality was observed among a subset originating from humans, wild birds, and wastewaters with SNP differences often ranging between 0–10 SNPs (Fig. 7); a trend

also observed in clade C2 within groups obtained from humans and wastewaters. No significant pattern in phylogenetic relatedness was observed amongst isolates sampled during different years. This was expected, since the sampling did not occur repeatedly at the same place.

There was no considerable close relatedness between source-diverse isolates from clades A and B; however, clustering of isolates from the same sources was visible. This was observed especially in clade B, with isolates originating from food animals, which were collected within a single study. Complete SNP heatmaps of all clades A, B, C0, C1, and C2 can be found in Supplementary files S6–S10, respectively.

4. Discussion

E. coli ST131 lineage in Europe was first reported around the year 2003 in eight different countries (Spain, Russia, Italy, Germany, France, Poland, Hungary, and Austria) and quickly started to spread further (United Kingdom, Netherlands, Sweden) (Marie-Hélène et al., Jul. 2014), (Rogers et al., Jan. 2011), (Cagnacci et al., Aug. 2008). First reports of this ST in the Czech Republic originated from the year 2006 and

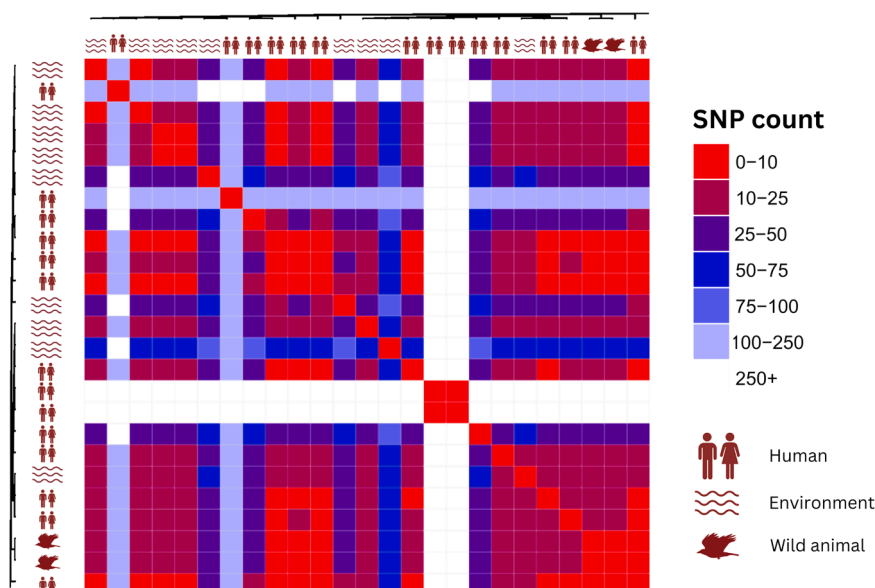


Fig. 7. Clade C1 heatmap cutout displaying close relatedness amongst isolates from humans, environment, and wild animals.

were associated with hospital outbreaks (Novais, Jul. 2012). After that, ST131 disseminated to other environments, such as wastewaters (Dolejska, Dec. 2011), wild birds (Tausova, May 2012), food animals (Röderova, Jan. 2017), and patients from the community (Papagiannitsis, Feb. 2015). Our research describes 898 *E. coli* ST131 genomes collected in the Czech Republic throughout the years 2005–2021, which shows even earlier emergence of this clone (Fig. 1). The four isolates obtained in the year 2005 originated from wild birds (three from gulls, one from a rook), which is notable given ST131 in the Czech Republic was reported only in hospital-related environments during that time (Novais, Jul. 2012), (Hrabák, Oct. 2009). Our human samples also displayed rising number of ESBL-positive isolates, with 32.20 % in the year 2007 and 88.57 % in the year 2021. Even though there are also several years where the number of ESBL-positive isolates decreased, the trend is overall increasing, same as in the samples obtained from the environment. It is also interesting to note that no isolates obtained from food-producing animals carried genes for ESBL production.

Our collection was divided into five clades: A, B, C0, C1, and C2, with the clade distribution correlating with trends from other parts of the world (Petty, Apr. 2014; Marie-Hélène et al., Jul. 2014), (Li, Dec. 2021; Ludden, Apr. 2020). Clades C1 and C2 proved to be the most resistant ones, carrying the largest amount of ARGs associated with resistance to at least seven different antibiotic classes in the majority of isolates. These clades were distinguished from the rest by the presence of both chromosomal point mutations *gyrA* and *parC*, encoding resistance to fluoroquinolones, providing them with further evolutionary advantage and causing the success of such strains (Marie-Hélène et al., Jul. 2014). We observed these mutations in other clades as well, however, except for one isolate in clade A, both of them never occurred together as in clades C1 and C2. High carriage of ESBL genes observed in clade C2 is based on the sole classification of the ST131 lineage. This clade was originally the first one to contain *bla*_{CTX-M-15} on an IncF plasmid, which has been associated with the rapid spread of this subclone (Stoesser, May 2016), (Ludden, Apr. 2020). During dissemination of ST131, clade C1 also obtained an IncF plasmid, this time carrying the *bla*_{CTX-M-27}, the predominant ESBL gene detected in our C1 genomes as well (Stoesser, May 2016). In our collection, clade A frequently contained ESBL genes, mostly *bla*_{CTX-M-15} or *bla*_{CTX-M-27}, as opposed to clades B and C0, which displayed lower carriage. This is linked to overall lower ESBL occurrence observed in these two clades, whereas clades A, C1, and C2 are known for high resistance levels and common production of CTX-M beta-lactamases (Price, Dec. 2013), (Li, Dec. 2021).

Clades C1 and C2 proved to be the most clonal, showing low SNP differences even between isolates from different sources and years. This correlates with their widest dissemination within the ST131 lineage, with their clones circulating in all kinds of environment, easily exchanging mobile genetic elements, and leading to a stabilization of the most successful strains. Close relatedness of source-diverse isolates from clades C1 and C2 correlates with previous studies, as opposed to research focused on clade B, where we observed low SNP differences only within isolates from the same source (Li, Dec. 2021), (Reid et al., Sep. 2019). In clade C1, we observed a high degree of genetic similarity among isolates originating from humans, wild birds, and wastewater. This suggests that anthropogenically impacted environments, particularly wastewater, act as a reservoir or transmission hub, facilitating the spread of resistant bacteria to wildlife through indirect exposure, such as contaminated foraging sites or water runoff. However, river water isolates showed lower genetic relatedness to those from other sources, indicating that while wastewater potentially plays a key role in the transmission chain, downstream environments could dilute the bacterial load or fail to select for resistant lineages.

Strong associations of clades with certain plasmids were observed in all clades, except C0. Genomes of clade A carried mostly F29:A:B10 plasmid, also known as pUTI89-like plasmid. This plasmid contains a conservative virulence region *cjrABC-senB* and is linked to *E. coli* strains

of various STs (ST95, ST127, ST131, ST141, ST961, ST963, ST1193), causing urinary tract infections (UTIs) with nosocomial potential (Wyrsh, Jun. 2022), (Elankumaran et al., Jun. 2022). Our clade A isolates carrying this plasmid originated from not only humans, but also municipal wastewaters, river water, and a gull, thus displaying a wide range of host environments outside of the hospital settings. Clade B showed high carriage of ColV plasmids, with the most prevalent plasmid ST C4/F18:A:B6. These plasmids are a subgroup of IncF plasmids, carrying large number of VAGs, and are commonly associated with APEC clones, found especially in poultry (Yang, Sep. 2015), (Cummins, Jun. 2023; Reid, Feb. 2022). Our clade B genomes displayed significantly higher content of VAGs compared to other clades, which was attributed to the presence of the ColV plasmid. They were commonly found in poultry with APEC infections, but other sources also included humans, wastewaters, and wild birds, such as gulls or rooks. Clade C1 carried mainly an IncF plasmid of F1:A2:B20 sequence type. This was anticipated, since this plasmid type is considered to be one of the determinants responsible for shaping the evolution of this clade (Johnson, Aug. 2016). The same situation can be observed in clade C2, with most of the genomes carrying the F2:A1:B- plasmid, which was again one of the main factors distinguishing this clade from the rest during its evolution (Johnson, Aug. 2016). Even though these plasmids transferred among isolates from clades C1 and C2 in the past, as well as now, there is still a clear association to be observed. A smaller part of our C2 genomes carried also an IncF plasmid with sequence type F31/F36:A4:B1. All these isolates, except for one, belonged to a virotype E and formed a highly clonal phylogenetic sublineage. This plasmid often carries a genomic island *GI-pheU* with *cnf1*/hemolysin genes, increasing its virulence and creating a pandemic clone out of this sublineage, while underscoring the need for further study (Omid, Aug. 2021; Hojabri, Nov. 2020). We observed higher presence of virotype E, as opposed to other studies in which this virotype was often not even detected, except for a single research (Blanco, Oct. 2013), (Kondratyeva et al., Jan. 2020). However, the rest of the overall distribution of virotypes roughly corresponded with previous research (Blanco, Oct. 2013), (Li, Dec. 2021). This was displayed especially in clade B, where all isolates belonged to virotype D, and vice versa (Li, Dec. 2021), (Hojabri, Nov. 2020). The connection between clade B and virotype D is crucial, as this combination is strongly linked to avian pathogenic *E. coli* (APEC) infections in poultry and other birds. Additionally, the frequent presence of ColV plasmids, carrying high amounts of VAGs, further strengthens this lineage and likely contributes to its efficient transmission within poultry production environments (Li, Dec. 2021), (Flament-Simon, Mar. 2020; Merino, Dec. 2020). The strong association between clade C1 and virotype C is also particularly important, as these isolates often represent UPEC strains capable of causing UTIs. The intrinsic fluoroquinolone resistance characteristic of the C1 clade, then further enhances their fitness, clinical impact, and potential for widespread dissemination (Merino, Dec. 2020).

5. Conclusion

This study presents a detailed genomic investigation of 898 isolates of *E. coli* ST131, collected in the Czech Republic between the years 2005–2021, revealing the early emergence and persistent spread of this multidrug-resistant lineage across diverse environments. The collection included isolates from human, animal, and environmental sources, and covered all five recognized ST131 clades. Notably, clades C1 and C2 showed high clonality and wide host and environmental distribution, reflecting their evolutionary success and adaptive potential. Distinct plasmid-clade associations were identified, with specific F-type plasmids contributing to the dissemination of resistance and virulence traits. The detection of highly similar isolates across unrelated sources highlights the importance of anthropogenic environments in bacterial transmission and underscores the role of mobile genetic elements and selective pressures in shaping the success of dominant ST131 subclones. These

findings provide valuable insight into the epidemiology of ST131 in a single European country and demonstrate the significance of exploring previously overlooked areas to better understand the mechanisms driving the spread of antimicrobial resistance.

Funding

This work was supported by the Ministry of Health of the Czech Republic (NU22-09-00645, DRO FNBr, 65269705), Internal Grant Agency of University of Veterinary Sciences Brno (209/2024/FVHE), Czech Science Foundation (24-12527S), and National Institute of Virology and Bacteriology (Programme EXCELES, ID Project No. LX22NPO5103) funded by the European Union-Next Generation EU.

Transparency declarations

None to declare.

CRediT authorship contribution statement

Ivana Karola: Writing – review & editing, Methodology, Investigation, Formal analysis. **Michaela Ruzickova:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ivo Papousek:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Jana Palkovicova:** Writing – review & editing, Software, Investigation, Formal analysis, Data curation. **Iva Sukkar:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Tomas Nohejl:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Monika Dolejska:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Steven P Djordjevic:** Writing – review & editing, Validation, Supervision, Conceptualization. **Max L Cummins:** Writing – review & editing, Software, Investigation, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Dana Cervinkova, Kristina Nesporova, Jana Siskova, Jiri Cejnar, Martin Klvana, Petra Sismova, Marketa Zelendova, Lenka Dostalova, Katerina Rehakova, Hassan Tarabai, Jiri Kolar, Adam Valcek and Eva Suchanova for the collection of isolates and corresponding lab work. We also want to thank Ivan Literak, Alois Cizek and Martina Masarikova for obtaining part of the samples, and to Lenka Davidova Gerzova and Jarmila Lausova for providing part of the WGS data.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ecoenv.2025.118320](https://doi.org/10.1016/j.ecoenv.2025.118320).

Data availability

Data accessibility is specified in the Data availability section of the manuscript. Bioproject numbers can be found in the Supplementary table S1 for each isolate.

References

Bache S. and Wickham H., “_magrittr: A Forward-Pipe Operator for R_ R package version 2.0.3,” 2022.

- Barrios-Villa, E., et al., Dec. 2018. Adherent/invasive *Escherichia coli* (AIEC) isolates from asymptomatic people: new *E. coli* ST131 O25:H4/H30-Rx virotypes. *Ann. Clin. Microbiol. Antimicrob.* 17 (1), 42. <https://doi.org/10.1186/s12941-018-0295-4>.
- Bayliss, S.C., Thorpe, H.A., Coyle, N.M., Sheppard, S.K., Feil, E.J., Oct. 2019. PIRATE: A fast and scalable pangenomics toolbox for clustering diverged orthologues in bacteria. *Gigascience* 8 (10). <https://doi.org/10.1093/gigascience/giz119>.
- Ben Zakour, N.L., et al., May 2016. Sequential Acquisition of Virulence and Fluoroquinolone Resistance Has Shaped the Evolution of *Escherichia coli* ST131. *mBio* 7 (2). <https://doi.org/10.1128/mBio.00347-16>.
- Biggel, M., Hoehn, S., Frei, A., Dassler, K., Jans, C., Stephan, R., Nov. 2023. Dissemination of ESBL-producing *E. coli* ST131 through wastewater and environmental water in Switzerland. *Environ. Pollut.* 337, 122476. <https://doi.org/10.1016/j.envpol.2023.122476>.
- Blanco, J., et al., Oct. 2013. Four Main Virotypes among Extended-Spectrum- β -Lactamase-Producing Isolates of *Escherichia coli* O25b:H4-B2-ST131: Bacterial, Epidemiological, and Clinical Characteristics. *J. Clin. Microbiol.* 51 (10), 3358–3367. <https://doi.org/10.1128/JCM.01555-13>.
- Bolger, A., Lohse, M., Usadel, B., Apr. 2014. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30. <https://doi.org/10.1093/bioinformatics/btu170>.
- Cagnacci, S., Gualco, L., Debbia, E., Schito, G.C., Marchese, A., Aug. 2008. European Emergence of Ciprofloxacin-Resistant *Escherichia coli* Clonal Groups O25:H4-ST 131 and O15:K52:H1 Causing Community-Acquired Uncomplicated Cystitis. *J. Clin. Microbiol.* 46 (8), 2605–2612. <https://doi.org/10.1128/JCM.00640-08>.
- Cho, S.-T., et al., Aug. 2024. Evolution of extended-spectrum β -lactamase-producing ST131 *Escherichia coli* at a single hospital over 15 years. *Sci. Rep.* 14 (1), 19750. <https://doi.org/10.1038/s41598-024-70540-1>.
- Croucher, N.J., et al., Feb. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *e15 Nucleic Acids Res* 43 (3), e15. <https://doi.org/10.1093/nar/gku1196>.
- Cummins, E.A., Snaith, A.E., McNally, A., Hall, R.J., 2021. The role of potentiating mutations in the evolution of pandemic *Escherichia coli* clones. *Springer Sci. Bus. Media Dtschl. GmbH*. <https://doi.org/10.1007/s10096-021-04359-3>.
- Cummins, M.L., et al., Jun. 2023. Whole Genome Sequencing of Avian Pathogenic *Escherichia coli* Causing Bacterial Chondronecrosis and Osteomyelitis in Australian Poultry. *Microorganisms* 11 (6), 1513. <https://doi.org/10.3390/microorganisms11061513>.
- Cummins, M.L., Reid, C.J., Djordjevic, S.P., Feb. 2022. F Plasmid Lineages in *Escherichia coli* ST95: Implications for Host Range, Antibiotic Resistance, and Zoonoses. *mSystems* 7 (1). <https://doi.org/10.1128/mSystems.01212-21>.
- Dahbi, G., et al., Nov. 2014. Molecular epidemiology and virulence of *Escherichia coli* O16:H5-ST131: Comparison with H30 and H30-Rx subclones of O25b:H4-ST131. *Int. J. Med. Microbiol.* 304 (8), 1247–1257. <https://doi.org/10.1016/j.ijmm.2014.10.002>.
- Davidova-Gerzova, L., et al., May 2023. Hospital and community wastewater as a source of multidrug-resistant ESBL-producing *Escherichia coli*. *Front Cell Infect. Microbiol.* 13. <https://doi.org/10.3389/fcimb.2023.1184081>.
- Dolejska, M., et al., Dec. 2011. CTX-M-15-producing *Escherichia coli* clone B2-O25b-ST131 and *Klebsiella* spp. isolates in municipal wastewater treatment plant effluents. *J. Antimicrob. Chemother.* 66 (12), 2784–2790. <https://doi.org/10.1093/jac/dkr363>.
- Elankumaran, P., Cummins, M.L., Browning, G.F., Marends, M.S., Reid, C.J., Djordjevic, S.P., Jun. 2022. Genomic and Temporal Trends in Canine ExPEC Reflect Those of Human ExPEC. *Microbiol Spectr.* 10 (3). <https://doi.org/10.1128/spectrum.01291-22>.
- Flament-Simon, S.-C., et al., Mar. 2020. High Prevalence of ST131 Subclades C2-H30Rx and C1-M27 Among Extended-Spectrum β -Lactamase-Producing *Escherichia coli* Causing Human Extraintestinal Infections in Patients From Two Hospitals of Spain and France During 2015. *Front Cell Infect. Microbiol.* 10. <https://doi.org/10.3389/fcimb.2020.00125>.
- Flament-Simon, S.-C., et al., Mar. 2020. Whole Genome Sequencing and Characteristics of mcr-1-Harboring Plasmids of Porcine *Escherichia coli* Isolates Belonging to the High-Risk Clone O25b:H4-ST131 Clade B. *Front Microbiol.* 11. <https://doi.org/10.3389/fmicb.2020.00387>.
- Hojabri, Z., et al., Nov. 2020. Expansion of a Subset Within the C2 Subclade of *Escherichia coli* Sequence Type 131 (ST131) Is Driving the Increasing Rates of Aminoglycoside Resistance. *Open Forum Infect. Dis.* 7 (11). <https://doi.org/10.1093/ofid/ofaa410>.
- Hrabák, J., et al., Oct. 2009. “International Clones of *Klebsiella pneumoniae* and *Escherichia coli* with Extended-Spectrum β -Lactamases in a Czech Hospital,”. *J. Clin. Microbiol.* 47 (10), 3353–3357. <https://doi.org/10.1128/JCM.00901-09>.
- Jamborova, I., et al., Jul. 2018. Extensive Genetic Commonality among Wildlife, Wastewater, Community, and Nosocomial Isolates of *Escherichia coli* Sequence Type 131 (H 30R1 and H 30Rx Subclones) That Carry bla CTX-M-27 or bla CTX-M-15. *p. AAC.00519-18 Antimicrob. Agents Chemother.* 62. <https://doi.org/10.1128/AAC.00519-18>.
- Johnson, T.J., et al., Aug. 2016. Separate F-Type Plasmids Have Shaped the Evolution of the H 30 Subclone of *Escherichia coli* Sequence Type 131. *mSphere* 1 (4). <https://doi.org/10.1128/msphere.00121-16>.
- Johnston, B.D., Johnson, J.R., 2017. Rapid and Specific Detection of *Escherichia coli* Sequence Type 131 (ST131) and its Key Subclones Using a Novel Single-tube Multiplex PCR Assay. *S599–S599 Open Forum Infect. Dis.* 4 (1). <https://doi.org/10.1093/ofid/ofx163.1573>.
- Kohlenberg, A., et al., Nov. 2024. Emergence of *Escherichia coli* ST131 carrying carbapenemase genes, European Union/European Economic Area, August 2012 to May 2024. *Eurosurveillance* 29 (47). <https://doi.org/10.2807/1560-7917.ES.2024.29.47.2400727>.

- Kolde Raivo, "pheatmap: Pretty Heatmaps. R package version 1.0.12," 2019.
- Kondratyeva, K., Salmon-Divon, M., Navon-Venezia, S., Jan. 2020. Meta-analysis of Pandemic *Escherichia coli* ST131 Plasmidome Proves Restricted Plasmid-clade Associations. *Sci. Rep.* 10 (1), 36. <https://doi.org/10.1038/s41598-019-56763-7>.
- Kutilova, I., et al., Feb. 2021. Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: Culture-based and metagenomic approaches. *Environ. Res.* 193, 110487. <https://doi.org/10.1016/j.envres.2020.110487>.
- Letunic, I., Bork, P., Apr. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkac268>.
- Li, D., et al., Dec. 2021. Genomic comparisons of *Escherichia coli* ST131 from Australia. *Micro Genom.* 7. <https://doi.org/10.1099/mgen.0.000721>.
- Ludden, C., et al., Apr. 2020. Genomic surveillance of *Escherichia coli* ST131 identifies local expansion and serial replacement of subclones. *Micro Genom.* 6 (4). <https://doi.org/10.1099/mgen.0.000352>.
- Marie-Hélène, N.-C., Xavier, B., Jean-Yves, M., Jul. 2014. "Escherichia coli ST131, an Intriguing Clonal Group." *Clin. Microbiol. Rev.* 27 (3), 543–574. <https://doi.org/10.1128/CMR.00125-13>.
- Mathers, A.J., Peirano, G., Pitout, J.D.D., Jul. 2015. The Role of Epidemic Resistance Plasmids and International High-Risk Clones in the Spread of Multidrug-Resistant Enterobacteriaceae. *Clin. Microbiol. Rev.* 28 (3), 565–591. <https://doi.org/10.1128/CMR.00116-14>.
- Merino, I., et al., Dec. 2020. Molecularly defined extraintestinal pathogenic *Escherichia coli* status predicts virulence in a murine sepsis model better than does virotype, individual virulence genes, or clonal subset among *E. coli* ST131 isolates. *Virulence* 11 (1), 327–336. <https://doi.org/10.1080/21505594.2020.1747799>.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., Jan. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* 32 (1), 268–274. <https://doi.org/10.1093/molbev/msu300>.
- Nicolas-Chanoine, M.-H., et al., 2017. The ST131 *Escherichia coli* H22 subclone from human intestinal microbiota: Comparison of genomic and phenotypic traits with those of the globally successful H30 subclone. *BMC Microbiol.* 17 (1), 71. <https://doi.org/10.1186/s12866-017-0984-8>.
- Novais, Á., et al., Jul. 2012. Characterization of Globally Spread *Escherichia coli* ST131 Isolates (1991 to 2010). *Antimicrob. Agents Chemother.* 56 (7), 3973–3976. <https://doi.org/10.1128/AAC.00475-12>.
- Omid, P., et al., Aug. 2021. Arrangements of Mobile Genetic Elements among Virotype E Subpopulation of *Escherichia coli* Sequence Type 131 Strains with High Antimicrobial Resistance and Virulence Gene Content, pp. 10.1128/msphere.00550-21 *msphere* 6 (4). <https://doi.org/10.1128/msphere.00550-21>.
- Papagiannitsis, C.C., et al., Feb. 2015. High prevalence of ST131 among CTX-M-producing *Escherichia coli* from community-acquired infections, in the Czech Republic. *Microb. Drug Resist.* 21 (1), 74–84. <https://doi.org/10.1089/mdr.2014.0070>.
- Paradis, E., Schliep, K., Feb. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35 (3), 526–528. <https://doi.org/10.1093/bioinformatics/bty633>.
- Paulshus, E., et al., Sep. 2019. Diversity and antibiotic resistance among *Escherichia coli* populations in hospital and community wastewater compared to wastewater at the receiving urban treatment plant. *Water Res.* 161, 232–241. <https://doi.org/10.1016/j.watres.2019.05.102>.
- Petty, N.K., et al., Apr. 2014. Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc. Natl. Acad. Sci.* 111 (15), 5694–5699. <https://doi.org/10.1073/pnas.1322678111>.
- Phan, M.D., et al., Apr. 2015. Molecular characterization of a multidrug resistance IncF plasmid from the globally disseminated *Escherichia coli* ST131 clone. *PLoS One* 10 (4). <https://doi.org/10.1371/journal.pone.0122369>.
- Pitout, J.D.D., Chen, L., Apr. 2023. The Significance of Epidemic Plasmids in the Success of Multidrug-Resistant Drug Pandemic Extraintestinal Pathogenic *Escherichia coli*. *Infect. Dis. Ther.* 12 (4), 1029–1041. <https://doi.org/10.1007/s40121-023-00791-4>.
- Pitout, J.D.D., Finn, T.J., 2020. The evolutionary puzzle of *Escherichia coli* ST131. *Infect., Genet. Evol.* 81, 104265. <https://doi.org/10.1016/j.meegid.2020.104265>.
- Price, L.B., et al., Dec. 2013. The Epidemic of Extended-Spectrum-β-Lactamase-Producing *Escherichia coli* ST131 Is Driven by a Single Highly Pathogenic Subclone, H30-Rx. *mBio* 4 (6). <https://doi.org/10.1128/mBio.00377-13>.
- Reid, C.J., et al., Feb. 2022. A role for ColV plasmids in the evolution of pathogenic *Escherichia coli* ST58. *Nat. Commun.* 13 (1), 683. <https://doi.org/10.1038/s41467-022-28342-4>.
- Reid, C.J., McKinnon, J., Djordjevic, S.P., Sep. 2019. Clonal ST131-H22 *Escherichia coli* strains from a healthy pig and a human urinary tract infection carry highly similar resistance and virulence plasmids. *Micro Genom.* 5 (9). <https://doi.org/10.1099/mgen.0.000295>.
- Revell, L.J., Jan. 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* 12, e16505. <https://doi.org/10.7717/peerj.16505>.
- Röderova, M., et al., Jan. 2017. Characteristics of Quinolone Resistance in *Escherichia coli* Isolates from Humans, Animals, and the Environment in the Czech Republic. *Front Microbiol.* 7. <https://doi.org/10.3389/fmicb.2016.02147>.
- Rogers, B.A., Sidjabat, H.E., Paterson, D.L., Jan. 2011. *Escherichia coli* O25b-ST131: a pandemic, multidrug-resistant, community-associated strain. *J. Antimicrob. Chemother.* 66 (1), 1–14. <https://doi.org/10.1093/jac/dkq415>.
- Sato, T., et al., Jun. 2024. Traces of pandemic fluoroquinolone-resistant *Escherichia coli* clone ST131 transmitted from human society to aquatic environments and wildlife in Japan. *One Health* 18, 100715. <https://doi.org/10.1016/j.onehlt.2024.100715>.
- Schürch, A.C., Arredondo-Alonso, S., Willems, R.J.L., Goering, R.V., Apr. 2018. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. *Clin. Microbiol. Infect.* 24 (4), 350–354. <https://doi.org/10.1016/j.cmi.2017.12.016>.
- Seemann, T., Jul. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30 (14), 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Stoesser, N., et al., May 2016. Evolutionary History of the Global Emergence of the *Escherichia coli* Epidemic Clone ST131. *mBio* 7 (2). <https://doi.org/10.1128/mBio.02162-15>.
- Tausova, D., et al., May 2012. *Escherichia coli* with extended-spectrum -lactamase and plasmid-mediated quinolone resistance genes in great cormorants and mallards in Central Europe. *J. Antimicrob. Chemother.* 67 (5), 1103–1107. <https://doi.org/10.1093/jac/dks017>.
- R.Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria.
- Tonkin-Hill, G., Lees, J.A., Bentley, S.D., Frost, S.D.W., Corander, J., Jun. 2019. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res.* 47 (11), 5539–5549. <https://doi.org/10.1093/nar/gkz361>.
- Tóth, K., et al., Apr. 2024. Genomic Epidemiology of C2/H30Rx and C1-M27 Subclades of *Escherichia coli* ST131 Isolates from Clinical Blood Samples in Hungary. *Antibiotics* 13 (4), 363. <https://doi.org/10.3390/antibiotics13040363>.
- Totsika, M., et al., Oct. 2011. Insights into a Multidrug Resistant *Escherichia coli* Pathogen of the Globally Disseminated ST131 Lineage: Genome Analysis and Virulence Mechanisms. *PLoS One* 6 (10), e26578. <https://doi.org/10.1371/journal.pone.0026578>.
- White, R.T., et al., Feb. 2024. Genomic epidemiology reveals geographical clustering of multidrug-resistant *Escherichia coli* ST131 associated with bacteraemia in Wales. *Nat. Commun.* 15 (1), 1371. <https://doi.org/10.1038/s41467-024-45608-1>.
- Whitmer, G.R., Moorthy, G., Arshad, M., 2019. The pandemic *Escherichia coli* sequence type 131 strain is acquired even in the absence of antibiotic exposure. *PLoS Pathog.* 15 (12). <https://doi.org/10.1371/journal.ppat.1008162>.
- Wickham, H., et al., Nov. 2019. Welcome to the Tidyverse. *J. Open Source Softw.* 4 (43), 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham H., François R., Henry L., Müller K., and Vaughan D., "dplyr: A Grammar of Data Manipulation. R package version 2.4.0," 2023.
- Wickham H., Hester J., and Bryan J., "readr: Read Rectangular Text Data. R package version 2.1.5," 2024.
- Wickham, Hadley, 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.
- Wyrsh, E.R., et al., Jun. 2022. Urban Wildlife Crisis: Australian Silver Gull Is a Bystander Host to Widespread Clinical Antibiotic Resistance. *mSystems* 7 (3). <https://doi.org/10.1128/mSystems.00158-22>.
- Yang, Q.-E., et al., Sep. 2015. IncF plasmid diversity in multi-drug resistant *Escherichia coli* strains from animals in China. *Front Microbiol.* 6. <https://doi.org/10.3389/fmicb.2015.00964>.
- Yu, G., Mar. 2020. Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc. Bioinforma.* 69 (1). <https://doi.org/10.1002/cpbi.96>.