Emergence and dissemination of antimicrobial resistance in Escherichia coli causing bloodstream infections in Norway in 2002–17

Gladstone, Rebecca A.; McNally, Alan; Pöntinen, Anna K.; Tonkin-Hill, Gerry; Lees, John A.; Skytén, Kusti; Cléon, François; Christensen, Martin O.K.; Haldorsen, Bjørg C.; Bye, Kristina K.; Gammelsrud, Karianne W.; Hjetland, Reidar; Kümmel, Angela; Larsen, Hege E.; Lindemann, Paul Christoffer; Löhr, Iren H.; Marvik, Åshild; Nilsen, Einar; Noer, Marie T.; Simonsen, Gunnar S.

DOI:
10.1016/S2666-5247(21)00031-8

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Gladstone, RA, McNally, A, Pöntinen, AK, Tonkin-Hill, G, Lees, JA, Skytén, K, Cléon, F, Christensen, MOK, Haldorsen, BC, Bye, KK, Gammelsrud, KW, Hjetland, R, Kümmel, A, Larsen, HE, Lindemann, PC, Löhr, IH, Marvik, Å, Nilsen, E, Noer, MT, Simonsen, GS, Steinbakk, M, Tofteland, S, Vattøy, M, Bentley, SD, Croucher, NJ, Parkhill, J, Johnsen, PJ, Samuelsen, Ø & Corander, J 2021, 'Emergence and dissemination of antimicrobial resistance in Escherichia coli causing bloodstream infections in Norway in 2002–17: a nationwide, longitudinal, microbial population genomic study', *The Lancet Microbe*, vol. 2, no. 7, pp. e331-e341. https://doi.org/10.1016/S2666-5247(21)00031-8

Emergence and dissemination of antimicrobial resistance in *Escherichia coli* causing bloodstream infections in Norway in 2002–17: a nationwide, longitudinal, microbial population genomic study

Rebecca A Gladstone, Alan McNally, Anna K Pöntinen, Gerry Tonkin-Hill, John A Lees, Kusti Skytén, François Cléon, Martin O K Christensen, Bjørg C Haldorsen, Kristina K Bye, Karianne W Gammelsrud, Reidar Hjetland, Angela Kümmel, Hege E Larsen, Paul Christoffer Lindemann, Iren H Löhr, Åshild Marvik, Einar Nilsen, Marie T Noer, Gunnar S Simonsen, Martin Steinbakk, Ståle Tofteland, Marit Vattøy, Stephen D Bentley, Nicholas J Croucher, Julian Parkhill, Pål J Johnsen, Ørjan Samuelsen, Jukka Corander**

Background The clonal diversity underpinning trends in multidrug resistant *Escherichia coli* causing bloodstream infections remains uncertain. We aimed to determine the contribution of individual clones to resistance over time, using large-scale genomics-based molecular epidemiology.

Methods This was a longitudinal, *E coli* population, genomic, cohort study that sampled isolates from 22 512 *E coli* bloodstream infections included in the Norwegian surveillance programme on resistant microbes (NORM) from 2002 to 2017. 15 of 22 laboratories were able to share their isolates, and the first 22·5% of isolates from each year were requested. We used whole genome sequencing to infer the population structure (PopPUNK), and we investigated the clade composition of the dominant multidrug resistant clonal complex (CC)131 using genetic markers previously reported for sequence type (ST)131, effective population size (BEAST), and presence of determinants of antimicrobial resistance (ARIBA, PointFinder, and ResFinder databases) over time. We compared these features between the 2002–10 and 2011–17 time periods. We also compared our results with those of a longitudinal study from the UK done between 2001 and 2011.

Findings Of the 3500 isolates requested from the participating laboratories, 3397 (97·1%) were received, of which 3254 (95·8%) were successfully sequenced and included in the analysis. A significant increase in the number of multidrug resistant CC131 isolates from 71 (5·6%) of 1277 in 2002–10 to 207 (10·5%) of 1977 in 2011–17 (p<0·0001), was the largest clonal expansion. CC131 was the most common clone in extended-spectrum β-lactamase (ESBL)-positive isolates (75 [58·6%] of 128) and fluoroquinolone non-susceptible isolates (148 [39·2%] of 378). Within CC131, clade A increased in prevalence from 2002, whereas the global multidrug resistant clade C2 was not observed until 2007. Multiple de-novo acquisitions of both bla_{CTX-M} ESBL-encoding genes in clades A and C1 and gain of phenotypic fluoroquinolone non-susceptibility across the clade A phylogeny were observed. We estimated that exponential increases in the effective population sizes of clades A, C1, and C2 occurred in the mid-2000s, and in clade B a decade earlier. The rate of increase in the estimated effective population size of clade A (N_e=3147) was nearly ten-times that of C2 (N_e=345), with clade A over-represented in Norwegian CC131 isolates (75 [27·0%] of 278) compared with the UK study (8 [5·4%] of 147 isolates).

Interpretation The early and sustained establishment of predominantly antimicrobial susceptible CC131 clade A isolates, relative to multidrug resistant clade C2 isolates, suggests that resistance is not necessary for clonal success. However, even in the low antibiotic use setting of Norway, resistance to important antimicrobial classes has rapidly been selected for in CC131 clade A isolates. This study shows the importance of genomic surveillance in uncovering the complex ecology underlying multidrug resistance dissemination and competition, which have implications for the design of strategies and interventions to control the spread of high-risk multidrug resistant clones.

Funding Trond Mohn Foundation, European Research Council, Marie Skłodowska-Curie Actions, and the Wellcome Trust.

Lancet Microbe 2021; 2: e331–41

Published Online May 10, 2021 https://doi.org/10.1016/S2666-5247(21)00031-8

This online publication has been corrected. The corrected version first appeared at thelancet.com/microbe on June 29, 2021

*Authors contributed equally

Department of Biostatistics (R A Gladstone PhD, A K Pöntinen PhD, K Skytén MSc, Prof J Corander PhD), **Institute of Clinical Medicine, Faculty of Medicine** (M T Noer MSc), **University of Oslo, Oslo, Norway; Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK** (Prof A McNally PhD); **Parasites and Microbes, Wellcome Sanger Institute, Cambridge, UK** (G Tonkin-Hill MSc, Prof S D Bentley PhD, Prof J Corander); **Faculty of Medicine, School of Public Health, Imperial College, London, UK** (J A Lees PhD, N J Croucher PhD); **Department of Pharmacy** (F Cléon PhD, Prof P J Johnsen PhD, Prof Ø Samuelsen PhD), **Department of Medical Biology** (Prof G S Simonsen PhD), **Faculty of Health Sciences UiT The Arctic University of Norway, Tromsø, Norway; Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway** (M O K Christensen MSc, B C Haldorsen MSc, Prof G S Simonsen, Prof Ø Samuelsen); **Laboratory of Microbiology, Department**

Introduction

Escherichia coli is a commensal of the gastrointestinal tract and a leading cause of bloodstream infections worldwide, which is associated with considerable morbidity and mortality.[1] *E coli* associated with bloodstream infections are a subset of extraintestinal pathogenic *E coli* (ExPEC), with bloodstream infections occurring as a result of underlying urinary tract

of Medical Biochemistry (K K Bye BSc), Oslo University Hospital Radiumhospitalet, Oslo, Norway; Institute of Medical Microbiology, Oslo University Hospital Rikshospitalet, Oslo, Norway (M T Noer); Department of Microbiology, Division of Laboratory Medicine, Oslo University Hospital Ullevål, Oslo, Norway (K W Gammelsrud PhD); Department of Microbiology, Førde General Hospital, Førde Health Trust, Førde, Norway (R Hjetland PhD); Department of Laboratory Medicine, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway (A Kümmel MD); Department of Microbiology, Nordland Hospital, Bodø, Norway (H E Larson BSc); Department of Microbiology, Haukeland University Hospital, Bergen, Norway (P C Lindemann MD); Department of Medical Microbiology, Stavanger University Hospital, Stavanger, Norway (I H Löhr MD); Department of Microbiology, Vestfold Hospital, Tønsberg, Norway (Å Marvik PhD); Department of Microbiology, Moere and Romsdal Hospital Trust, Molde, Norway (E Nilsen MD); Norwegian Institute of Public Health, Oslo, Norway (Prof G S Simonsen); Centre for Laboratory Medicine, Sections for Microbiology, Østfold Hospital, Kalnes, Norway (M Steinbakk MD); Department of Medical Microbiology, Sørlandet Hospital, Kristiansand, Norway (S Tofteland PhD); Department of Microbiology, Akershus University Hospital, Lørenskog, Norway (M Vattøy BSc); Department of Veterinary Medicine, University of Cambridge, Cambridge, UK (Prof J Parkhill PhD)

Correspondence to:
Dr Rebecca A Gladstone, Department of Biostatistics, University of Oslo, P 0317 Oslo, Norway
r.a.gladstone@medisin.uio.no

Research in context

Evidence before this study

Following initial reports in the early 2000s, numerous surveillance and epidemiological studies have reported *Escherichia coli* clonal complex (CC)131 as the dominant cause of *E coli* urinary tract and bloodstream infections worldwide. We searched PubMed from inception of the database to May 22, 2020, with the search terms "coli" AND "ST131" AND "infection" AND "urinary" OR "blood". Searches were restricted to primary research articles published in English. This returned a total of 957 publications describing the isolation and relative frequency of CC131 in urinary tract infections and bloodstream infections. Of these, only one was designed as a longitudinal survey of isolates on a UK collection of bloodstream infections *E coli* isolates in 2001–11 from the British Society for Antimicrobial Chemotherapy. By doing a densely sampled longitudinal study, the team behind the UK collection were able to conclude that although CC131 increased in frequency in the early 2000s, this increase was not as large as suggested by previous single-site, single-timepoint studies. The study also suggested that multidrug resistance alone did not explain the increase in CC131 prevalence, a finding supported by a subsequent evolutionary genomics study in 2019. Thus, a number of important questions remain regarding the epidemiology of *E coli* from bloodstream infections, including whether the results of the UK study are representative of the wider geographical picture and what has happened to the *E coli* population from bloodstream infections after 2011.

Added value of this study

Our study used a longitudinal sampling framework over an extended period of 16 years that provides a clear picture of the underlying epidemiology of *E coli* from bloodstream infections. By doing this in another country, we had the opportunity to compare the findings with the longitudinal UK study and determine whether the findings in one setting were generalisable, even when antimicrobial usage and resistance differs. These data also allowed for comparisons of the growth and expansion of clones that cause bloodstream infections beyond the raw prevalence numbers available at any one time.

Implications of all the available evidence

A full picture of the emergence of *E coli* as a major cause of bloodstream infections is required, driven by longitudinal understanding of the epidemiology and multidrug resistance of these organisms. Our data suggests that CC131 clade A is an increasingly common cause of multidrug resistance bloodstream infections, which is a novel finding not reported in previous studies and merits surveillance for this clone worldwide. This study shows that the findings of individual, longitudinal surveys of *E coli* epidemiology are not generalisable across countries, not even between resource-similar nations within Europe. The observation of both emerging multidrug resistant clones and the acquisition of multidrug resistance in established clones, has implications for our efforts to control multidrug resistance in this pathogen.

infection, gastrointestinal colonisation, or hepatobiliary infections.[2] Although most ExPEC belong to phylogroup B2 they are not a monophyletic group; ExPEC are a diverse collection of clones with the capacity to invade and cause disease through the acquisition of virulence factors.[3] A small number of globally dispersed ExPEC clones are responsible for most infections,[4] and increases in the incidence of *E coli*-associated bloodstream infections have been reported in multiple countries.[5–7] Increasing incidence is exacerbated by an increasing proportion of multidrug resistant ExPEC infections.[5–7] One study estimated that *E coli* resistant to third generation cephalosporins caused 25–43 million cases of bloodstream infections and other serious infections globally in 2014.[8]

In 2014, a single ExPEC clone—clonal complex (CC)131—was reported to predominate globally.[9] A retrospective genomic study of bloodstream infections in the UK showed that CC131 had emerged quickly around 2003.[10] Whole genome sequencing has allowed the delineation of a number of phylogenetic subgroups—known as clades—within CC131; each clade has different associations with antimicrobial resistance.[11–13] Clades C1 (alternatively referred to as H30-R) and C2 (H30-Rx) are characterised by fluoroquinolone resistance. Clade C2 is also associated with extended-spectrum β-lactamases (ESBL) of the CTX-M class conferring resistance to third generation cephalosporins.[11–13] Clades C1 and C2 were reported to be the leading cause of multidrug resistant *E coli* infections in the USA in 2011–12.[6] Additional CC131 clades—A, B, B0, and C0—have been reported as minor clades with less resistance.[11–13]

Much of the literature proposes that antimicrobial resistance might have driven the emergence and success of CC131. However, there are other successful *E coli* bloodstream infections clones—including CC73 and CC95—that have less extensive antimicrobial resistance.[10] Evidence suggests that success, through competition between *E coli* strains, is a product of the relative frequency of the full complement of accessory genes, not just antimicrobial resistance, termed negative frequency-dependent selection.[10,14]

E coli bloodstream infections have been increasing in Norway, with the prevalence of ESBL-producing and fluoroquinolone non-susceptible *E coli* causing bloodstream infections also increasing.[5] We used whole genome sequencing to analyse the clonal diversity of *E coli* causing bloodstream infections in Norway over 16 years. This large nationwide longitudinal genomic collection allowed us to observe the emergence of clones and determine their contribution to rising bloodstream infections incidence and antimicrobial resistance prevalence.

Methods

Study design and bacterial isolates

In this nationwide, longitudinal, microbial population genomic study, a dataset of all 22 512 *E coli* isolates collected from 2002 to 2017, as part of the Norwegian surveillance programme on resistant microbes (NORM), was collated.[5] Sample identifier, laboratory, year, sample collection date, and antimicrobial susceptibility profile were reported for each isolate. We selected the first 22·5% of isolates annually from 15 of the 22 laboratories participating in NORM to fill our capacity to sequence 3500 isolates. (appendix 1 pp 1, 6). The collection of data within NORM follows a standard protocol defined in the yearly surveillance reports (appendix 1 p 1).[5]

Procedures

Isolates were plated on MacConkey agar no 3 (Oxoid, Thermo Fisher Scientific, Waltham, MA, USA); one colony was subsequently used to inoculate 1·6 mL Luria Bertani broth (Becton Dickinson, NJ, USA) and incubated overnight at 37°C. Genomic DNA was extracted by QIAGEN (Hilden, Germany) with the DNeasy 96 Blood and Tissue kit (QIAGEN, Hilden, Germany). Samples were sequenced at the Wellcome Sanger Institute (Hinxton, UK) on the Illumina HiSeq platform (San Diego, CA, USA) using the NEB Ultra II custom kit (Emsworth, UK) with 392-plexing and a 150 base-pair read length. Samples failed quality control if there was insufficient DNA to be sequenced, depth of coverage was less than 20, or there was evidence of mixed strain or species contamination. Sequence data are available on the European Nucleotide Archive (appendix 2). Data in NORM are disidentified. Ethics approval was not required after evaluation by the Regional Ethical Committee (REC North 93528).

Genome sequences were assembled and annotated using default parameters of a published pipeline,[15,16] and the pangenome was defined using Panaroo (version 1.0.2).[17] Phylogenetic group and multilocus sequence type were determined using published tools (appendix pp 2–3).[18,19] PopPUNK (version 2.0.2) was used to cluster isolates with shared ancestry into clones.[20] We defined a CC as the most prevalent sequence type (ST) within a PopPUNK grouping. ARIBA (version 2.12.2) was used with ResFinder and PointFinder (*E coli*) databases[21,22] to detect bacterial resistance determinants, including $bla_{CTX\text{-}M}$ genes and fluoroquinolone resistance associated mutations in *gyrA*, *parC*, or *parE*.[21–23] $bla_{CTX\text{-}M}$ genes served as a proxy for ESBL-production and mutations in genes associated with ciprofloxacin efflux were screened for in the Panaroo gene alignments.[24–26]

PopPUNK4 (CC131) isolates were mapped to EC958 (GenBank HG941718.1) and recombination was detected and removed using GUBBINS (version 2.4.0). A recombination-free phylogeny was produced with RAxML (version 8.2.8).[27,28] CC131 clades were assigned using clade-specific single nucleotide polymorphism and *fimH* alleles, and then corrected for the phylogenetic distribution of clades.[11–13] K-mers associated with recombination hotspots were excluded before generating a reference-free alignment with SKA (version 1.0) for each CC131 clade.[29] BEAST (version 1.10.4) was subsequently used with the generalised time reversible substitution and gamma site heterogeneity models for 100 000 000 generations, with 10% discarded as burn-in, sampled every 10 000 states.[30] External collections of *E coli* CC131 genomes used by McNally and colleagues[14] and a longitudinal collection of bloodstream infections[10] were used to compare with Norwegian isolates, the fastq data was processed using the methods we described here. Additional details on isolate culturing; DNA extraction, sequencing, and annotation; and genome sequence analyses are described in appendix 1 (pp 1–4).

See Online for appendix 1

Statistical analysis

To determine if differences in proportions between two groups were significant two-sided Fisher's exact test was used when any counts were less than five, otherwise χ^2 test was used. Both tests had a $p<0·05$ threshold. We split the collection into two periods (2002–10 and 2011–17) to assess differences in proportions to reduce any bias contained in any one year. We chose to split between 2010 and 2011 because this was approximately the half way point of the sample collection period and because we observed that many key changes occurred after 2010. Using these two periods we compared the prevalence of popPUNK4-CC131 as a proportion of the collection, each CC131 clade as a proportion of the collection, and of each CC131 clade within CC131. We also compared the overall prevalence of $bla_{CTX\text{-}M}$-positive and ciprofloxacin non-susceptible CC131 isolates (on the basis of NORM data) and compared the isolates that belonged to each clade of CC131 as proportion of the collection, as a proportion of CC131, and as a proportion of each CC131 clade. Additionally, we compared the prevalence of CC131, and the prevalence of $bla_{CTX\text{-}M}$-positive and ciprofloxacin non-susceptible isolates between different health regions in Norway. Finally, we compared the prevalence of CC131, each CC131 clade, $bla_{CTX\text{-}M}$-positive isolates, and isolates in which *gyrA* or *parC* mutations associated with reduced ciprofloxacin susceptibility were detected, between our study and a UK study.[10] The BEAST Bayesian skyline model, which uses phylogenetic history to provide estimates of the median effective population size over time with a 95% highest posterior density, was used to identify the main exponential increase in population size and the slope representing the rate of increase in each clade (appendix p 4). To assess the relative invasiveness of CC131 clades we used the ratio of effective population size and incidence in bloodstream infections in 2017. The incidence of each clade was estimated from the isolate counts adjusted for the proportion of total *E coli* bloodstream infections in 2017 that was sampled in this study, reported per million

See Online for appendix 2

population. Statistical analyses were done with R (version 3.6.3).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The 15 laboratories contributing to NORM provided 15 552 samples collected from 2002 to 2017, of which we selected 3500 (22·5%) for inclusion. Of the 3500 isolates, 3397 (97·1%) were received and sequenced. 3254 (95·8%) passed sequencing quality control criteria and had a mean sequencing depth of 52·1 (SD 9·9). Excluded samples had no effect on the interpretation of the results (appendix 1 p 5). 1277 (39·2%) of 3254 samples were from the first half of the study (2002–10) and 1977 (60·8%) samples were from the second half of the study (2011–17).

2180 (67·0%) of 3254 samples belonged to phylogroup B2. The collection was delineated into 136 clonal groups by PopPUNK (appendix 1 p 7). PopPUNK1 (CC73), PopPUNK2 (CC95), PopPUNK3 (CC69), PopPUNK4 (CC131), and PopPUNK5 (CC14) were the most common clones, accounting for 1761 (54·1%) of 3254 samples. CC131 showed the largest proportional gain in the collection between 2002 and 2017, occurring gradually from 2002 to its peak prevalence in 2013 (figure 1; appendix 1 p 8), with a significant increase between 2002–10 (71 [5·6%] of 1277 samples) and 2011–17 (207 [10·5%] of 1977 samples; p<0·0001; figure 1A). CC131 clades A (p=0·0043), C1 (p=0·0002), and C2 (p=0·0009) increased as a proportion of the collection between 2002 and 2010 and 2011 and 2017, while clade B remained stable (p=0·55; figure 1B; appendix 1 p 18). Clades B0 and C0 combined were present in eight (2·9%) of 278 CC131 samples (appendix 1 p 9). Within CC131, we only detected a change in proportion for clade B, which significantly decreased as a proportion of CC131 between 2002 and 2010 (32 [45·1%] of 71 samples) and 2011–17 (42 [20·3%] of 207 samples; p=0·0001; appendix 1 p 18).

We considered CTX-M-encoding genes to be markers for ESBL-production and detected nine $bla_{\text{CTX-M}}$ variants in 128 (3·9%) of 3254 isolates (appendix 1 p 10). We did not detect a significant difference in $bla_{\text{CTX-M}}$ gene prevalence between 2002 (one [1·2%] of 83 isolates) and the peak prevalence in 2016 (21 [7·0%] of 300; p=0·058). CC131 was the single largest contributor of $bla_{\text{CTX-M}}$-positive isolates in the collection (75 [58·6%] of 128 isolates; figure 2A). The remaining 53 (41·4%) CTX-M positive isolates belonged to 18 other PopPUNK clones. The number of clones with $bla_{\text{CTX-M}}$-positive isolates in 2017 was not significantly different to 2002 (p=0·19). Clades C2, C1, and A accounted for all CC131 $bla_{\text{CTX-M}}$-positive isolates and for more than half (71 [63·4%]) of all 112 $bla_{\text{CTX-M}}$-positive isolates during 2011–17 (C2: 33 [29·5%]; C1: 27 [24·1%]; and A: 11 [9·8%]; figure 2B). $bla_{\text{CTX-M}}$-positive isolates increased significantly overall in the 2002–10 and 2011–17 periods, as did $bla_{\text{CTX-M}}$-positive isolates that belonged to CC131 and to the CC131 clades C1, C2, and A as a proportion of the collection (appendix 1 p 19). Clade C2 had the highest proportion of $bla_{\text{CTX-M}}$-positive isolates (33 [76·7%] of 43) in 2011–17.

$bla_{\text{CTX-M-15}}$ was the most common $bla_{\text{CTX-M}}$ type (appendix p 10). Within CC131, 36 (70·6%) of 51 clade C2 isolates were $bla_{\text{CTX-M-15}}$-positive, and one (2·0%) of 51 clade C2 isolates was $bla_{\text{CTX-M-89}}$-positive. The most common $bla_{\text{CTX-M}}$ type in clade C1 was $bla_{\text{CTX-M-27}}$ (15 [21·4%] of 70 isolates). $bla_{\text{CTX-M-15}}$ (7 [9·3%] of 75 isolates) was the most common in clade A (appendix 1 p 20). All 74 clade B isolates were $bla_{\text{CTX-M}}$-negative. De-novo $bla_{\text{CTX-M}}$ acquisitions were observed in clades A and C1 (appendix 1 pp 11–13). Within clade C1 we estimated that there were nine phylogenetically independent acquisitions of $bla_{\text{CTX-M}}$ genes contributing to the 27 $bla_{\text{CTX-M}}$-positive isolates. A single acquisition of $bla_{\text{CTX-M-27}}$ accounted for 14 (51·9%) of the 27 isolates. However, most acquisitions (five [55·6%] of nine) were detected in a single isolate. Within clade A we estimated that around seven de-novo acquisitions of CTX-M-encoding genes accounted for most $bla_{\text{CTX-M}}$-positive isolates, with most acquisitions (four [57·1%] of seven) detected in only one isolate and all acquisitions detected in a maximum of two isolates.

Ciprofloxacin non-susceptibility (determined on the basis of susceptibility profiles from NORM) was observed in 378 (11·6%) of 3254 isolates, increasing from three (3·6%) of 83 isolates in 2002 to a peak of 67 (18·9%) of 355 isolates in 2017. There was a significant increase from 77 (6·0%) of 1277 isolates during the 2002–10 period to 301 (15·2%) of 1977 isolates during the 2011–17 (p<0·0001). Only 94 (24·9%) of 378 ciprofloxacin non-susceptible isolates were also $bla_{\text{CTX-M}}$-positive, whereas most (94 [73·4%] of 128 isolates) of $bla_{\text{CTX-M}}$-positive isolates were also ciprofloxacin non-susceptible. CC131 made the single largest contribution to ciprofloxacin non-susceptibility in the collection, contributing 148 (39·2%) of 378 isolates overall and 25 (37·3%) of 67 isolates in 2017 alone. CC131, CC14, and CC69 combined accounted for 41 (61·2%) of 67 ciprofloxacin non-susceptible isolates in 2017. The remaining 26 (38·8%) belonged to 34 other PopPUNK clones (figure 3A). Of the 301 ciprofloxacin non-susceptible isolates reported between 2011 and 2017, 53 (17·6%) were CC131 clade C1, 42 (14·0%) were clade C2 contributed, and 25 (8·3%) were clade A. The proportion of isolates with ciprofloxacin non-susceptibility within CC131 significantly increased (p=0·0002) between 2002–10 and 2011–17 (figure 3B; appendix 1 pp 14, 21). Clades C1 and C2 were near uniformly ciprofloxacin resistant (114 [94·2%] of 121 isolates) and all had Ser83Leu and Asp87Asn substitutions in GyrA and Ser80Ile and Glu84Val substitutions in ParC. The non-susceptible phenotypes in clade A were represented by multiple independent phylogenetic clusters and were not well

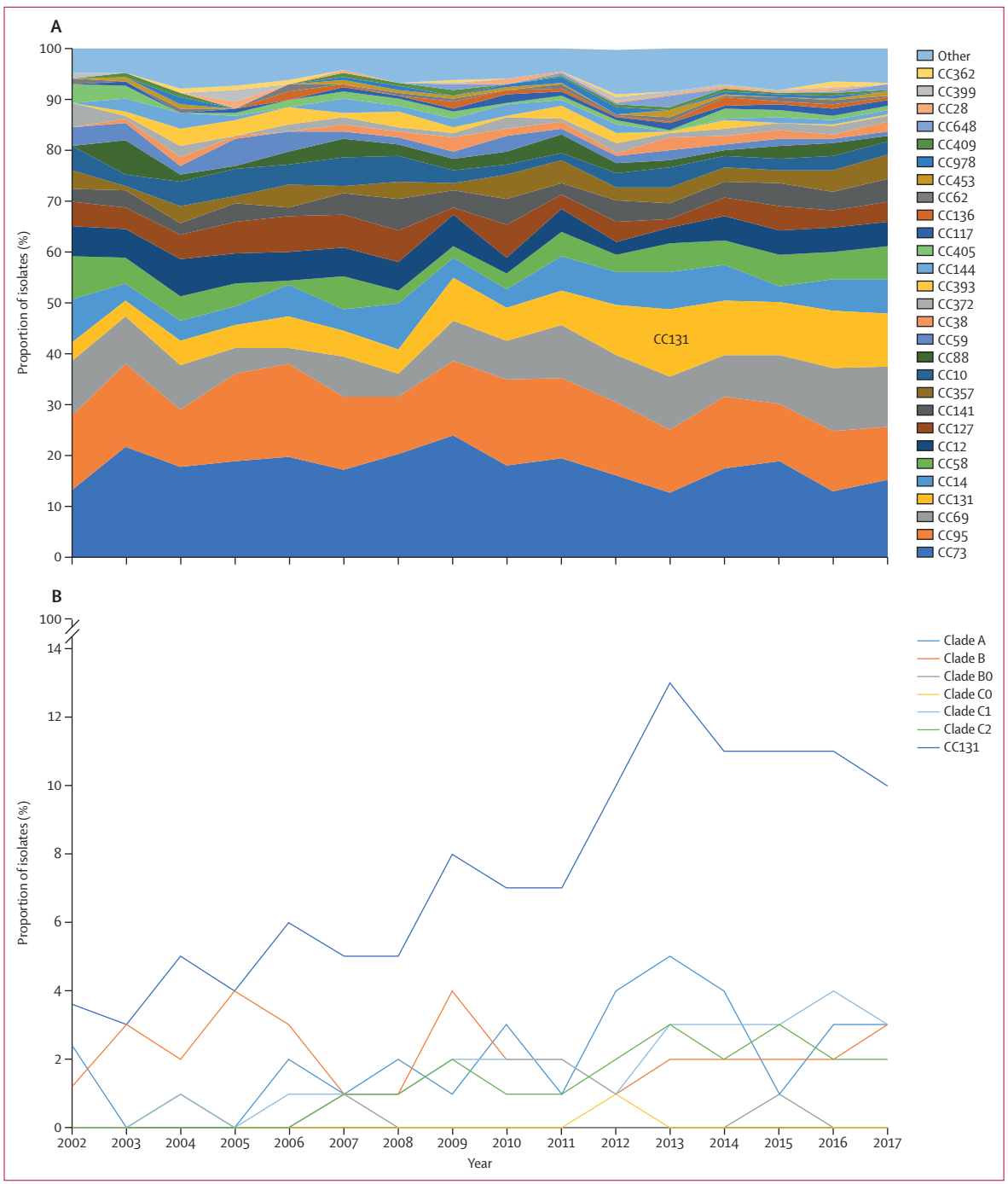

Figure 1: **Proportion of isolates by clone and by CC131 clades over time**
(A) Clones as a proportion of all 3254 isolates by year. CC131 had the largest gain of all lineages as proportion of the collection between 2002 and 2017. (B) CC131 and its clades as a proportion of all 3254 isolates by year. Clades C1, C2, and A all contributed to the observed increase of CC131 as a proportion of collection. CC=clonal complex.

explained by detected mutations in *gyrA*, *parC*, or *parE* listed in the PointFinder (*E coli*) resistance database (appendix 1 p 15). 23 (82·1%) of the 28 ciprofloxacin non-susceptible isolates in clade A had resistance mutations in *gyrA* but not *parC*, and most clade A isolates (71 [94·7%] of 75) had the ParE Ile529Leu substitution regardless of phenotype. Non-susceptible isolates in clade A did not have known mutations in efflux-associated genes, pump-encoding genes, or plasmid-mediated quinolone resistance genes.

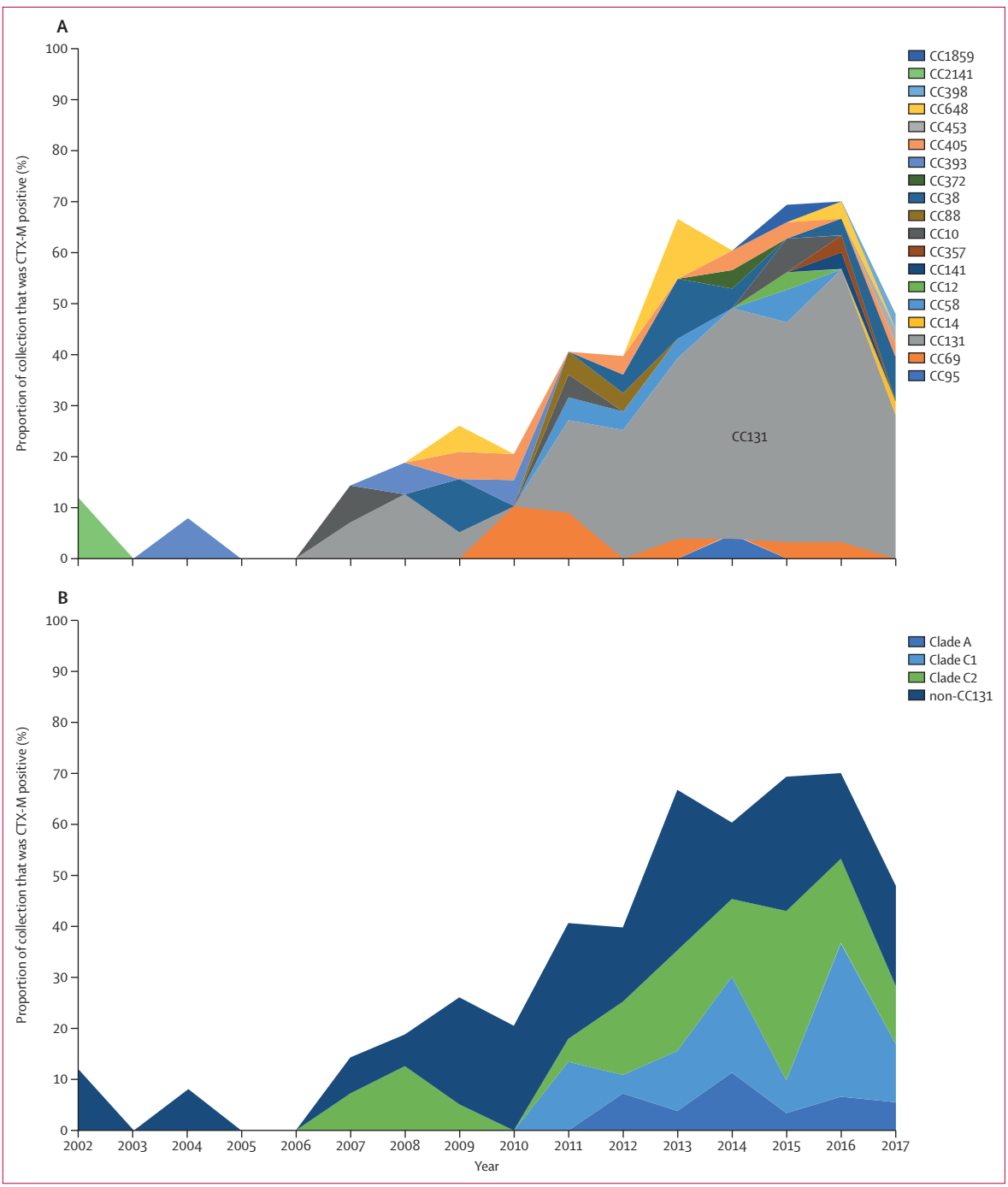

Figure 2: **Proportion *bla*$_{CTX-M}$-positive isolates by clone and by CC131 clades over time**
(A) The 128 *bla*$_{CTX-M}$-positive isolates by clone as a proportion of all 3254 isolates by year. (B) *bla*$_{CTX-M}$-positive isolates in CC131 (n=75) by clade as a proportion of the total collection (n=3254), over time. There were no *bla*$_{CTX-M}$-positive CC131 isolates in clades B0, B, or C0. CC=clonal complex.

We estimated an exponential increase in the effective population sizes of clades A, C1, and C2 in the mid-2000s, whereas the increase in clade B occurred more than a decade earlier, around 1990 (figure 4). The rate of expansion of the effective population size over time (per year) in the exponential phase was estimated to be nearly ten-times higher for clade A (N_e=3147) than for clade C2 (N_e=345) and nearly 25-times higher than for C1 (N_e=131; figure 4; table).

These four clades (A, B, C1, and C2) each accounted for about 2–3% of the collection in 2017 and were estimated (adjusting for sampling) to have similar bloodstream infections incidence per million population in Norway

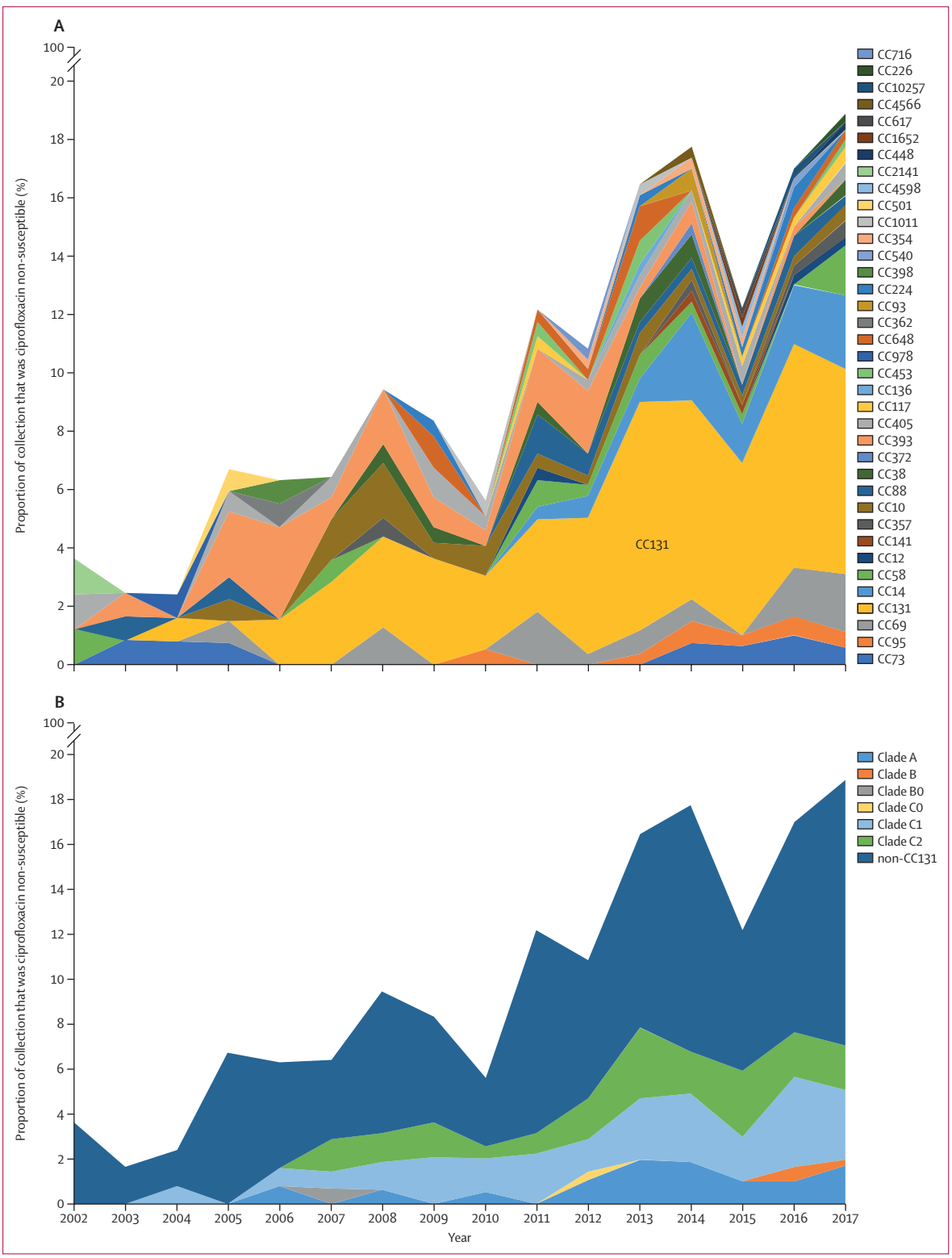

***Figure 3*: Proportion of isolates that were ciprofloxacin non-susceptible by clone and by CC131 clades over time**
(A) 378 ciprofloxacin non-susceptible isolates by clone as a proportion of all 3254 isolates by year. CC131 made the single largest contribution of any clone to ciprofloxacin non-susceptibility. (B) 148 ciprofloxacin non-susceptible isolates in CC131 by clade as a proportion of all 3254 isolates by year. CC=clonal complex.

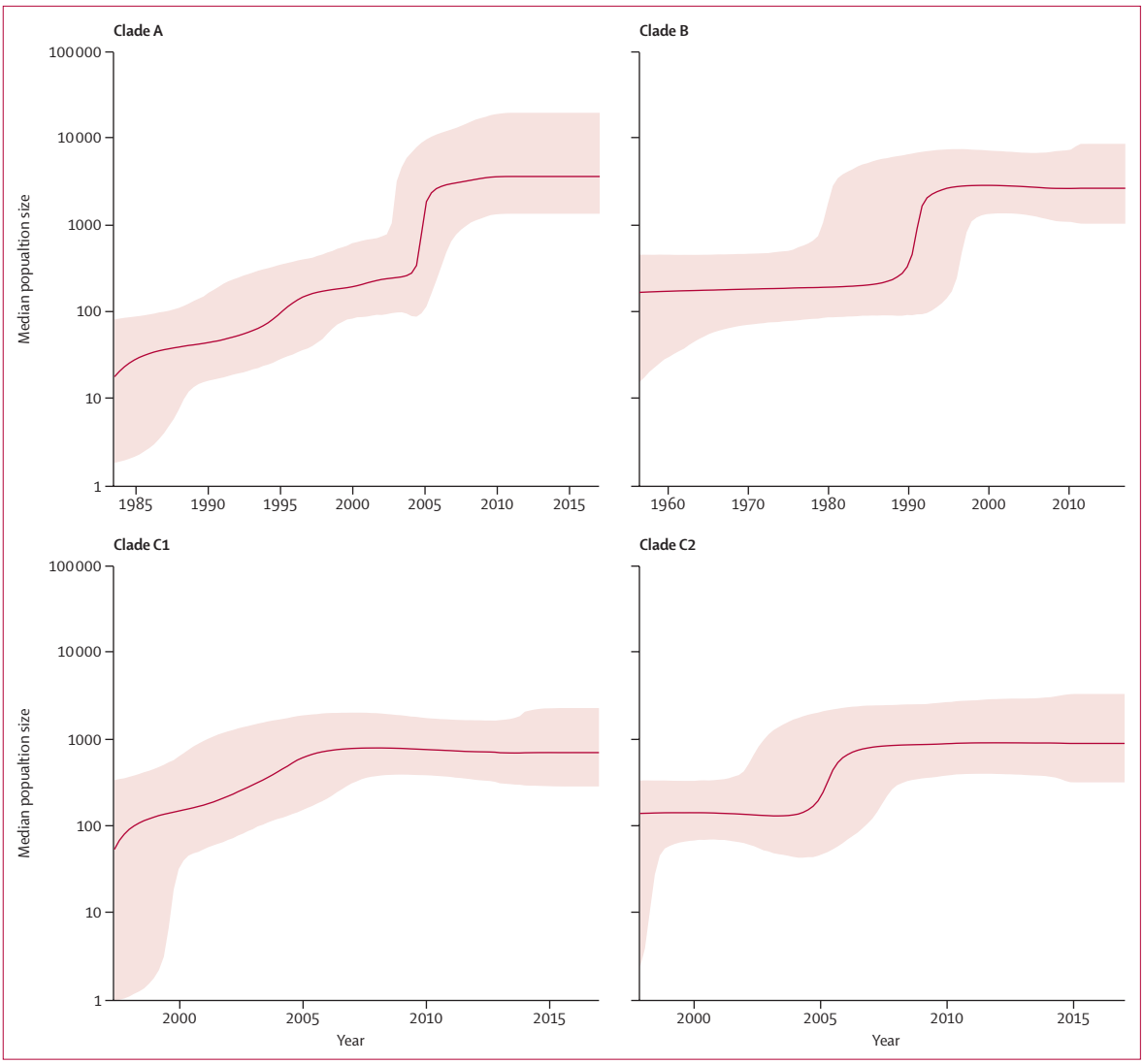

Figure 4: **Bayesian skyline plots of estimated median effective population size of clonal complex 131 clades A, B, C1, and C2 over time**
Shaded areas indicate 95% highest posterior density intervals.

in 2017 (table; appendix 1 pp 4, 22). The ratio of each clade's estimated population size (used as a proxy for the carriage population) to incidence of disease (bloodstream infections) in the human population was used as a measure of invasiveness relative to the other clades. For clades C1 and C2 we estimated smaller population sizes and higher population size to incidence ratios, whereas clade A had the largest population size and lowest ratio. The time to the most recent common ancestors of our Norwegian isolates in CC131 clades C1 and C2 predate the study period by more than a decade, whereas clades A and B are estimated to be older (table).

There was no discernible spatiotemporal spread or phylogenetic structure within Norway during the expansion of CC131 (appendix 1 pp 16, 23). In 2002, clade A was already observed in blood cultures in the north and west health regions. C2 was the most recently observed clade in bloodstream infections: first reported in 2007. CC131 (p=0·031) and $bla_{CTX\text{-}M}$-positive isolates (p=0·032) were significantly under-represented in the central health region compared with the rest of the collection. Ciprofloxacin resistance was under-represented in the north (p=0·041) and over-represented in the southeast (p=0·015; appendix 1 p 24).

Prevalence of CC131 in our collection was compared with CC131 isolates from another longitudinal collection from the UK, representing isolates from 2002–11.[10] Compared with the UK, a significantly larger proportion of Norwegian CC131 isolates belonged to clade A (UK 8 [5·4%] of 147 isolates *vs* Norway 75 [27·0%] of 278 isolates; p<0·0001) and clade C1 (UK 12 [8·2%] of 147 *vs* Norway 70 [25·2%] of 278; p<0·0001; appendix 1 pp 25–27). The

	Population size in 2017 (95% HPDI)	Exponential time period	N_e (95% HPDI), per year*	Estimated incidence per million population in 2017† (95% CI)	Population size to incidence ratio in 2017	tMRCA (95% HPDI)
Clade A	3667 (1365–19872)	2004·6–2005·3	3147 (2909–3385)	20·58 (20·35–20·82)	0·006	1977 (1968–1983)
Clade B	2711 (1058–8812)	1989·8–1993·9	650 (608–692)	22·87 (22·63–23·10)	0·008	1926 (1884–1965)
Clade C1	730 (297–2382)	2002·8–2005·8	131 (126–137)	25·15 (24·92–25·39)	0·028	1993 (1988–1997)
Clade C2	883 (314–3284)	2004·7–2006·7	345 (326–364)	16·01 (15·77–16·24)	0·022	1992 (1985–1997)

N_e=growth rate. HPDI=highest posterior density interval. tMRCA=time to most recent common ancestor. *Total cases were estimated for each CC131 clade by adjusting the number of genomes of a clade by the proportion of total 2017 cases that were sequenced, before calculating incidence per million population in Norway in 2017. †During exponential time period.

Table: **Median effective population size of clonal complex 131 clades**

opposite was true for clade C2 (UK 89 [60·5%] of 147 *vs* Norway 51 [18·3%] of 278; p<0.0001). As such, resistant isolates within these clades made different contributions to resistance between the two countries. Clade C2 made a significantly smaller contribution (p<0·0001), and clades C1 and A made significantly larger contributions to both CC131 $bla_{CTX\text{-}M}$-positive isolates (C1 p<0·0001 and A p=0·0028), and CC131 isolates with *gyrA* and *parC* mutations (C1 p<0·0001 and A p<0·0001) in Norway than in the UK (appendix 1 pp 26–27).

Finally, we contextualised our Norwegian isolates within a global genomic collection.[14] Aside from the over-representation of Norwegian isolates in clade A and under-representation in clade C2, Norwegian isolates were phylogenetically spread across the sampled genetic diversity from other locations within each of the CC131 clades (appendix 1 p 17).

Discussion

Cases of *E coli* bloodstream infections have been increasing in Norway, the UK, and elsewhere.[5,7] Of greater concern is the increasing proportion of bloodstream infections caused by multidrug resistant isolates. ESBL prevalence in *E coli* bloodstream infections in Norway increased from 0·3% in 2002 to 6·6% by 2017.[5] Fluoroquinolone resistance in Norway has also continued to increase from 3·3% in 2002 to 18·0% in *E coli* causing bloodstream infections in 2017.[5] We observed similar prevalence and general increasing trends. A clear correlation between the total usage of fluoroquinolones and non-susceptibility to these drugs has been reported in Norway, where antimicrobial usage in both animals and people is generally low and tightly regulated.[5]

Our data show the emergence of CC131 as the dominant contributor to the prevalence of ESBLs and fluoroquinolone resistance in *E coli* causing bloodstream infections in Norway. The emergence of CC131 in Norway appears to have been a gradual process spanning 2002–13 compared with that reported for the UK where the expansion occurred quickly between 2002–04.[10] Additionally, we show that the prevalence of the individual CC131 clades are different between the two countries, with a higher prevalence and resistance for clade A and C1 and lower prevalence and resistance in C2 in Norway. Despite these differences we detected the same predominant CC131 $bla_{CTX\text{-}M}$ types reported by others: $bla_{CTX\text{-}M\text{-}15}$ in clade C2, and $bla_{CTX\text{-}M\text{-}27}$ and $bla_{CTX\text{-}M\text{-}14}$ in clade C1.[10,31] Moreover, the CC131 diversity we observed overlapped with other datasets, and our estimated time to most recent common ancestor for CC131 clades are in line with previous estimates.[11–13,31,32]

Of note, in Norway, CC131 clades have different antimicrobial resistance prevalence, estimated population sizes, growth rates, and potentially invasiveness. In-vivo quantification of clade virulence combined with genomics could reveal underlying variation driving any differences in invasiveness. Evidence of varying evolutionary pressures also warrants elaboration of the negative frequency-dependent selection model for ExPEC population evolution.[14] Although responsible for less resistance than the multidrug resistant clade C2, increasing resistance facilitated by multiple de-novo acquisition events in other clades of CC131 is of clinical concern. Clade A, which was the most frequently observed clade during the study period, and clade B are far less resistant than clade C2 suggesting drug resistance is not essential for their success. Furthermore, we show resistance being acquired multiple times sporadically within clades A and C1. The clades expanded without preferential expasion of the resistant isolates, which again suggest resistance is being selected for, even in a low usage setting, after their establishment. Our data underscores that a longitudinal framework is pivotal to obtaining a thorough epidemiological understanding of the population dynamics of *E coli*. Furthermore, clones and clades can differ significantly between European countries that are resource similar, highlighting the need for country-specific intensified surveillance efforts to circumvent the threat posed by the current evolutionary trajectory of *E coli* causing bloodstream infections. The dominance of CC131 also offers the possibility of targeted diagnostics in relation to infection control to restrict the spread of this high-risk multidrug resistant clone.

NORM isolates represent a subset of all bloodstream infection cases in Norway used for surveillance of antimicrobial resistance; however, NORM isolates are unbiased in terms of resistance profiles and are considered to be representative of all bloodstream infections in Norway.[5] Although we only sampled the first 22·5% of isolates annually, this subset should retain the representativeness of the NORM collection because *E coli*

bloodstream infections do not exhibit seasonality. Comparing two periods (2002–10 and 2011–17) or trends over time limits any bias in an individual year and the general trends in prevalence observed here concur with the NORM reports.[5] For the analysis using phenotypic antimicrobial susceptibility testing (AST) data, we used categorical AST data reported in NORM based on the breakpoints for each year. Reported AST data were acquired by multiple methods during the time period (appendix 1 p 1). Thus, the breakpoints could not be calibrated over the whole study period. We used an earlier characterisation of ESBL-producers in the NORM surveillance, which showed that in 2017 100% of ESBL-producing *E coli* bloodstream infections were $bla_{CTX\text{-}M}$-positive,[5] to justify using CTX-M-encoding genes as a proxy for ESBL producers, as such we did not include isolates in which only bla_{SHV} or bla_{TEM} were detected in the ESBL and CTX-M analyses. Finally, the published UK collection only represents isolates up until 2011. Given that the second half of our study (2011–17), during which 207 (74·5%) of all 278 CC131 isolates were observed, extends well beyond sampling years of the UK report, it is possible that the importance of clade A within CC131 observed in Norway could have occurred in other locations since 2011.

Our study confirms that the clones of the extra-intestinal pathogenic *E coli* population is under strong balancing selection and that resistance is not necessary for clonal success. However, we also show that even in a low antibiotic use setting, resistance to important antimicrobial classes has rapidly been selected for in a previously susceptible clade. This highlights the importance of future genomic surveillance in uncovering the complex ecology underlying multidrug resistance dissemination and competition, to support the design of strategies and interventions to control the spread of high-risk multidrug resistant clones.

Contributors

AM, PJJ, ØS, and JC conceptualised and sought funding for the study. RAG, AM, PJJ, ØS, and JC developed the methodology. KKB, KWG, RH, AK, HEL, PCL, IHL, ÅM, EN, MTN, GSS, MS, ST, and MV obtained resources for the study. RAG, KS, GGS, and ØS curated the data. ØS and JC verified the underlying data. RAG, FC, MOKC, and BCH did the investigation. GT-H developed and supported Panaroo. JAL developed and supported PopPUNK. RAG analysed and visualised the data. JC analysed the data. RAG, AM, and JC drafted the manuscript. AKP, GT-H, JAL, KS, FC, MOKC, BCH, KKB, KWG, RH, AK, HEL, PCL, IHL, ÅM, EN, MTN, GSS, MS, ST, MV, SDB, NJC, JP, PJJ, and ØS reviewed and edited the manuscript.

Declaration of interests

NJC reports grants from GlaxoSmithKline and personal fees from Antigen Discovery, outside the submitted work. JP reports grants from Wellcome Trust, during the study. All other authors declare no competing interest. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Data sharing

Fastq sequence data are available on the European Nucleotide Archive (ENA). Metadata for samples including ENA accession are included appendix 2. Phylogenetic trees overlaid with metadata are hosted in Microreact and URLs listed in appendix 2.

Acknowledgments

This work was supported by a Trond Mohn Foundation (antimicrobial resistance grant TMS2019TMT04; to RAG, AKP, JC, ØS, and PJJ), Marie Skłodowska-Curie Actions (801133; to AKP), and European Research Council (grant 742158; to JC). Sequencing at the Wellcome Sanger Institute was supported by a core Wellcome Trust grant (206194). We are grateful for technical assistance from Ellen Josefsen, Miriam Nilsen, Lennart Maximillian van Ligtenberg, and all those that prepared and shipped isolates for the study, and thank the Wellcome Sanger Institute sequencing facility and the Wellcome Sanger Institute pathogens informatics team.

References

1 Kern WV, Rieg S. Burden of bacterial bloodstream infection—a brief update on epidemiology and significance of multidrug-resistant pathogens. *Clin Microbiol Infect* 2020; **26:** 151–57.
2 Day MJ, Doumith M, Abernethy J, et al. Population structure of *Escherichia coli* causing bacteraemia in the UK and Ireland between 2001 and 2010. *J Antimicrob Chemother* 2016; **71:** 2139–42.
3 Köhler C-D, Dobrindt U. What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol* 2011; **301:** 642–47.
4 Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, Pitout JDD. Global extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. *Clin Microbiol Rev* 2019; **32:** e00135-18.
5 Norwegian Veterinary Institute. NORM-VET reports. 2017. https://www.vetinst.no/en/surveillance-programmes/norm-norm-vet-report (accessed Dec 22, 2020).
6 Johnson JR, Porter S, Thuras P, Castanheira M. The pandemic H30 subclone of sequence type 131 (ST131) as the leading cause of multidrug-resistant *Escherichia coli* infections in the United States (2011-2012). *Open Forum Infect Dis* 2017; **4:** ofx089.
7 Gerver R, Mihalkova M, Abernethy J, et al. Annual epidemiological commentary: mandatory MRSA, MSSA and *E coli* bacteraemia and *C difficile* infection data, 2014/15. 2015. https://webarchive.nationalarchives.gov.uk/20180412112957/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/442952/Annual_Epidemiological_Commentary_FY_2014_2015.pdf (accessed April 21, 2021).
8 Temkin E, Fallach N, Almagor J, et al. Estimating the number of infections caused by antibiotic-resistant *Escherichia coli* and *Klebsiella pneumoniae* in 2014: a modelling study. *Lancet Glob Health* 2018; **6:** e969–79.
9 Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev* 2014; **27:** 543–74.
10 Kallonen T, Brodrick HJ, Harris SR, 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* 2017; **27:** 1437–49.
11 Roer L, Tchesnokova V, Allesøe R, et al. Development of a web tool for *Escherichia coli* subtyping based on *fimH* alleles. *J Clin Microbiol* 2017; **55:** 2538–43.
12 Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, et al. Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. *MBio* 2016; **7:** e00347–16.
13 Price LB, Johnson JR, Aziz M, et al. The epidemic of extended-spectrum-β-lactamase-producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, H30-Rx. *MBio* 2013; **4:** e00377–13.
14 McNally A, Kallonen T, Connor C, et al. Diversification of colonization factors in a multidrug-resistant *Escherichia coli* lineage evolving under negative frequency-dependent selection. *MBio* 2019; **10:** e00644–19.
15 Page AJ, De Silva N, Hunt M, et al. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Microb Genom* 2016; **2:** e000083.
16 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30:** 2068–69.
17 Tonkin-Hill G, MacAlasdair N, Ruis C, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol* 2020; **21:** 180.
18 Waters NR, Abram F, Brennan F, Holmes A, Pritchard L. Easy phylotyping of *Escherichia coli* via the EzClermont web app and command-line tool. *Access Microbiol* 2020; **2:** acmi000143.

19 Inouye M, Dashnow H, Raven L-A, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 2014; **6:** 90.

20 Lees JA, Harris SR, Tonkin-Hill G, et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res* 2019; **29:** 304–16.

21 Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother* 2017; **72:** 2764–68.

22 Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012; **67:** 2640–44.

23 Hunt M, Mather AE, Sánchez-Busó L, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genom* 2017; **3:** e000131.

24 Huseby DL, Pietsch F, Brandis G, Garoff L, Tegehall A, Hughes D. Mutation supply and relative fitness shape the genotypes of ciprofloxacin-resistant *Escherichia coli*. *Mol Biol Evol* 2017; **34:** 1029–39.

25 Podnecky NL, Fredheim EGA, Kloos J, et al. Conserved collateral antibiotic susceptibility networks in diverse clinical strains of *Escherichia coli*. *Nat Commun* 2018; **9:** 3673.

26 Pietsch F, Bergman JM, Brandis G, et al. Ciprofloxacin selects for RNA polymerase mutations with pleiotropic antibiotic resistance effects. *J Antimicrob Chemother* 2017; **72:** 75–84.

27 Croucher NJ, Page AJ, Connor TR, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015; **43:** e15.

28 Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006; **22:** 2688–90.

29 Harris SR. SKA: Split Kmer Analysis Toolkit for Bacterial Genomic Epidemiology. *bioRxiv* 2018; published online Oct 25. https://doi.org/10.1101/453142 (preprint).

30 Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007; **7:** 214.

31 Decano AG, Downing T. An *Escherichia coli* ST131 pangenome atlas reveals population structure and evolution across 4071 isolates. *Sci Rep* 2019; **9:** 17394.

32 Stoesser N, Sheppard AE, Pankhurst L, et al. Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio* 2016; **7:** e02162.