Global *Escherichia coli* Sequence Type 131 Clade with $bla_{CTX\text{-}M\text{-}27}$ Gene

Yasufumi Matsumura, Johann D.D. Pitout, Ryota Gomi, Tomonari Matsuda, Taro Noguchi, Masaki Yamamoto, Gisele Peirano, Rebekah DeVinney, Patricia A. Bradford, Mary R. Motyl, Michio Tanaka, Miki Nagao, Shunji Takakura, Satoshi Ichiyama

The *Escherichia coli* sequence type (ST) 131 C2/*H*30Rx clade with the $bla_{CTX\text{-}M\text{-}15}$ gene had been most responsible for the global dissemination of extended-spectrum β-lactamase (ESBL)–producing *E. coli*. ST131 C1/*H*30R with $bla_{CTX\text{-}M\text{-}27}$ emerged among ESBL-producing *E. coli* in Japan during the late 2000s. To investigate the possible expansion of a single clade, we performed whole-genome sequencing for 43 Japan and 10 global ST131 isolates with $bla_{CTX\text{-}M\text{-}27}$ (n = 16), $bla_{CTX\text{-}M\text{-}14}$ (n = 16), $bla_{CTX\text{-}M\text{-}15}$ (n = 13), and others (n = 8). We also included 8 ST131 genomes available in public databases. Core genome-based analysis of 61 isolates showed that ST131 with $bla_{CTX\text{-}M\text{-}27}$ from 5 countries formed a distinct cluster within the C1/*H*30R clade, named C1-M27 clade. Accessory genome analysis identified a unique prophage-like region, supporting C1-M27 as a distinct clade. Our findings indicate that the increase of ESBL-producing *E. coli* in Japan is due mainly to emergence of the C1-M27 clade.

The global increase in resistance to the third-generation cephalosporins and fluoroquinolones among extraintestinal pathogenic *Escherichia coli* (ExPEC) is a public health concern because of the importance of these drugs in treating serious infections (*1*). The extended-spectrum β-lactamases (ESBLs), especially CTX-M types, contribute to third-generation cephalosporin resistance among ExPEC, and specific mutations in quinolone resistance–determining regions in *gyrA* and *parC* mainly contribute to fluoroquinolone resistance (*2*). The increase in resistance among ExPEC has resulted mainly from the recent expansion of a pandemic clonal group known as *E. coli* sequence type (ST) 131, which is usually multidrug resistant and is associated with CTX-M-15, the most prevalent β-lactamase among ESBL-producing ExPEC (*2*). ST131 harbors more virulence factors than other antimicrobial-resistant ExPEC and can cause severe infections (*2,3*).

Recent studies using whole-genome sequencing (WGS) analysis revealed that ST131 comprises different lineages or clades (*4,5*). Price et al. found a dominant fluoroquinolone-resistant lineage (named *H*30R) in North America that contains the *fimH* 30 allele and was associated with characteristic quinolone resistance–determining region mutations (*2,4*). ST131 with the $bla_{CTX\text{-}M\text{-}15}$ gene formed a distinct cluster within the *H*30R lineage, referred to as the *H*30Rx clade (*4*). Petty et al. confirmed these findings using a collection of strains from 6 countries (*5*). In their study, *H*30R and *H*30Rx clades correspond to clade C and clade C2 (subset of clade C), respectively. The other clade C subset, clade C1, included ST131 isolates with different CTX-Ms than $bla_{CTX\text{-}M\text{-}15}$.

Globally, the CTX-M-15–producing C2/*H*30Rx clade is mostly responsible for the pandemic of ExPEC with ESBLs (*2*), but in Japan, ExPEC with $bla_{CTX\text{-}M\text{-}15}$ is rare despite the predominance of ST131 among ESBL-producing isolates (*6*). Before 2005, ST131 C1/*H*30R negative for Rx containing $bla_{CTX\text{-}M\text{-}14}$ predominated among Japanese ST131 (*6*). In 2006, ST131 C1/*H*30R with $bla_{CTX\text{-}M\text{-}27}$ was detected in Japan, and the numbers of this lineage escalated since 2010 and are responsible for the substantial increase of ESBL-producing ExPEC in Japan (*6*). Moreover, $bla_{CTX\text{-}M\text{-}27}$ is confined to ST131, whereas other CTX-Ms, such as $bla_{CTX\text{-}M\text{-}14}$ and $bla_{CTX\text{-}M\text{-}15}$, are equally present among ST131 and non-ST131 *E. coli* isolates (*3*).

$bla_{CTX\text{-}M\text{-}27}$ is an infrequent global $bla_{CTX\text{-}M}$ allele that differs by only 1 nt from $bla_{CTX\text{-}M\text{-}14}$, which results in 1 aa change at position 240 (*1,6*). ST131 with CTX-M-27 had previously been reported from other countries, such as Korea (isolation year 2008), China (2013–2014), Australia (2009–2010), Nepal (2013–2014), Cambodia (2004–2005), Israel (2008–2009), Czech Republic (2008–2011), Switzerland (2011), Spain (2012), France (2012), Portugal (2013–2014), Netherlands (2011), Canada (2005), and United States (2013) (*2, 5–15*). Because of the rapid

Author affiliations: Kyoto University Graduate School of Medicine, Kyoto, Japan (Y. Matsumura, T. Noguchi, M. Yamamoto, M. Tanaka, M. Nagao, S. Takakura, S. Ichiyama); University of Calgary, Calgary, Alberta, Canada (J.D.D. Pitout, G. Peirano, R. DeVinney); Kyoto University Graduate School of Engineering, Kyoto (R. Gomi, T. Matsuda); AstraZeneca Pharmaceuticals LP, Waltham, Massachusetts, USA (P.A. Bradford); Merck & Co., Inc., Rahway, New Jersey, USA (M.R. Motyl)

DOI: http://dx.doi.org/10.3201/eid2211.160519

increase of CTX-M-27–producing ST131 in Japan since 2010 (*6*), we designed a study to characterize these isolates using WGS techniques.

Materials and Methods

Bacterial Isolates

We selected 43 nonduplicate ST131 clinical isolates collected from 2 multicenter surveillance programs in Japan for WGS to represent 3 major ESBL-producing ST131 (CTX-M-27–producing *H*30R, 13 isolates; CTX-M-14–producing *H*30R, 9 isolates; CTX-M-15–producing *H*30Rx, 11 isolates) and other ST131 (CTX-M-14+CTX-M-15–producing *H*30Rx, 2 isolates; CTX-M-14–producing *H*30Rx, 1 isolate; CTX-M-14–producing *H*22, 1 isolate; CTX-M-2–producing *H*22, 1 isolate; TEM-producing *H*30, 2 isolates; non–ESBL-producing *H*30R, 3 isolates) in Japan (*6*) (Table). One of the surveillance programs collected ESBL-producing *E. coli* isolates during 2001–2010 at 10 acute-care hospitals in the Kyoto and Shiga prefectures of Japan (*6*); the other program collected all *E. coli* isolates during December 2014 at 10 acute-care hospitals in the 5 prefectures in central Japan. ST131 isolates were identified by PCR specific for *mdh* and *gyrB* alleles, O25b or O16 *rfb* variants, *fimH* allele, and *H*30Rx status (*6*). The selection process of the Japanese ST131 ensured equal representation by geographic location, specimen type, and date of isolation.

In addition to isolates from Japan, we obtained 10 CTX-M–producing ST131 isolates from global collections that previously had been characterized by multilocus sequence typing (MLST) (Table; online Technical Appendix Table 1, http://wwwnc.cdc.gov/EID/article/22/11/16-0519-Techapp1.pdf). We selected all of the CTX-M-27 producers, 1 CTX-M-14 producer per country, and 2 CTX-M-15 producers. We also sought public databases for ST131 *H*30 and included sequence data for 8 isolates from countries other than Japan: CTX-M-27 producers (3 raw reads, 2 draft genomes); CTX-M-14 producer (1 raw read); and CTX-M-15–producing C2/*H*30Rx (2 complete genomes) (Table; online Technical Appendix Table 1) (*5,9,16–18*).

WGS

We used the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, USA) to prepare libraries for sequencing. Samples were multiplexed and sequenced on an Illumina MiSeq for 600 cycles (300-bp paired-end) or NextSeq500 for 300 cycles (151-bp paired-end). The ST131 genomes were sequenced at an average depth of 44.03 (SD ± 14.70) and an average coverage of 97.73% (SD ± 0.93%) using the 5,109,767-bp EC958 chromosome as previously described (*16*).

Core Genome Analysis

We used a core genome single-nucleotide polymorphism (SNP)–based approach to create a phylogenetic tree. We identified SNPs using raw read mapping followed by duplicate read removal, realignment, quality score recalibration, and variant filtering (online Technical Appendix). Reads from 53 isolates sequenced in this study and 4 isolates (S100EC, S107EC, S108EC, and S135EC) (*5*) were aligned against a reference genome of EC958, and SNPs were called. The remaining 4 draft or complete genomes underwent whole-genome alignment against EC958 to make EC958-like pseudo-chromosomes that contained only SNPs. The SNP-only core genome was identified as the blocks of >500 bp common to all 61 study isolates to ensure that each block represented a common segment from good alignment in each isolate and that the block had enough length to enable identification (*5*). A maximum-likelihood tree was built using RAxML (*19*). A recombination-free tree was also build by excluding recombination sites identified using a Bayesian analysis software BRATNextGen (*20*).

Comparative Genomic Analysis

To define presence of genes and their alleles, we used SRST2 with trimmed reads or BLAST+ (executables [http://blast.ncbi.nlm.nih.gov/]) with assembled draft genomes and following databases or typing schemes: ResFinder antimicrobial resistance gene database, VFDB and VirulenceFinder virulence gene databases, serotypeFinder O:H typing database, PlasmidFinder plasmid replicon database, MLST (http://mlst.ucc.ie/mlst/dbs/Ecoli), plasmid MLST, *fimH* typing, *gyrA*/*parC* typing, ST131 virotyping, and detection of *H*30Rx-specific *ybbW* SNPs, plasmid addiction systems, and bla_{CTX-M} genetic environment (online Technical Appendix). We used pangenome analysis to identify clade specific segments among draft or complete genomes. BRIG was used to visualize similarity of genomes to ST131 genomic islands (*16*) and to the ST131 reference plasmid pEC958 (*21*).

Statistical Analysis and Sequence Data Accession Numbers

We compared categorical variables using Fisher exact test. A p value <0.05 was considered statistically significant. We conducted our statistical analysis using Stata, version 13.1 (StataCorp, College Station, TX, USA). The sequences were deposited in the DDBJ Sequence Read Archive database (accession no. DRA004266 and DRA004267).

Results

Bacterial Isolates

The study comprised 60 clinical and 1 environmental ST131 isolates (Table; online Technical Appendix Table

Table. ST131 isolates included in study of EBSL-producing *Escherichia coli*, Japan*

	Country/prefecture of isolation (no. isolates; year)			
	*H*30			
	*H*30R			
Type of ESBL	*H*30R, n = 39	*H*30Rx, n = 18	*H*30, n = 2	*H*22, n = 2
CTX-M-27, n = 21	Japan/Kyoto, Shiga, Aichi (13; 2004–2014), Australia (3; 2009, 2010),† United States (2; 2013, 2014),‡ Canada (1; 2008), Thailand (1; 2013), Vietnam (1; 2011)			
CTX-M-14, n = 17	Japan/Kyoto, Shiga, Hyogo (9; 2002–2014), Canada (2; 2005, 2009),† France (1; 2008), New Zealand (1; 2010), South Africa (1; 2008), United States (1; 2008)	Japan/Kyoto (1; 2009)		Japan/Kyoto (1; 2007)
CTX-M-15, n = 15		Japan/Kyoto, Shiga, Osaka (11; 2006–2014), Canada (2; 2009), UK (1; 2005),§ United States (1; 2008)¶		
CTX-M-14 and CTX-M-15, n = 2		Japan/Kyoto (2; 2010, 2014)		
CTX-M-2, n = 1				Japan/Kyoto (1; 2004)
TEM, n = 2			Japan/Kyoto (2; 2005, 2009)	
Negative, n = 3	Japan/Shiga, Hyogo, Osaka (3; 2014)			

*ESBL, extended-spectrum β-lactamase; ST, sequence type.
†Raw reads were downloaded from European Nucleotide Archive (accession no. ERA118286) for 3 isolates from Australia (S100EC, S107EC, S108EC) and 1 isolate from Canada (S135EC).
‡MRSN17749 draft genome (GenBank assembly accession no. GCA_000770275.1) and IEH71520 draft genome (GenBank assembly accession no. GCA_000681435.1).
§EC958 complete genome (GenBank assembly accession no. GCA_000285655.3).
¶JJ1886 complete genome (GenBank assembly accession no. GCA_000493755.1).

1). We confirmed the types of β-lactamase genes, ST131 status, *fimH* allele numbers, and *H*30Rx status using draft genomes.

Core Genome SNP-based Phylogenetic Tree

We identified a 4,086,650-bp core genome that included 5,280 SNPs by mapping and alignment of the 61 study isolates to EC958 (Figure 1). The ciprofloxacin-resistant isolates with *gyrA* 1AB and *parC* 1aAB alleles formed the C/*H*30R cluster that comprised the C2/*H*30Rx and C1/*H*30R clades. The C2/*H*30Rx clade included isolates with $bla_{\text{CTX-M-15}}$ (n = 15) and $bla_{\text{CTX-M-14}}$ (n = 1) and isolates with both $bla_{\text{CTX-M-15}}$ and $bla_{\text{CTX-M-14}}$ (n = 2) (Figure 1). The C1/*H*30R clade included isolates with $bla_{\text{CTX-M-27}}$ (n = 21) and $bla_{\text{CTX-M-14}}$ (n = 14) and isolates without ESBLs (n = 3) (Figure 1). Within the C1/*H*30R clade, 19 of 21 CTX-M-27–producing isolates clustered into a distinct group, named the C1-M27 clade (Figure 1). *E. coli* ST131 C1-M27 comprised isolates from Japan (n = 13; isolation years 2004–2014), Australia (n = 2; 2009–2010), United States (n = 2; 2013–2014), Canada (n = 1, 2008), and Thailand (n = 1, 2013).

Analysis of the core genome showed that 79 segments (i.e., 304,782 bp, including 3,453 SNPs) were associated with recombination sites (online Technical Appendix Figure 1). This finding suggests that recombinant segments contained 65% of SNPs with subsequent higher frequency of SNPs compared with nonrecombinant regions (average 11 vs. 0.48 SNPs/kb, respectively). The phylogenetic tree created without recombination sites showed the same results as the phylogenetic tree obtained with recombination sites (online Technical Appendix Figure 2). In addition to the core genome–based phylogeny with or without recombination sites, the C1-M27 clade was defined by a unique accessory genome of the M27PP1.

The C1-M27 Clade–Specific Region

The pangenome analysis of genomes from all the isolates identified an 11,894-bp region named M27PP1 that was specific to all the isolates from the C1-M27 clade. Further analysis using the BLAST database and Sanger sequencing for gap filling showed that this region was identical to a prophage-like genomic island (GenBank accession no. CP006632) in *E. coli* PCN033 that belonged to phylogenetic group D and was isolated from a pig in China. The BLAST database also identified 2 similar sequences (i.e., 99.9% homology): A CMY-2 containing plasmid pEQ011 (GenBank accession no. NC_023315) in an *E. coli* isolate from a horse in Ireland (*22*) and a multidrug-resistant plasmid pSD853_88 (GenBank accession no. JF267652) found in a bovine *Salmonella enterica* isolate in the United States. M27PP1 was inserted into chromosome creating a 7-bp direct repeat region (Figure 2). PCN033 had the same flanking structure as the M27PP1,

whereas the 2 plasmids (pEQ011, pSD853_89) contained only a 44-bp similar segment at 5′ side and other parts of these plasmids were not found in the C1-M27 clade isolates.

Two *E. coli* ST131 C1-M27 isolates (i.e., KUN5781 and Ec 24) had an additional insertion region of 19,352 bp, named M27PP2, situated upstream of M27PP1. M27PP2 was accommodated within the same 7-bp direct repeat region (Figure 2). M27PP2 included a 15,555-bp region that showed 88.9% homology to a prophage-like sequence in the chromosome of the γ proteobacterium HdN1 (GenBank

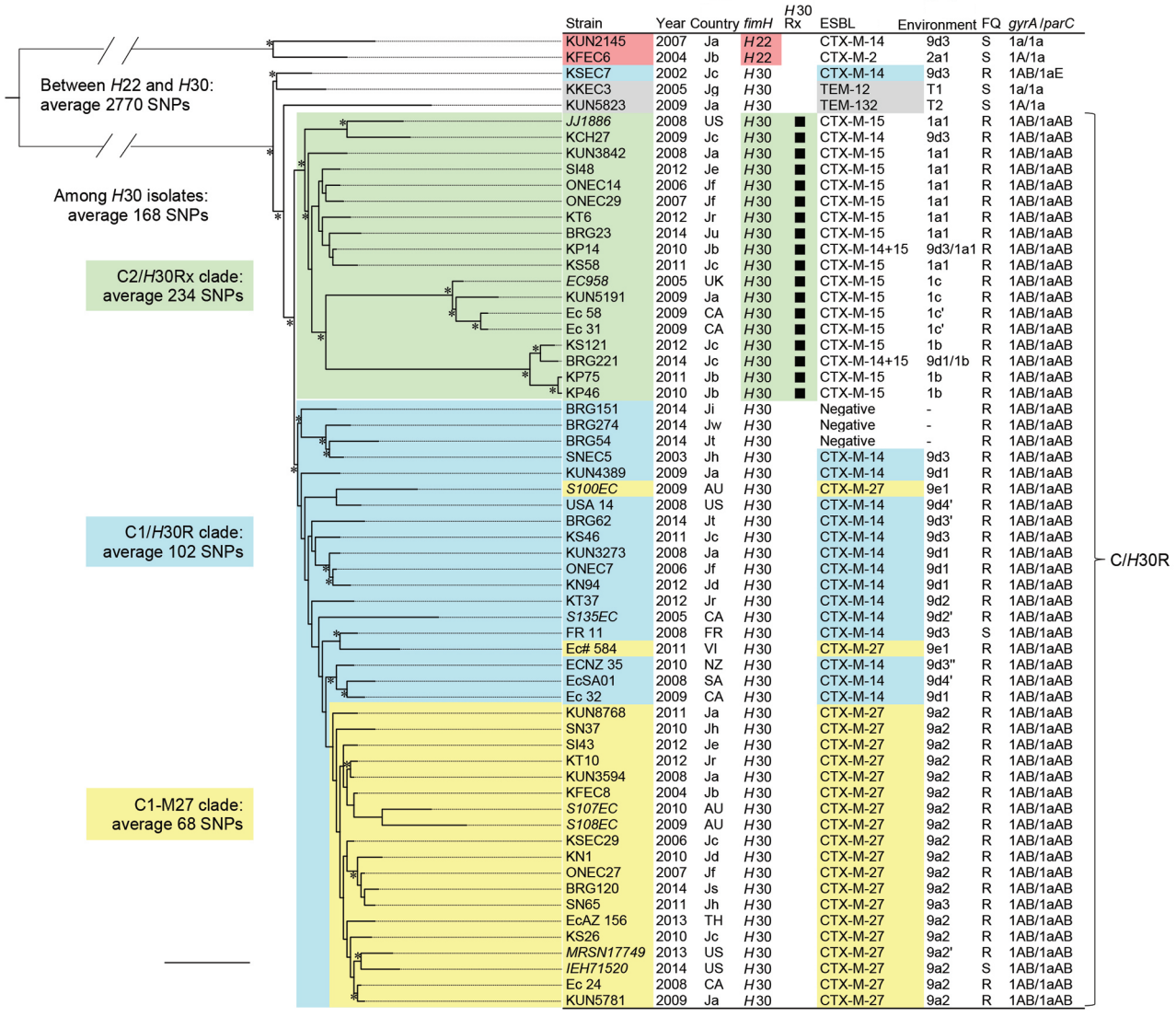


Figure 1. Core genome single-nucleotide polymorphism (SNP)–based phylogenetic tree of *Escherichia coli* sequence type (ST) 131 isolates. This maximum-likelihood phylogram is based on a 4,086,650-bp core genome and a total of 5,280 SNPs. The tree is rooted by using the outgroup *H*22 isolates, and asterisks indicate bootstrap support >90% from 100 replicates. Strains that had previously been sequenced are in italics. The Country columns indicate places of isolation: Ja to Jw, Japan (a to w indicates hospitals); AU, Australia; CA Canada; FR, France; NZ, New Zealand; SA, South Africa; TH, Thailand; UK, United Kingdom; US, United States; VI, Vietnam. Environment column shows a type of genetic environment of ESBL genes (online Technical Appendix Table 2, http://wwwnc.cdc.gov/EID/article/22/11/16-0519-Techapp1.pdf). FQ columns indicate ciprofloxacin susceptibilities (S, susceptible; R, resistant). KSEC7 had a *parC* 1aE allele including G250A (S80K) mutation in addition to a 1a allele. The ciprofloxacin-resistant C/*H*30R cluster comprised the C2/*H*30Rx and C1/*H*30R clades. All of the *H*30Rx isolates belonged to the C2/*H*30Rx clade. The C1/*H*30R clade included CTX-M-14–producing *H*30R, non–ESBL-producing *H*30R, and CTX-M-27–producing *H*30R isolates. CTX-M-27–producing isolates belonged to the C1-M27 clade within the C1/*H*30R clade except 2 isolates (S100EC and EC# 584). The bootstrap value for the root of the C1-M27 clade was 64%. An average of 68 SNPs was found among the C1-M27 clade, whereas an average of 158 SNPs was found between the C1-M27 clade and 2 non–C1-M27 clade isolates with $bla_{CTX-M-27}$. Scale bar indicates 100 SNPs.

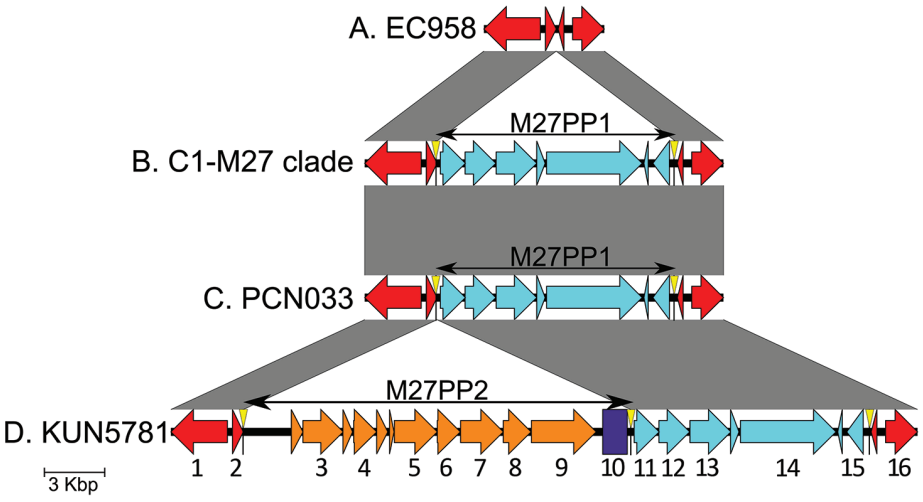


Figure 2. Genetic environments of the C1-M27 clade–specific region of *Escherichia coli*. All isolates other than the C1-M27 clade isolates had the type A structure in their chromosome (red arrows; gene locus tags shown in the bottom are annotated according to EC958). The C1-M27 clade isolates except 2 isolates (KUN5781 and Ec 24) had the type B structure. A 11,894-bp region (M27PP1; predicted genes shown in light blue arrows) is inserted into the type A structure creating the 7-bp direct repeat (CCGTTCT; yellow triangle). The inserted sequence M27PP1 is identical to a prophage-like genomic island in *E. coli* PCN033 chromosome (GenBank accession no. CP006632), which had the similar flanking structure (structure C, 98.8% similarity). M27PP1 included phage-like integrase and recombinase. The identical M27PP1 sequence was found in all of the C1-M27 isolates with the use of additional Sanger sequencing. Only the draft genome of IEH71520 had 98.7% coverage to the M27PP1 sequence because of contig discontinuity. KUN5781 and Ec 24 had the type D structure, of which an additional 19,352-bp region (M27PP2) is inserted into the type B structure by creating the same 7-bp direct repeat (yellow triangle). The M27PP2 includes a total of 15,555-bp region (genes shown in orange arrows), which was 88.9% similar to a prophage-like region in γ proteobacterium HdN1 chromosome (GenBank accession no. FP929140) and a following 1,221-bp region is 99.8% similar to IS*Sen4* (purple box). Code to gene locus tags: 1, 958RS23365; 2, 958RS23370; 3, HDN1F03950; 4, HDN1F03970; 5, HDN1F04000; 6, HDN1F04010; 7, HDN1F04020; 8, HDN1F04030; 9, HDN1F04040; 10, ISSen4; 11, 033RS22420; 12, 033RS22425; 13, 033RS22430; 14, 033RS22440; 15, 033RS22450; 16, 958RS23380.

accession no. FP929140) and 99.8% homology to the insertion element IS*Sen4*.

Genomic Comparison of the ST131 Genomic Islands and Virulence Genes

The sequences of the study isolates were similar to the ST131 genomic islands in EC958 and JJ1886 (a CTX-M-15–producing C2/*H*30Rx strain obtained in the United States from a patient with fatal urosepsis) (Figure 3) (*17*). The C1-M27 clade isolates lacked the prophage 1 region present in EC958 (Figure 3). This prophage 1 region, specific for ST131, was present among the non–C1-M27 ST131 isolates in this study, except for BRG23 and EcSA01. The presence of ExPEC-associated virulence genes is shown in online Technical Appendix Figure 3. The *senB* enterotoxin gene was more common in C1/*H*30R (than in C2/*H*30Rx). No significant differences existed in the distribution of virulence genes between *E. coli* ST131 C1-M27 and other isolates.

Plasmid Replicons, Addiction Systems, and Antimicrobial Drug Resistance Genes

We compared the study isolates with pEC958, the plasmid present in EC958 that carries $bla_{\text{CTX-M-15}}$ (online Technical Appendix Figure 4). The C1-M27 clade lacked the first part of the transfer region (*tra*) present in pEC958. Some regions common to both C2/*H*30Rx and C1/*H*30R clades were present in pEC958. The C1/*H*30R clade producing CTX-M-27 or CTX-M-14 (including C1-M27) contained mostly F1:A2:B20 replicons, whereas the C2/*H*30Rx clade producing CTX-M-15 contained mainly F2:A1:B- replicons (online Technical Appendix Figure 5). The C1-M27 clade was negative for Tn*2* containing $bla_{\text{TEM-1}}$. Two C1-M27 isolates from Thailand and the United States were also positive for $bla_{\text{NDM-1}}$ (online Technical Appendix Figure 5).

Discussion

A previously unreported clade named C1-M27 within C1/*H*30R clade is responsible for the epidemic of ESBL-producing ExPEC in Japan and has already been disseminated to 5 countries on 3 continents. ST131 containing $bla_{\text{CTX-M-27}}$ responsible for human infections has been reported from various continents (*2*) and is especially common among ESBL-producing ExPEC in certain countries, such as Israel, the Czech Republic, and Switzerland (*2,13,14*). CTX-M-27–producing ST131 also is present among nonclinical and nonhuman *E. coli* isolates, including fecal specimens of healthy children attending day care centers in France; fecal specimens of healthy adults in China, Portugal, and the Netherlands; samples from sick dogs and cats in Japan; samples from water birds from central Europe and Swiss rivers and lakes; and samples of well water from China (*2,10,11,15,23–25*). The most common ESBL among *E. coli* ST131 in nonhuman samples is CTX-M-27 (*2,23–25*). ST131 with $bla_{\text{CTXM-15}}$ is rare among animal and environmental *E. coli* isolates (*26*). Our analysis of IEH71520,

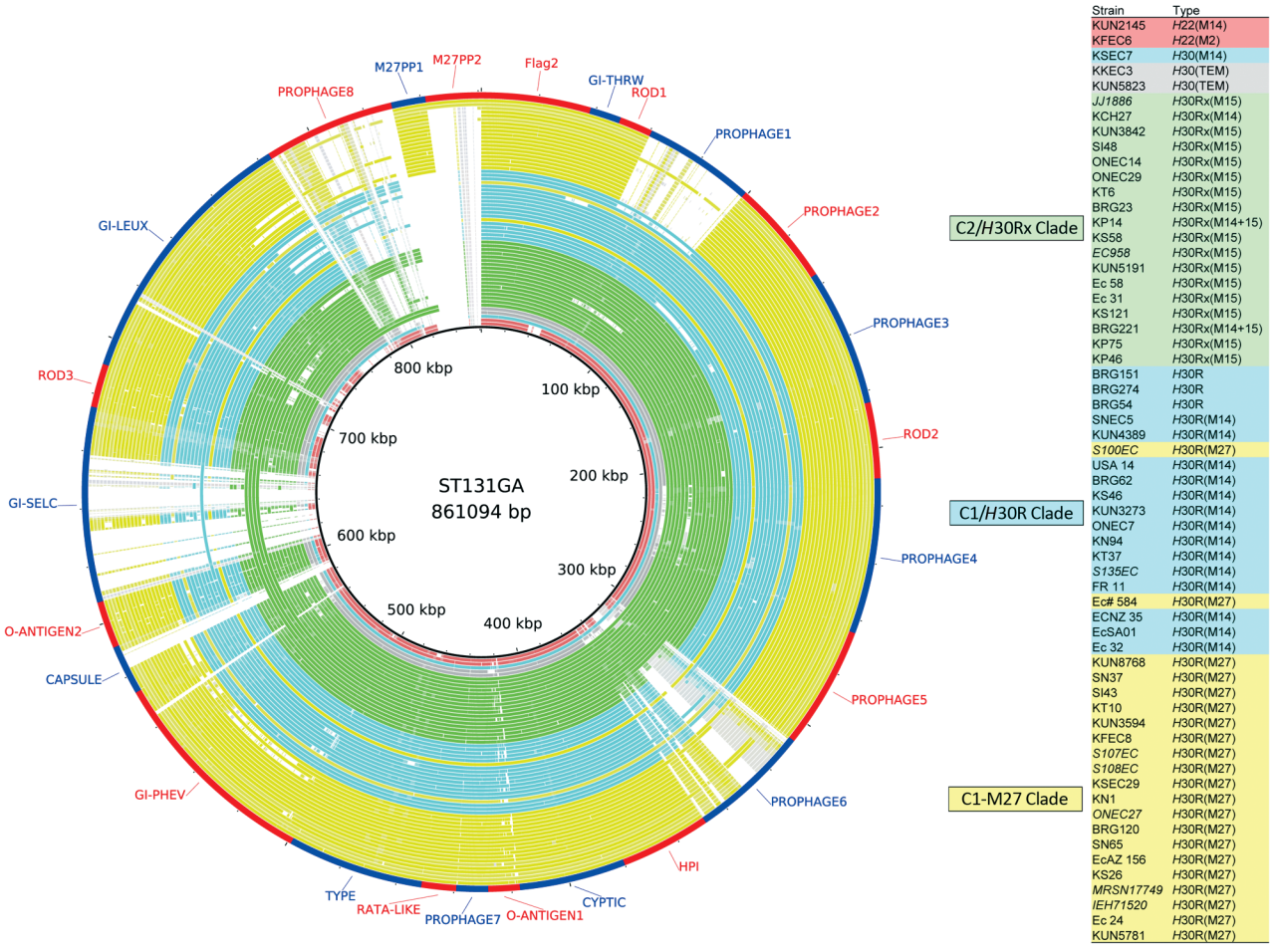


Figure 3. Genome similarities to the *Escherichia coli* sequence type (ST) 131 genomic islands and the C1-M27 clade–specific region. Rings drawn by BRIG show the presence of these regions. Colored segments indicate >90% similarity and gray segments indicate >70% similarity by BLAST comparison between the regions of interest and each genome. Extended-spectrum β-lactamase types are indicated in parentheses of Type column. The regions from Flag2 to GI-*lueX* were found in EC958, the prophage 8 region was found in JJ1886, and the M27PP1 and M27PP2 were found as the C1-M27 clade–specific regions in this study. Prophage 6, capsule, GI-*selC*, and prophage 8 regions were present in some C2/*H*30Rx isolates but were absent in C1/*H*30R isolates.

an *E. coli* isolate from vacuum cleaner dust in the United States (*15*), showed that this ST131 isolate belong to the C1-M27 clade. The C1-M27 clade is likely to be present among animal and environmental ST131, and such isolates might act as a hidden reservoir for the introduction of ST131 containing $bla_{CTX-M-27}$ into human medicine.

E. coli ST131 C1-M27 had an additional, unique prophage-like region (M27PP1) within its chromosome, lacked the prophage 1 genomic island previously identified in ST131 C2/*H*30Rx, and were negative for the transposon Tn*2* containing bla_{TEM-1} (Figure 3; online Technical Appendix Figure 5). The direct flanking repeat sequences surrounding M27PP1 suggest that this region was introduced into *E. coli* ST131 C1/*H*30R with $bla_{CTX-M-27}$ by a recombination event that was then followed by the clonal expansion of the C1-M27 clade.

Recent studies focusing on evolutionary history of ST131 suggested that C1/*H*30R and C2/*H*30Rx clades emerged ≈30 years ago, after their acquisition of *gyrA*-1AB and *parC*-1aAB alleles from C0/*H*30 (non-R) clade (*27,28*). The phylogeny and smaller numbers of SNPs in the C1-M27 clade (Figure 1) suggest that this clade was recently diverged from the C1/*H*30R. In the time-scaled phylogeny presented by Stoesser et al. (*27*), a cluster that included 6 CTX-M-27–producing isolates from Cambodia, Thailand, and Laos in 2007–2011 was present within the C1/*H*30R clade. This cluster, supposed to be the C1-M27 clade, diverged in the early 2000s, supporting our hypothesis.

CTX-M-27–producing ST131 that belongs to the *H*41 lineage previously had been described from Japan (*6*) and China (*15*). The characterization of the Japanese ST131

*H*41 isolates showed different genetic structures flanking the $bla_{\text{CTX-M-27}}$ from those structures present in *E. coli* ST131 *H*30R (*6*). The flanking regions previously characterized in ST131 *H*41 were identical to the flanking regions in ST131 non–C1-M27 from this study. It seems there are 2 types of structures flanking the $bla_{\text{CTX-M-27}}$ among *E. coli* ST131; 1 type is confined to clade C1-M27 (i.e., 208 bp of ΔIS*Ecp1* upstream and ΔIS*903D* downstream), whereas another structure (i.e., 388 bp of ΔIS*Ecp1* upstream and full IS*903D* downstream) is distributed among non–C1-M27 isolates, including ST131 *H*41 (*6*). Therefore, ST131 *H*41, through horizontal transfer of $bla_{\text{CTX-M-27}}$, is unlikely to have played a substantial role in the emergence of the C1-M27 clade.

Two ST131 isolates with $bla_{\text{CTX-M-27}}$ from Australia and Vietnam did not belong to the C1-M27 clade (Figure 1). These 2 isolates differ from the C1-M27 clade in that their core genomes had more SNPs (158 vs. 68), contained the prophage 1 ST131-specific region, and lacked the M27PP1 and M27PP2 elements. Moreover, the genetic environment surrounding the $bla_{\text{CTX-M-27}}$ differed from *E. coli* ST131 C1-M27 (as described previously). The isolate from Vietnam lacked *mph*(A)-*mrx*-*mphR*, *tetR*-*tet*(A), *sul2*-*strA*-*strB*, and In*54* resistance genes, compared with the C1-M27 clade (online Technical Appendix Figure 5). These differences indicate that some ST131 isolates might acquire $bla_{\text{CTX-M-27}}$ independently from the C1-M27 clade.

Our study has several limitations. Most isolates originated from Japan. However, we were able to include ST131 C1-M27 isolates from 5 countries on 3 continents and C1/*H*30R isolates producing CTX-M-14 or CTX-M-15 from 6 countries on 4 continents. Another limitation was that we were able to obtain only 1 environmental ST131 isolate with $bla_{\text{CTX-M-27}}$ (IEH71520). Future studies that include environmental isolates will provide additional insights into molecular epidemiology and evolutionary history of the C1-M27 clade. We could not analyze plasmid contents of $bla_{\text{CTX-M-27}}$ because $bla_{\text{CTX-M-27}}$–containing contigs were too short. The sequencing of plasmids that contain $bla_{\text{CTX-M-27}}$ obtained from various ST131 clades (including the C1-M27 clade) should also be undertaken.

In conclusion, we showed that the recent increase in ESBL-producing *E. coli* from Japan resulted from emergence of a ST131 C1/*H*30R subclade with $bla_{\text{CTX-M-27}}$. This clade, named C1-M27, had unique genomic characteristics and was present in ST131 from Thailand, Australia, Canada, and the United States. Our findings suggest that the C1-M27 clade is contributing to the global success of ST131. *E. coli* ST131 C1-M27 poses a major new public health threat because of its global distribution and association with the very dominant C/*H*30 lineage. We urgently need rapid cost-effective detection methods for *E. coli* ST131 C1-M27 and well-designed epidemiologic and molecular studies to understand the dynamics of transmission, risk factors, and reservoirs for ST131 C1-M27. These efforts will provide insight into the emergence and spread of this multidrug-resistant clade that will lead to information essential for preventing the spread of ST131.

Acknowledgments

We thank the members of the Kyoto-Shiga Clinical Microbiology Study Group (Naohisa Fujita, Toshiaki Komori, Yukiji Yamada, Tsunehiro Shimizu, Akihiko Hayashi, Tamotsu Ono, Harumi Watanabe, Naoko Fujihara, Takeshi Higuchi, Kunihiko Moro, Masayo Shigeta, Kaneyuki Kida, Hiromi Terada, Fusayuki Tsuboi, Yoshihisa Sugimoto, and Chiyoko Fukumoto) and the members of the 89JAID–BRG (89th Annual Meeting of the Japanese Association for Infectious Diseases–drug-resistant Bacteria Research Group; Ooi Yukimasa, Tomohiro Higashiyama, Masaaki Sasano, Harumi Watanabe, Tsunehiro Shimizu, Akihiko Hayashi, Yasufumi Matsumura, Taro Noguchi, Miki Nagao, Masaki Yamamoto, Michio Tanaka, Takeshi Higuchi, Toru Kanahashi, Naohisa Fujita, Toshiaki Komori, Yukiji Yamada, Kohei Ueda, Go Yamamoto, Kunihiko Moro, and Shigenobu Yabu) for collection of isolates. We also thank Sayo Shitashiro and Karen Poon for technical assistance.

This work was supported by Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Young Scientists (B) Grant No. 15K19179 (Y.M.), the John Mung Program from Kyoto University, Japan (Y.M.), and a research grant from the Calgary Laboratory Services (#10006465; J.D.D.P). J.D.D.P had previously received research funds from Merck and Astra Zeneca.

Dr. Matsumura is an assistant professor at the Clinical Laboratory Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan. His main research focuses on the detection and molecular epidemiology of antimicrobial drug resistance mechanisms among gram-negative bacteria.

References

1. Pitout JD, Laupland KB. Extended-spectrum beta-lactamase–producing *Enterobacteriaceae*: an emerging public-health concern. Lancet Infect Dis. 2008;8:159–66. http://dx.doi.org/10.1016/S1473-3099(08)70041-0
2. Nicolas-Chanoine MH, Bertrand X, Madec JY. *Escherichia coli* ST131, an intriguing clonal group. Clin Microbiol Rev. 2014;27:543–74. http://dx.doi.org/10.1128/CMR.00125-13
3. Matsumura Y, Yamamoto M, Nagao M, Ito Y, Takakura S, Ichiyama S, et al. Association of fluoroquinolone resistance, virulence genes, and IncF plasmids with extended-spectrum-β-lactamase–producing *Escherichia coli* sequence type 131 (ST131) and ST405 clonal groups. Antimicrob Agents Chemother. 2013;57:4736–42. http://dx.doi.org/10.1128/AAC.00641-13
4. Price LB, Johnson JR, Aziz M, Clabots C, Johnston B, Tchesnokova V, 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.

5. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. Proc Natl Acad Sci U S A. 2014;111:5694–9. http://dx.doi.org/10.1073/pnas.1322678111
6. Matsumura Y, Johnson JR, Yamamoto M, Nagao M, Tanaka M, Takakura S, et al. CTX-M-27– and CTX-M-14–producing, ciprofloxacin-resistant *Escherichia coli* of the *H*30 subclonal group within ST131 drive a Japanese regional ESBL epidemic. J Antimicrob Chemother. 2015;70:1639–49.
7. Sherchan JB, Hayakawa K, Miyoshi-Akiyama T, Ohmagari N, Kirikae T, Nagamatsu M, et al. Clinical epidemiology and molecular analysis of extended-spectrum-β-lactamase–producing *Escherichia coli* in Nepal: characteristics of sequence types 131 and 648. Antimicrob Agents Chemother. 2015;59:3424–32. http://dx.doi.org/10.1128/AAC.00270-15
8. Clermont O, Dhanji H, Upton M, Gibreel T, Fox A, Boyd D, et al. Rapid detection of the O25b-ST131 clone of *Escherichia coli* encompassing the CTX-M-15–producing strains. J Antimicrob Chemother. 2009;64:274–7. http://dx.doi.org/10.1093/jac/dkp194
9. McGann P, Snesrud E, Ong AC, Appalla L, Koren M, Kwak YI, et al. War wound treatment complications due to transfer of an IncN plasmid harboring $bla_{OXA-181}$ from *Morganella morganii* to CTX-M-27–producing sequence type 131 *Escherichia coli*. Antimicrob Agents Chemother. 2015;59:3556–62. http://dx.doi.org/10.1128/AAC.04442-14
10. Rodrigues C, Machado E, Fernandes S, Peixe L, Novais Â. An update on faecal carriage of ESBL-producing *Enterobacteriaceae* by Portuguese healthy humans: detection of the *H*30 subclone of B2–ST131 *Escherichia coli* producing CTX-M-27. J Antimicrob Chemother. 2016;71:1120–2. http://dx.doi.org/10.1093/jac/dkv443
11. Reuland EA, Al Naiemi N, Kaiser AM, Heck M, Kluytmans JA, Savelkoul PH, et al. Prevalence and risk factors for carriage of ESBL-producing *Enterobacteriaceae* in Amsterdam. J Antimicrob Chemother. 2016;71:1076–82. http://dx.doi.org/10.1093/jac/dkv441
12. Peirano G, van der Bij AK, Freeman JL, Poirel L, Nordmann P, Costello M, et al. Characteristics of *Escherichia coli* sequence type 131 isolates that produce extended-spectrum β-lactamases: global distribution of the *H*30-Rx sublineage. Antimicrob Agents Chemother. 2014;58:3762–7. http://dx.doi.org/10.1128/AAC.02428-14
13. Micenková L, Sišková P, Bosák J, Jamborová I, Cernohorská L, Smajs D. Characterization of human uropathogenic ESBL-producing *Escherichia coli* in the Czech Republic: spread of CTX-M-27–producing strains in a university hospital. Microb Drug Resist. 2014;20:610–7. http://dx.doi.org/10.1089/mdr.2014.0013
14. Adler A, Gniadkowski M, Baraniak A, Izdebski R, Fiett J, Hryniewicz W, et al. Transmission dynamics of ESBL-producing *Escherichia coli* clones in rehabilitation wards at a tertiary care centre. Clin Microbiol Infect. 2012;18:E497–505. http://dx.doi.org/10.1111/j.1469-0691.2012.03999.x
15. Zhong YM, Liu WE, Liang XH, Li YM, Jian ZJ, Hawkey PM. Emergence and spread of O16–ST131 and O25b-ST131 clones among faecal CTX-M–producing *Escherichia coli* in healthy individuals in Hunan Province, China. J Antimicrob Chemother. 2015;70:2223–7. http://dx.doi.org/10.1093/jac/dkv114
16. Forde BM, Ben Zakour NL, Stanton-Cook M, Phan MD, Totsika M, Peters KM, et al. The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant *E. coli* O25b:H4–ST131 clone. PLoS ONE. 2014;9:e104400. http://dx.doi.org/10.1371/journal.pone.0104400
17. Andersen PS, Stegger M, Aziz M, Contente-Cuomo T, Gibbons HS, Keim P, et al. Complete genome sequence of the epidemic and highly virulent CTX-M-15–producing *H*30-Rx subclone of *Escherichia coli* ST131. Genome Announc. 2013;1:e00988–13. http://dx.doi.org/10.1128/genomeA.00988-13
18. Kutumbaka KK, Han S, Mategko J, Nadala C, Buser GL, Cassidy MP, et al. Draft denome sequence of bla_{NDM-1}-positive *Escherichia coli* O25b-ST131 clone isolated from an environmental sample. Genome Announc. 2014;2:e00462–14. http://dx.doi.org/10.1128/genomeA.00462-14
19. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3. http://dx.doi.org/10.1093/bioinformatics/btu033
20. Marttinen P, Hanage WP, Croucher NJ, Connor TR, Harris SR, Bentley SD, et al. Detection of recombination events in bacterial genomes from large population samples. Nucleic Acids Res. 2012;40:e6. http://dx.doi.org/10.1093/nar/gkr928
21. Phan MD, Forde BM, Peters KM, Sarkar S, Hancock S, Stanton-Cook M, et al. Molecular characterization of a multidrug resistance IncF plasmid from the globally disseminated *Escherichia coli* ST131 clone. PLoS ONE. 2015;10:e0122369. http://dx.doi.org/10.1371/journal.pone.0122369
22. Karczmarczyk M, Wang J, Leonard N, Fanning S. Complete nucleotide sequence of a conjugative IncF plasmid from an *Escherichia coli* isolate of equine origin containing bla_{CMY-2} within a novel genetic context. FEMS Microbiol Lett. 2014;352:123–7. http://dx.doi.org/10.1111/1574-6968.12364
23. Zhang H, Zhou Y, Guo S, Chang W. Prevalence and characteristics of extended-spectrum beta-lactamase (ESBL)–producing Enterobacteriaceae isolated from rural well water in Taian, China, 2014. Environ Sci Pollut Res Int. 2015;22:11488–92. http://dx.doi.org/10.1007/s11356-015-4387-9
24. Müller A, Stephan R, Nüesch-Inderbinen M. Distribution of virulence factors in ESBL-producing *Escherichia coli* isolated from the environment, livestock, food and humans. Sci Total Environ. 2016;541:667–72. http://dx.doi.org/10.1016/j.scitotenv.2015.09.135
25. Harada K, Nakai Y, Kataoka Y. Mechanisms of resistance to cephalosporin and emergence of O25b-ST131 clone harboring CTX-M-27 β-lactamase in extraintestinal pathogenic *Escherichia coli* from dogs and cats in Japan. Microbiol Immunol. 2012;56:480–5. http://dx.doi.org/10.1111/j.1348-0421.2012.00463.x
26. Mathers AJ, Peirano G, Pitout JD. The role of epidemic resistance plasmids and international high-risk clones in the spread of multidrug-resistant *Enterobacteriaceae*. Clin Microbiol Rev. 2015;28:565–91. http://dx.doi.org/10.1128/CMR.00116-14
27. Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, Sebra R, et al. Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. MBio. 2016;7:e02162–15.
28. Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, Khanh Nhu NT, Roberts LW, Stanton-Cook M, et al. Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. MBio. 2016;7:e00347–16.

Address for correspondence: Yasufumi Matsumura, Department of Clinical Laboratory Medicine, Kyoto University Graduate School of Medicine, 54 Shogoinkawahara-cho, Sakyo-ku, Kyoto 6068507, Japan; email: yazblood@kuhp.kyoto-u.ac.jp

Article DOI: http://dx.doi.org/10.3201/eid2211.160519

Global *Escherichia coli* Sequence Type 131 Clade with *bla*$_{CTX-M-27}$ Genes

Technical Appendix

Supplementary Methods

We used a core genome single-nucleotide polymorphism (SNP)–based approach to create a phylogenetic tree using the current standard procedure (*1*). SNPs were identified using raw read mapping followed by duplicate read removal, realignment, quality score recalibration, and variant filtering (*2*).

Core Genome Analysis

Reads from 53 isolates sequenced in this study and 4 isolates (S100EC, S107EC, S108EC, and S135EC) (*3*) underwent quality trimming using ERNE-FILTER (*4*). Trimmed reads were aligned against a reference genome of EC958 using Burrows-Wheeler Aligner (*5*). SNPs were called by using GATK Best Practices workflow (*6*) and SAMtools (coverage >10 and Phred-score >20) (*7*). The remaining 4 draft or complete genomes were aligned against EC958 by using ProgressiveMauve (*8*) to make EC958-like pseudo-chromosomes that contained only SNPs. The SNP-only core genome was identified as the common blocks of >500 bp to all 61 study isolates by using in-house Perl script. A maximum-likelihood tree was build using RAxML with GTR GAMMA substitution model and 100 rapid bootstrap replicates (*9*). We also separately analyzed the phylogeny of the sequence type (ST) 131 isolates excluding recombination sites. Bacterial recombination occurs more frequently than spontaneous mutations, and a phylogenetic tree that includes recombination sites could potentially distort phylogenetic inference (*10*), although this is not universally accepted as dogma (*11*). A recombination-free tree was also build by excluding recombination sites identified using a Bayesian analysis software BRATNextGen (*12*). A cutoff in the proportion of shared ancestry tree was chosen to enable separation of clades found in core genome-based tree. Twenty iterations of hidden Markov model parameter estimation were performed, and 100 permutations resampling was performed to determine the statistically significant recombination segments ($p<0.05$).

Genome Assembly

Trimmed reads were assembled by using Velvet and VelvetOptimizer (*13*) with k-mer values ranging from 31 to 73. The best assembly results in terms of the highest N50 value of each isolate underwent refinement of draft genomes using PAGIT (*14*).

Comparative Genomic Analysis

To define presence of genes and their alleles, we mapped trimmed reads to reference genes using SRST2 (*15*) and used BLAST+ (*16*) for draft or complete genomes. We used the following databases or typing schemes: ResFinder antimicrobial resistance gene database (*17*) VFDB (*18*) and VirulenceFinder (*19*) virulence gene databases, serotypeFinder O:H typing database (*20*), PlasmidFinder plasmid replicon database (*21*), MLST (http://mlst.ucc.ie/mlst/dbs/Ecoli), plasmid MLST (*21*), *fimH* typing (*22*), *gyrA*/*parC* typing (*22*), ST131 virotyping (*23*), and detection of *H*30Rx-specific *ybbW* SNP (*24*), plasmid addiction systems (*25*), and $bla_{\text{CTX-M}}$ genetic environment (*26*). Gegenees (*27*) was used to identify clade-specific segments among draft or complete genomes and visualized with EasyFig (*28*). BRIG (*29*) was used to visualize similarity of genomes to ST131 genomic islands (*30*) and to the ST131 reference plasmid pEC958 (*31*).

References

1. Robinson ER, Walker TM, Pallen MJ. Genomics and outbreak investigation: from sequence to consequence. Genome Med. 2013;5:36. PubMed http://dx.doi.org/10.1186/gm440

2. Olson ND, Lund SP, Colman RE, Foster JT, Sahl JW, Schupp JM, et al. Best practices for evaluating single nucleotide variant calling methods for microbial genomics. Front Genet. 2015;6:235. PubMed http://dx.doi.org/10.3389/fgene.2015.00235

3. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. Proc Natl Acad Sci U S A. 2014;111:5694–9. PubMed http://dx.doi.org/10.1073/pnas.1322678111

4. Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. An extensive evaluation of read trimming effects on Illumina NGS data analysis. PLoS ONE. 2013;8:e85024. PubMed http://dx.doi.org/10.1371/journal.pone.0085024

5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60. PubMed http://dx.doi.org/10.1093/bioinformatics/btp324

6. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303. PubMed http://dx.doi.org/10.1101/gr.107524.110

7. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9. PubMed http://dx.doi.org/10.1093/bioinformatics/btp352

8. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE. 2010;5:e11147. PubMed http://dx.doi.org/10.1371/journal.pone.0011147

9. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3. PubMed http://dx.doi.org/10.1093/bioinformatics/btu033

10. Rannala B, Yang Z. Phylogenetic inference using whole genomes. Annu Rev Genomics Hum Genet. 2008;9:217–31. PubMed http://dx.doi.org/10.1146/annurev.genom.9.081307.164407

11. Hedge J, Wilson DJ. Bacterial phylogenetic reconstruction from whole genomes is robust to recombination but demographic inference is not. MBio. 2014;5:e02158. PubMed

12. Marttinen P, Hanage WP, Croucher NJ, Connor TR, Harris SR, Bentley SD, et al. Detection of recombination events in bacterial genomes from large population samples. Nucleic Acids Res. 2012;40:e6. PubMed http://dx.doi.org/10.1093/nar/gkr928

13. Zerbino DR. Using the Velvet de novo assembler for short-read sequencing technologies. Curr Protoc Bioinformatics. 2010 Sep; Chapter 11:Unit 11.5. http://dx.doi.org/10.1002/0471250953.bi1105s31 **PMID: 20836074**

14. Swain MT, Tsai IJ, Assefa SA, Newbold C, Berriman M, Otto TD. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. Nat Protoc. 2012;7:1260–84. PubMed http://dx.doi.org/10.1038/nprot.2012.068

15. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2: rapid genomic surveillance for public health and hospital microbiology labs. Genome Med. 2014;6:90. PubMed http://dx.doi.org/10.1186/s13073-014-0090-6

16. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421. PubMed http://dx.doi.org/10.1186/1471-2105-10-421

17. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67:2640–4. PubMed http://dx.doi.org/10.1093/jac/dks261

18. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 2012;40:D641–5. PubMed http://dx.doi.org/10.1093/nar/gkr989

19. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol. 2014;52:1501–10. PubMed http://dx.doi.org/10.1128/JCM.03617-13

20. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol. 2015;53:2410–26. PubMed http://dx.doi.org/10.1128/JCM.00008-15

21. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58:3895–903. PubMed http://dx.doi.org/10.1128/AAC.02412-14

22. Johnson JR, Tchesnokova V, Johnston B, Clabots C, Roberts PL, Billig M, et al. Abrupt emergence of a single dominant multidrug-resistant strain of *Escherichia coli*. J Infect Dis. 2013;207:919–28. PubMed http://dx.doi.org/10.1093/infdis/jis933

23. Blanco J, Mora A, Mamani R, López C, Blanco M, Dahbi G, et al. 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. 2013;51:3358–67. PubMed http://dx.doi.org/10.1128/JCM.01555-13

24. Banerjee R, Robicsek A, Kuskowski MA, Porter S, Johnston BD, Sokurenko E, et al. Molecular epidemiology of *Escherichia coli* sequence type 131 and Its H30 and H30-Rx subclones among extended-spectrum-β-lactamase-positive and -negative *E. coli* clinical isolates from the Chicago region, 2007 to 2010. Antimicrob Agents Chemother. 2013;57:6385–8. PubMed http://dx.doi.org/10.1128/AAC.01604-13

25. Mnif B, Vimont S, Boyd A, Bourit E, Picard B, Branger C, et al. Molecular characterization of addiction systems of plasmids encoding extended-spectrum beta-lactamases in *Escherichia coli*. J Antimicrob Chemother. 2010;65:1599–603. PubMed http://dx.doi.org/10.1093/jac/dkq181

26. Matsumura Y, Johnson JR, Yamamoto M, Nagao M, Tanaka M, Takakura S, et al. CTX-M-27– and CTX-M-14–producing, ciprofloxacin-resistant *Escherichia coli* of the *H*30 subclonal group within ST131 drive a Japanese regional ESBL epidemic. J Antimicrob Chemother. 2015;70:1639–49. PubMed

27. Agren J, Sundström A, Håfström T, Segerman B. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS ONE. 2012;7:e39107. PubMed http://dx.doi.org/10.1371/journal.pone.0039107

28. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27:1009–10. PubMed http://dx.doi.org/10.1093/bioinformatics/btr039

29. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 2011;12:402. PubMed http://dx.doi.org/10.1186/1471-2164-12-402

30. Forde BM, Ben Zakour NL, Stanton-Cook M, Phan MD, Totsika M, Peters KM, et al. The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant *E. coli* O25b:H4–ST131 clone. PLoS ONE. 2014;9:e104400. PubMed http://dx.doi.org/10.1371/journal.pone.0104400

31. Phan MD, Forde BM, Peters KM, Sarkar S, Hancock S, Stanton-Cook M, et al. Molecular characterization of a multidrug resistance IncF plasmid from the globally disseminated *Escherichia coli* ST131 clone. PLoS ONE. 2015;10:e0122369. PubMed http://dx.doi.org/10.1371/journal.pone.0122369

32. Andersen PS, Stegger M, Aziz M, Contente-Cuomo T, Gibbons HS, Keim P, et al. Complete genome sequence of the epidemic and highly virulent CTX-M-15–producing H30-Rx subclone of *Escherichia coli* ST131. Genome Announc. 2013;1:e00988-13. PubMed http://dx.doi.org/10.1128/genomeA.00988-13

33. Peirano G, van der Bij AK, Freeman JL, Poirel L, Nordmann P, Costello M, et al. Characteristics of *Escherichia coli* sequence type 131 isolates that produce extended-spectrum β-lactamases: global distribution of the *H*30-Rx sublineage. Antimicrob Agents Chemother. 2014;58:3762–7. PubMed http://dx.doi.org/10.1128/AAC.02428-14

34. Peirano G, Bradford PA, Kazmierczak KM, Badal RE, Hackel M, Hoban DJ, et al. Global incidence of carbapenemase-producing *Escherichia coli* ST131. Emerg Infect Dis. 2014;20:1928–31. PubMed http://dx.doi.org/10.3201/eid2011.141388

35. McGann P, Snesrud E, Ong AC, Appalla L, Koren M, Kwak YI, et al. War wound treatment complications due to transfer of an IncN plasmid harboring *bla*$_{\text{OXA-181}}$ from Morganella morganii to CTX-M-27–producing sequence type 131 *Escherichia coli*. Antimicrob Agents Chemother. 2015;59:3556–62. PubMed http://dx.doi.org/10.1128/AAC.04442-14

36. Kutumbaka KK, Han S, Mategko J, Nadala C, Buser GL, Cassidy MP, et al. Draft genome sequence of *bla*$_{\text{NDM-1}}$-positive *Escherichia coli* O25b-ST131 clone isolated from an environmental sample. Genome Announc. 2014;2:e00462-14. PubMed http://dx.doi.org/10.1128/genomeA.00462-14

Technical Appendix Table 1. Strain information, mapping, and assembly statistics

							Mapping to EC958 genome		De novo assembly			
Strain	Year	Country (hospital, prefecture)	Location	Sample	*fimH* allele, *H*30Rx status	ESBL*	Sequencing depth	% Coverage	No. contigs	N50	Genome size	Reference
KUN2145	2007	Japan (a, Kyoto)	Community	Blood	*H*22	CTX-M-14	36.3	96.0	201	77275	5261498	
KFEC6	2004	Japan (b, Kyoto)	Unknown	Unknown	*H*22	CTX-M-2	24.1	96.4	239	70346	5290551	
KSEC7	2002	Japan (c, Kyoto)	Unknown	Unknown	*H*30	CTX-M-14	26.0	98.5	187	69032	5165720	
KKEC3	2005	Japan (g, Shiga)	Unknown	Unknown	*H*30	TEM-12	24.9	97.7	174	71954	5128144	
KUN5823	2009	Japan (a, Kyoto)	Hospital	Urine	*H*30	TEM-132	39.9	96.3	103	153130	5093967	
JJ1886	2008	United States	Community	Blood	*H*30Rx	CTX-M-15	NA*	NA	6†	NA	5308284	(*32*)
KCH27	2009	Japan (c, Kyoto)	Unknown	Unknown	*H*30Rx	CTX-M-14	42.1	98.7	189	86306	5227333	
KUN3842	2008	Japan (a, Kyoto)	Community	Urine	*H*30Rx	CTX-M-15	38.3	98.6	153	84772	5216555	
SI48	2012	Japan (e, Shiga)	Unknown	Urine	*H*30Rx	CTX-M-15	37.0	98.4	171	101146	5253375	
ONEC14	2006	Japan (f, Shiga)	Unknown	Unknown	*H*30Rx	CTX-M-15	62.3	98.3	130	141652	5290408	
ONEC29	2007	Japan (f, Shiga)	Unknown	Unknown	*H*30Rx	CTX-M-15	40.7	98.6	176	87726	5151579	
KT6	2012	Japan (r, Kyoto)	Unknown	Urine	*H*30Rx	CTX-M-15	32.6	97.1	178	76197	5112138	
BRG23	2014	Japan (u, Osaka)	Hospital	Urine	*H*30Rx	CTX-M-15	55.8	97.4	151	125485	5151404	
KP14	2010	Japan (b, Kyoto)	Unknown	Urine	*H*30Rx	CTX-M-14+15	38.9	98.6	208	93237	5226221	
KS58	2011	Japan (c, Kyoto)	Unknown	Urine	*H*30Rx	CTX-M-15	40.6	98.6	179	70831	5140718	
EC958	2005	United Kingdom	Community	Urine	*H*30Rx	CTX-M-15	NA	NA	2‡	NA	5245369	(*30*)
KUN5191	2009	Japan (a, Kyoto)	Community	Urine	*H*30Rx	CTX-M-15	37.8	99.9	135	129047	5170127	
Ec 58	2009	Canada	Hospital	Blood	*H*30Rx	CTX-M-15	84.4	100.0	116	311639	5299942	(*33*)
Ec 31	2009	Canada	Healthcare-associated	Blood	*H*30Rx	CTX-M-15	86.8	100.0	94	282486	5300483	(*33*)
KS121	2012	Japan (c, Kyoto)	Unknown	Urine	*H*30Rx	CTX-M-15	34.7	98.0	216	75907	5346990	
BRG221	2014	Japan (c, Kyoto)	Community	Urine	*H*30Rx	CTX-M-14+15	36.5	96.7	197	87770	5233600	
KP75	2011	Japan (b, Kyoto)	Hospital	Blood	*H*30Rx	CTX-M-15	36.9	98.0	213	96467	5310621	
KP46	2010	Japan (b, Kyoto)	Unknown	Urine	*H*30Rx	CTX-M-15	37.5	98.0	204	91282	5315779	
BRG151	2014	Japan (i, Shiga)	Community	Urine	*H*30	Negative	32.8	97.8	151	94973	5134331	
BRG274	2014	Japan (w, Osaka)	Hospital	Urine	*H*30	Negative	45.4	96.9	179	95884	5187623	
BRG54	2014	Japan (t, Hyogo)	Community	Blood	*H*30	Negative	34.9	97.4	193	83029	5070132	
SNEC5	2003	Japan (h, Shiga)	Unknown	Unknown	*H*30	CTX-M-14	29.3	97.5	277	64189	5232935	
KUN4389	2009	Japan (a, Kyoto)	Hospital	Urine	*H*30	CTX-M-14	41.0	97.9	207	96127	5243970	
S100EC§	2009	Australia	Unknown	Rectal swab	*H*30	CTX-M-27	51.1	98.5	97	243020	5153420	(*3*)
USA 14	2008	United States	Community	Urine	*H*30	CTX-M-14	70.9	98.3	114	220820	5220386	(*33*)
BRG62	2014	Japan (t, Hyogo)	Community	Urine	*H*30	CTX-M-14	52.0	98.9	159	102298	5158745	
KS46	2011	Japan (c, Kyoto)	Unknown	Urine	*H*30	CTX-M-14	40.8	97.7	207	84629	5116558	
KUN3273	2008	Japan (a, Kyoto)	Community	Urine	*H*30	CTX-M-14	38.6	97.6	181	105906	5128255	
ONEC7	2006	Japan (f, Shiga)	Unknown	Unknown	*H*30	CTX-M-14	33.3	98.1	220	80399	5196172	
KN94	2012	Japan (d, Kyoto)	Unknown	Urine	*H*30	CTX-M-14	48.8	97.5	131	119924	5223561	
KT37	2012	Japan (r, Kyoto)	Unknown	Urine	*H*30	CTX-M-14	24.2	97.8	289	47475	5125852	
S135EC§	2005	Canada	Community	Blood	*H*30	CTX-M-14	58.9	98.8	116	152032	5275688	(*3*)
FR 11	2008	France	Community	Urine	*H*30	CTX-M-14	36.7	97.6	106	169392	5131354	(*33*)
Ec# 584	2011	Vietnam	Unknown	Intraabdominal	*H*30	CTX-M-27	48.4	98.1	346	75552	5412329	(*34*)
ECNZ 35	2010	New Zealand	Hospital	Blood	*H*30	CTX-M-14	63.9	98.4	115	159997	5309674	(*33*)
EcSA01	2008	South Africa	Community	Urine	*H*30	CTX-M-14	68.3	97.5	80	194881	5198654	(*33*)
Ec 32	2009	Canada	Community	Blood	*H*30	CTX-M-14	61.5	98.3	174	184369	5352796	(*33*)

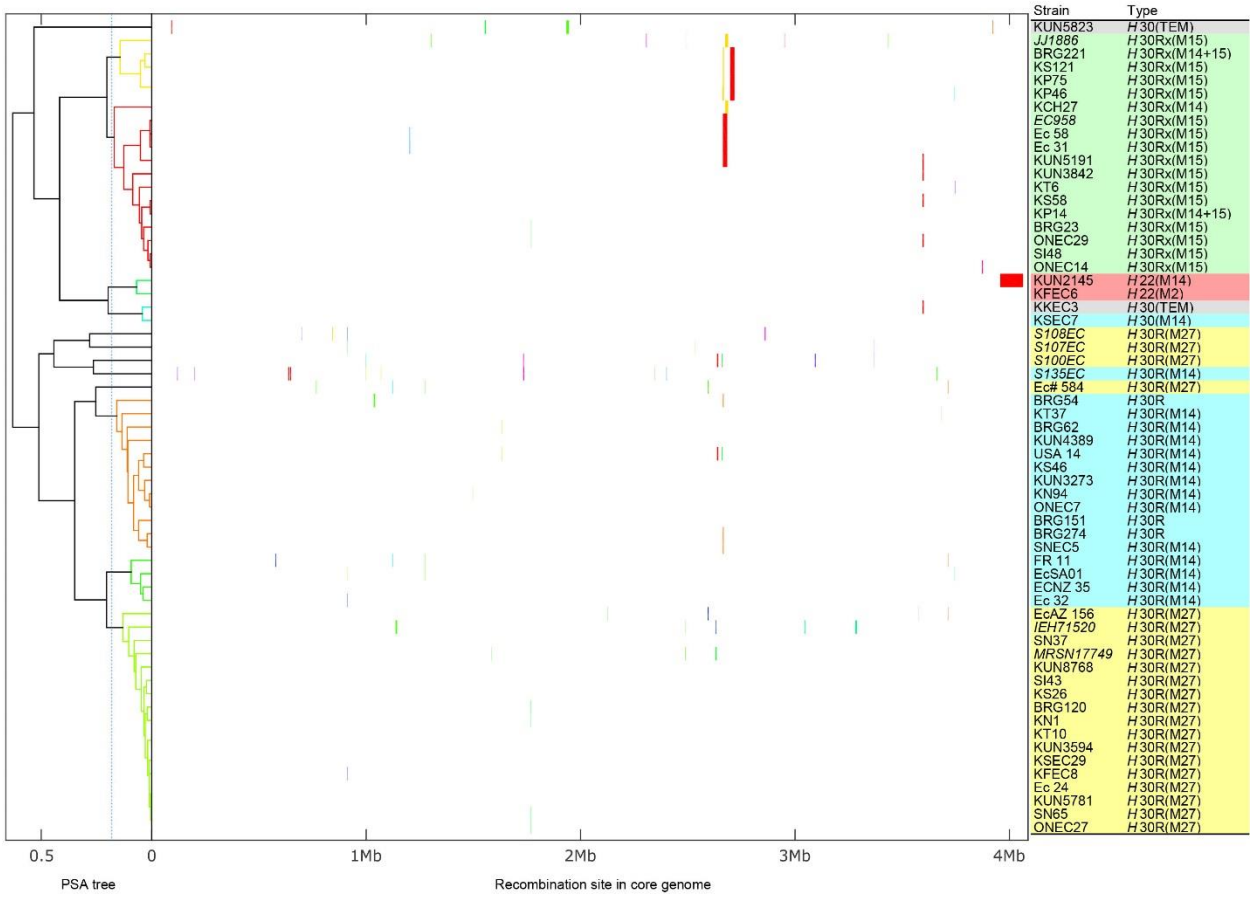
Strain	Year	Country (hospital, prefecture)	Location	Sample	*fimH* allele, *H*30Rx status	ESBL*	Mapping to EC958 genome		De novo assembly			Reference
							Sequencing depth	% Coverage	No. contigs	N50	Genome size	
KUN8768	2011	Japan (a, Kyoto)	Community	Urine	*H*30	CTX-M-27	35.5	97.5	170	93073	5184482	
SN37	2010	Japan (h, Shiga)	Unknown	Sputum	*H*30	CTX-M-27	35.3	96.3	164	79492	5039880	
SI43	2012	Japan (e, Shiga)	Unknown	Urine	*H*30	CTX-M-27	38.5	97.2	171	82979	5017470	
KT10	2012	Japan (r, Kyoto)	Unknown	Blood	*H*30	CTX-M-27	53.3	95.0	119	133403	4954097	
KUN3594	2008	Japan (a, Kyoto)	Hospital	Urine	*H*30	CTX-M-27	35.8	97.1	124	93754	5021116	
KFEC8	2004	Japan (b, Kyoto)	Unknown	Unknown	*H*30	CTX-M-27	32.2	97.1	235	55958	5097618	
S107EC§	2010	Australia	Unknown	Urine	*H*30	CTX-M-27	46.8	97.1	97	191225	5092957	(*3*)
S108EC§	2009	Australia	Unknown	Blood	*H*30	CTX-M-27	61.9	97.4	95	192487	5121514	(*3*)
KSEC29	2006	Japan (c, Kyoto)	Unknown	Unknown	*H*30	CTX-M-27	39.3	97.3	138	102764	5064236	
KN1	2010	Japan (d, Kyoto)	Unknown	Urine	*H*30	CTX-M-27	46.2	97.3	141	120021	5179897	
ONEC27	2007	Japan (f, Shiga)	Unknown	Unknown	*H*30	CTX-M-27	37.8	97.2	190	76207	5046307	
BRG120	2014	Japan (s, Aichi)	Hospital	Sputum	*H*30	CTX-M-27	41.3	97.6	124	140794	5094906	
SN65	2011	Japan (h, Shiga)	Unknown	Pus	*H*30	CTX-M-27	47.0	98.5	148	124167	5105443	
EcAZ 156	2013	Thailand	Unknown	Urine	*H*30	CTX-M-27	54.0	97.2	128	159590	5130003	(*34*)
KS26	2010	Japan (c, Kyoto)	Unknown	Urine	*H*30	CTX-M-27	51.3	97.2	162	104636	5105370	
MRSN17749¶	2013	United States	Hospital	Groin swab	*H*30	CTX-M-27	NA	NA	92	191197	5046460	(*35*)
IEH71520¶	2014	United States	House environment	Vacuum cleaner dust	*H*30	CTX-M-27	NA	NA	202	67135	5153432	(*36*)
Ec 24	2008	Canada	Hospital	Blood	*H*30	CTX-M-27	84.9	97.3	77	216849	5077997	(*33*)
KUN5781	2009	Japan (a, Kyoto)	Hospital	Blood	*H*30	CTX-M-27	36.3	97.2	147	116898	5066641	

*ESBL, extended-spectrum β-lactamase; NA, not applicable.
†Chromosome and 5 plasmids.
‡Chromosome and 1 plasmid.
§Short reads were mapped and assembled using the same methods as our sequenced isolates.
¶Draft genome.

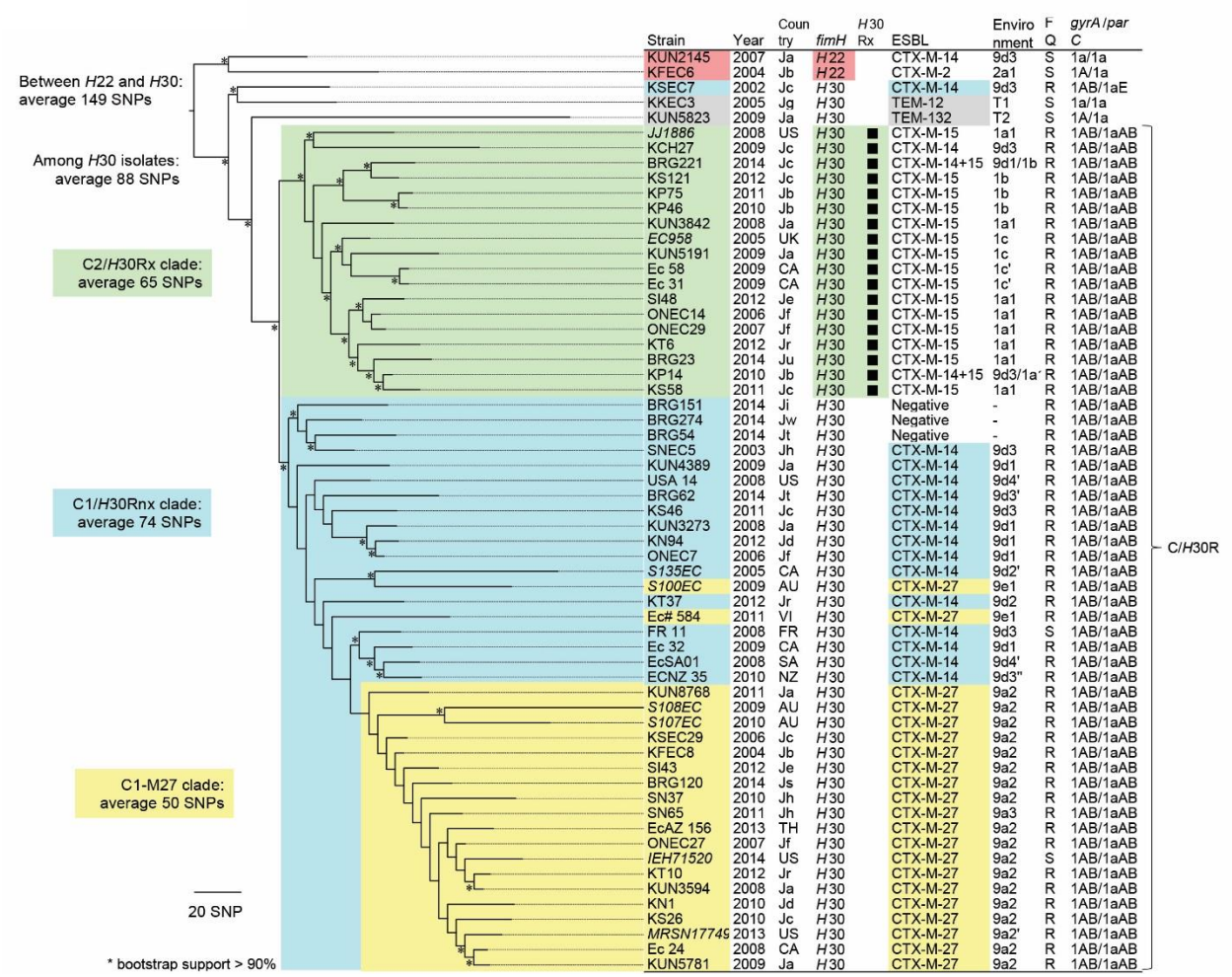
Technical Appendix Table 2. Genetic structures that flank ESBL genes*

ESBL type: group, subtype	Genetic structure: Upstream (bp)	Genetic structure: Downstream (bp)	Genetic structure: Type in Figure 1 and Technical Appendix Figure 2†	Genetic structure: GenBank accession no.	Clade in the tree or group, number of isolates: C/*H*30R; C1/*H*30R; C1-M27, n = 19	C1/*H*30R; Others, n = 16	C2/*H*30Rx, n = 18	Other than C/*H*30R, n = 5
bla$_{CTX-M-9}$ group								
bla$_{CTX-M-27}$	IS*26*-ΔIS*Ecp1* (208)	ΔIS*903D* (391)-IS*26*	9a2, 9a2′‡	AB976590	18[b]			
	IS*26*-ΔIS*Ecp1* (208)	ΔIS*903D* (226)-IS*26*	9a3	AB985520	1			
	IS*26*-ΔIS*Ecp1* (388)	IS*903D*	9e1	LC091535§		2		
bla$_{CTX-M-14}$	IS*Ecp1*	ΔIS*903D*	9d3, 9d3′¶, 9d3″#	AB976599, LC091534¶, LC107627[#]		6¶#	2**	2
	IS*Ecp1*	IS*903D*	9d1	AB976598		4	1[g]	
	ΔIS*Ecp1*	ΔIS*903D*	9d2, 9d2′‡‡	AB976605		2		
	ΔIS*Ecp1*	IS*903D*	9d4′§§	AB976604		2		
bla$_{CTX-M-1}$: *bla*$_{CTX-M-15}$	IS*Ecp1*	*orf477*	1a1	AB976566			9**	
	IS*26*-ΔIS*Ecp1* (497)	*orf477*	1b	AB976569			4††	
	IS*26*-ΔIS*Ecp1* (24)	*orf477*	1c, 1c′¶¶	AB976574, LC107628[j]			4	
bla$_{CTX-M-2}$: *bla*$_{CTX-M-2}$	IS*Ecp1*	Downstream *bla*$_{KLUA}$##	2a1	AB976588				1
bla$_{TEM}$								
bla$_{TEM-12}$	Tn*2*	Tn*2*	T1	LC091536§				1
bla$_{TEM-132}$	Tn*2*	Tn*2*	T2	LC091537§				1

*ESBL, extended-spectrum β-lactamase.
†The classification and numbering of the structures follows these in our previous publication (*26*).
‡One isolate from another study (MRSN17749) had a contig of ΔIS*Ecp1* (208bp)-*bla*$_{CTX-M-27}$-ΔIS*903D* (391bp) without the IS*26* flanking structure. However, the lengths of the truncated IS*Ecp1* and IS*903D* structures suggest the isolate had the 9a2 structure.
§New sequence found in this study (no identical sequence deposited in GenBank).
¶One isolate (BRG62) had the 9d3′ structure, a variant of 9d3. The only difference between 9d3 and 9d3□ is one nucleotide (1 aa) change in *tnpA* of ΔIS*903D*.
#One isolate (ECNZ 35) had the 9d3” structure, a variant of 9d3. The 9d3” structure has 1 nt change (synonymous substitution) in *bla*$_{CTX-M-14}$. The isolate may have IS*Ecp1*-*bla*$_{CTX-M-14}$-IS*903D* structure because the *bla*$_{CTX-M-14}$-containing contig included 5′ truncated IS*903D* but remaining sequence of IS*903D* was found in another contig.
**One isolate was positive for both 9d1 and 1b.
††One isolate was positive for both 9d3 and 1a1.
‡‡One isolate (S135EC) may have IS*Ecp1*-*bla*$_{CTX-M-14}$-ΔIS*903D* structure because the *bla*$_{CTX-M-14}$-containing contig included 3′ truncated IS*Ecp1* but remaining sequence of IS*Ecp1* was found in another contig.
§§These 2 isolates (USA 14 and EcSA01) may have IS*Ecp1*-*bla*$_{CTX-M-14}$-IS*903D* (9d1) structure because the *bla*$_{CTX-M-14}$-containing contig included 5′ truncated IS*903D* but remaining sequence of IS*903D* was found in another contig.
¶¶One isolate (BRG62) had the 9d3” structure, a variant of 9d3. The only difference between 9d3 and 9d3” is one nucleotide (1 aa) change in *tnpA* of ΔIS*903D*.
##The nucleotide sequence was identical to the region between *kluA-1* and *orf3* of *Kluyvera ascorbata* (GenBank accession no. AJ272538).

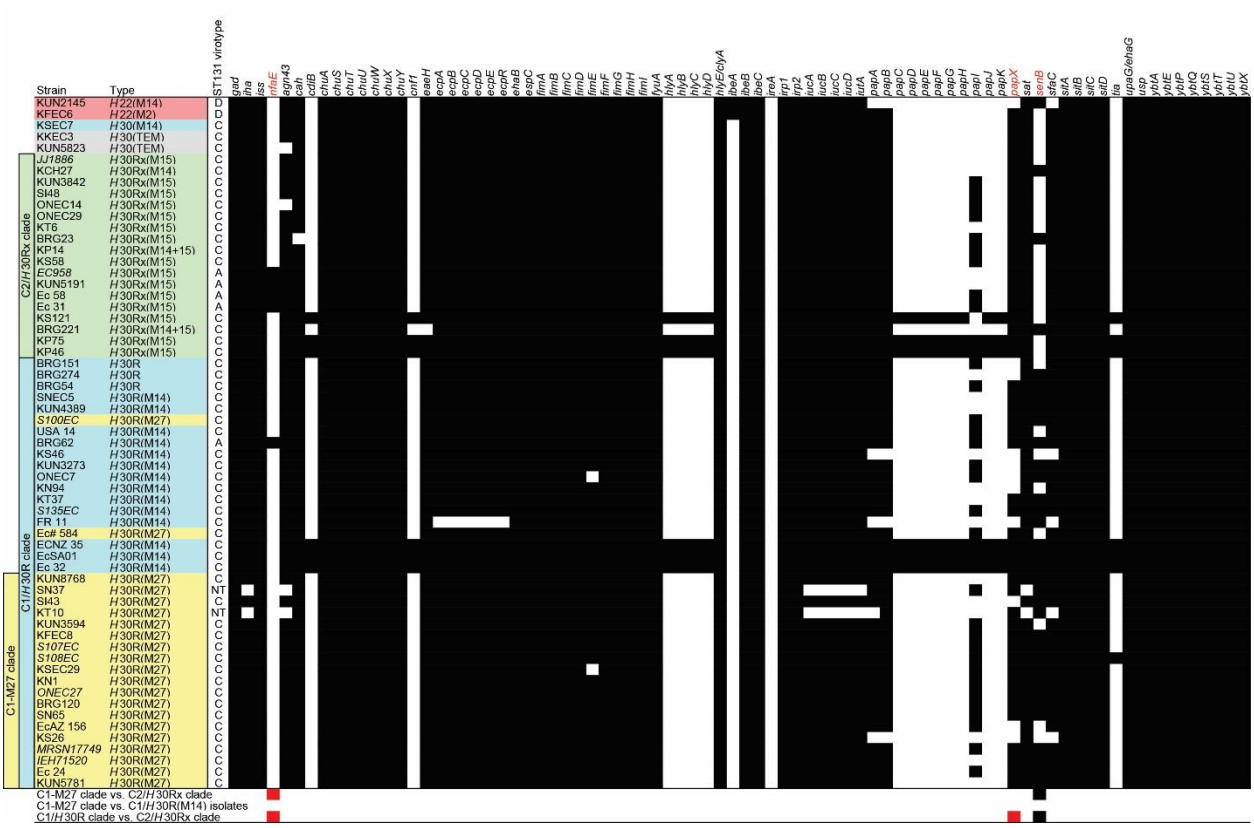


Technical Appendix Figure 1. Recombinant regions identified by BRATNextGen. The same core genome used for construction of the single-nucleotide polymorphism–based phylogenetic tree (Figure 1 in main text) was used for the analysis. The tree in the left is a proportion of shared ancestry tree. A cutoff value of 0.15 was chosen to form clusters of the C1-M27 clade, C1/*H*30R isolates other than those of the C1-M27 clade, and C2/*H*30Rx clade. The strain names and types are colored as same as those in Figure 1. ESBL types are indicated in parentheses of Type column. The middle panel shows a horizontal representation of the recombinant segments using color bars. Segments of the same color and the same column derived from the same origin. A total of 79 segments (304,782 bp) including 3,453 SNPs were associated with recombination.

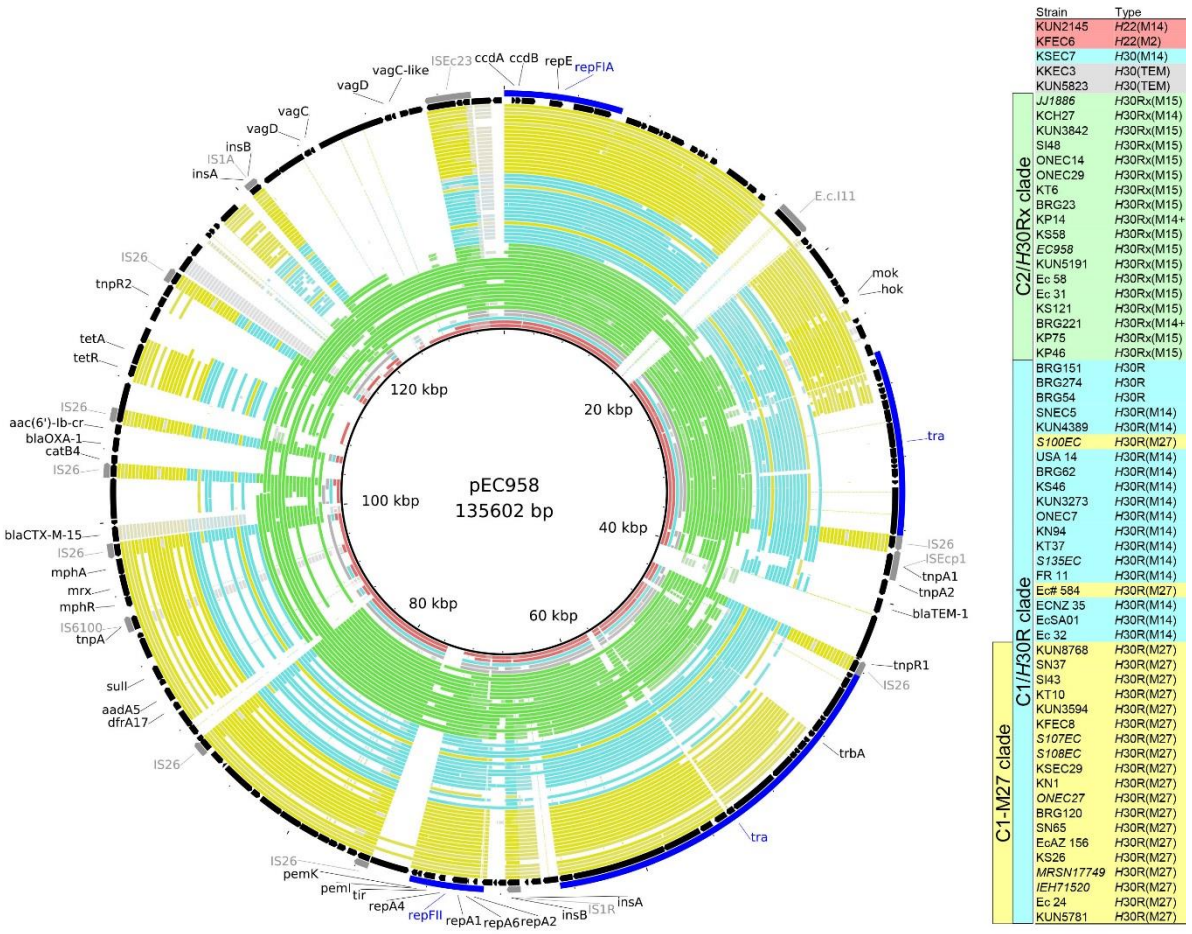


Strain	Year	Country	fimH	H30Rx	ESBL	Environment	FQ	gyrA/parC
KUN2145	2007	Ja	H22		CTX-M-14	9d3	S	1a/1a
KFEC6	2004	Jb	H22		CTX-M-2	2a1	S	1A/1a
KSEC7	2002	Jc	H30		CTX-M-14	9d3	R	1AB/1aE
KKEC3	2005	Jg	H30		TEM-12	T1	S	1a/1a
KUN5823	2009	Ja	H30		TEM-132	T2	S	1A/1a
JJ1886	2008	US	H30	■	CTX-M-15	1a1	R	1AB/1aAB
KCH27	2009	Jc	H30	■	CTX-M-14	9d3	R	1AB/1aAB
BRG221	2014	Jc	H30	■	CTX-M-14+15	9d1/1b	R	1AB/1aAB
KS121	2012	Jc	H30	■	CTX-M-15	1b	R	1AB/1aAB
KP75	2011	Jb	H30	■	CTX-M-15	1b	R	1AB/1aAB
KP46	2010	Jb	H30	■	CTX-M-15	1b	R	1AB/1aAB
KUN3842	2008	Ja	H30	■	CTX-M-15	1a1	R	1AB/1aAB
EC958	2005	UK	H30	■	CTX-M-15	1c	R	1AB/1aAB
KUN5191	2009	Ja	H30	■	CTX-M-15	1c	R	1AB/1aAB
Ec 58	2009	CA	H30	■	CTX-M-15	1c'	R	1AB/1aAB
Ec 31	2009	CA	H30	■	CTX-M-15	1c'	R	1AB/1aAB
SI48	2012	Je	H30	■	CTX-M-15	1a1	R	1AB/1aAB
ONEC14	2006	Jf	H30	■	CTX-M-15	1a1	R	1AB/1aAB
ONEC29	2007	Jf	H30	■	CTX-M-15	1a1	R	1AB/1aAB
KT6	2012	Jr	H30	■	CTX-M-15	1a1	R	1AB/1aAB
BRG23	2014	Ju	H30	■	CTX-M-15	1a1	R	1AB/1aAB
KP14	2010	Jb	H30	■	CTX-M-14+15	9d3/1a'	R	1AB/1aAB
KS58	2011	Jc	H30	■	CTX-M-15	1a1	R	1AB/1aAB
BRG151	2014	Ji	H30		Negative	-	R	1AB/1aAB
BRG274	2014	Jw	H30		Negative	-	R	1AB/1aAB
BRG54	2014	Jt	H30		Negative	-	R	1AB/1aAB
SNEC5	2003	Jh	H30		CTX-M-14	9d3	R	1AB/1aAB
KUN4389	2009	Ja	H30		CTX-M-14	9d1	R	1AB/1aAB
USA 14	2008	US	H30		CTX-M-14	9d4'	R	1AB/1aAB
BRG62	2014	Jt	H30		CTX-M-14	9d3'	R	1AB/1aAB
KS46	2011	Jc	H30		CTX-M-14	9d3	R	1AB/1aAB
KUN3273	2008	Ja	H30		CTX-M-14	9d1	R	1AB/1aAB
KN94	2012	Jd	H30		CTX-M-14	9d1	R	1AB/1aAB
ONEC7	2006	Jf	H30		CTX-M-14	9d1	R	1AB/1aAB
S135EC	2005	CA	H30		CTX-M-14	9d2'	R	1AB/1aAB
S100EC	2009	AU	H30		CTX-M-27	9e1	R	1AB/1aAB
KT37	2012	Jr	H30		CTX-M-14	9d2	R	1AB/1aAB
Ec# 584	2011	VI	H30		CTX-M-27	9e1	R	1AB/1aAB
FR 11	2008	FR	H30		CTX-M-14	9d3	S	1AB/1aAB
Ec 32	2009	CA	H30		CTX-M-14	9d1	R	1AB/1aAB
EcSA01	2008	SA	H30		CTX-M-14	9d4'	R	1AB/1aAB
ECNZ 35	2010	NZ	H30		CTX-M-14	9d3"	R	1AB/1aAB
KUN8768	2011	Ja	H30		CTX-M-27	9a2	R	1AB/1aAB
S108EC	2009	AU	H30		CTX-M-27	9a2	R	1AB/1aAB
S107EC	2010	AU	H30		CTX-M-27	9a2	R	1AB/1aAB
KSEC29	2006	Jc	H30		CTX-M-27	9a2	R	1AB/1aAB
KFEC8	2004	Jb	H30		CTX-M-27	9a2	R	1AB/1aAB
SI43	2012	Je	H30		CTX-M-27	9a2	R	1AB/1aAB
BRG120	2014	Js	H30		CTX-M-27	9a2	R	1AB/1aAB
SN37	2010	Jh	H30		CTX-M-27	9a2	R	1AB/1aAB
SN65	2011	Jh	H30		CTX-M-27	9a3	R	1AB/1aAB
EcAZ 156	2013	TH	H30		CTX-M-27	9a2	R	1AB/1aAB
ONEC27	2007	Jf	H30		CTX-M-27	9a2	R	1AB/1aAB
IEH71520	2014	US	H30		CTX-M-27	9a2	S	1AB/1aAB
KT10	2012	Jr	H30		CTX-M-27	9a2	R	1AB/1aAB
KUN3594	2008	Ja	H30		CTX-M-27	9a2	R	1AB/1aAB
KN1	2010	Jd	H30		CTX-M-27	9a2	R	1AB/1aAB
KS26	2010	Jc	H30		CTX-M-27	9a2	R	1AB/1aAB
MRSN17749	2013	US	H30		CTX-M-27	9a2'	R	1AB/1aAB
Ec 24	2008	CA	H30		CTX-M-27	9a2	R	1AB/1aAB
KUN5781	2009	Ja	H30		CTX-M-27	9a2	R	1AB/1aAB

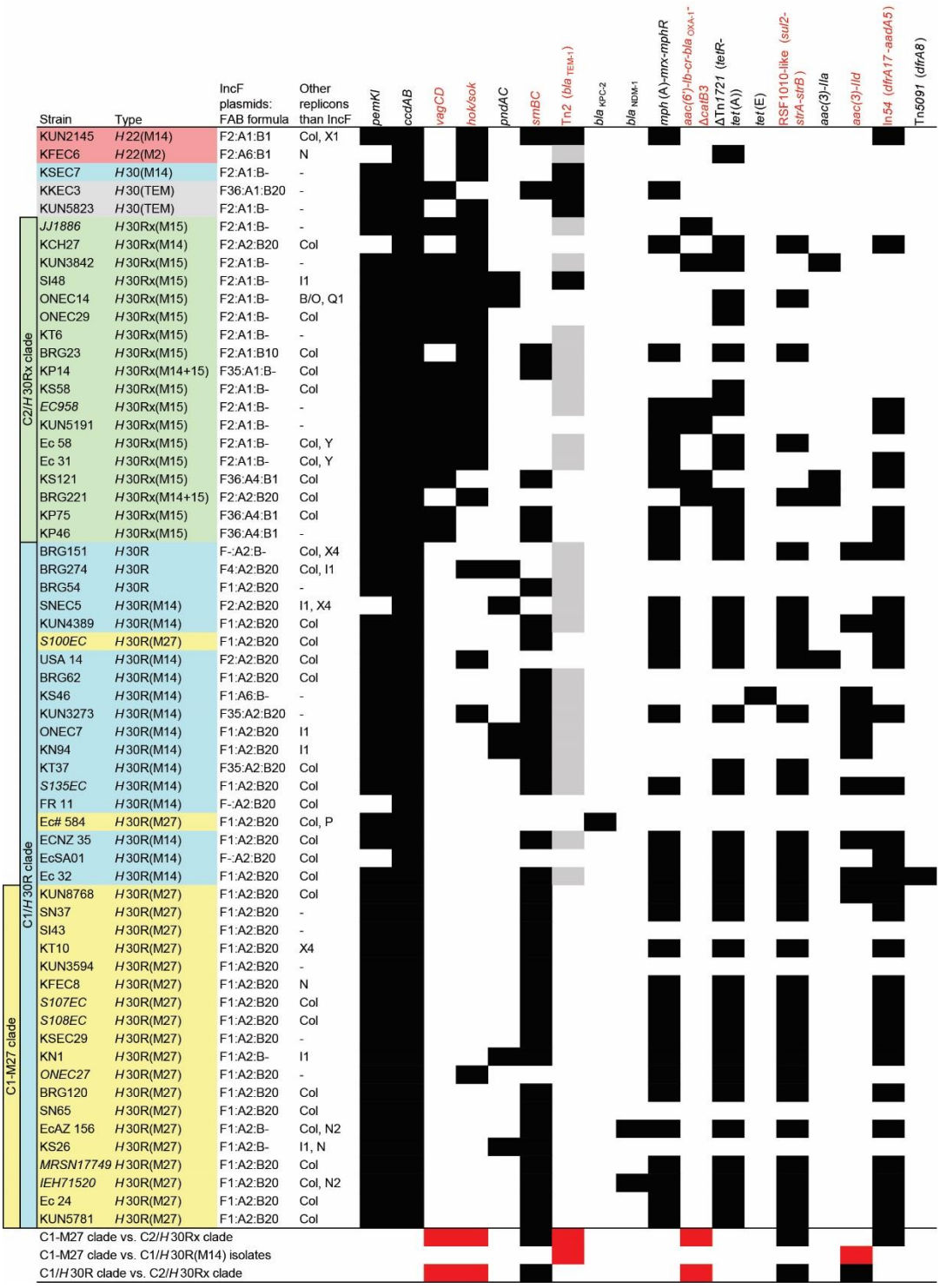
Technical Appendix Figure 2. Phylogenetic tree build from recombination-free core genome. This maximum-likelihood phylogram is based on a 3,781,868-bp recombination-free core genome and a total of 1,827 single-nucleotide polymorphisms. The tree is rooted by using the outgroup *H*22 isolates and asterisks indicates bootstrap support >90% from 100 replicates. The clustering results were as same as the tree built from the whole core genome shown in Figure 1. The ciprofloxacin-resistant C/*H*30R cluster comprised the C2/*H*30Rx and C1/*H*30R clades. All of the *H*30Rx isolates belonged to the C2/*H*30Rx clade. The C1/*H*30R clade included CTX-M-14-produing *H*30R, non-ESBL–producing *H*30R, and CTX-M-27–producing *H*30R isolates. CTX-M-27–producing isolates belonged to the C1-M27 clade within the C1/*H*30R clade except two isolates (S100EC and Ec #584). The bootstrap value for the root of the M27 clade was 76%.



Technical Appendix Figure 3. *Escherichia coli* sequence type (ST) 131 virotypes and virulence genes. Black squares indicate presence of each gene. Results of statistical tests for gene prevalence comparison between clades are shown at the bottom rows; black indicates high prevalence of the former clade, and red indicates high prevalence of the latter clade. ST131 virotype C was prevalent in common. Virotype NT indicates nontypeable. The C1-M27 clade isolates more frequently had *senB* enterotoxin gene than C2/*H*30Rx isolates but the other C1/*H*30R isolates also frequently had it. Two genes (*nfaE* and *papX*) were prevalent in the C2/*H*30Rx clade than the C1/*H*30R clade.



Technical Appendix Figure 4. Comparison of genomes of *Escherichia coli* sequence type (ST) 131 isolates with the pEC958 plasmid of CTX-M-15–producing ST131 C2/*H*30Rx reference strain EC958. Rings drawn by BRIG show the presence of the pEC958-like regions and colored according to colors in Figure 1. Colored segments indicate >90% similarity, and gray segments indicate >70% similarity by BLAST comparison between the regions of interest and each genome. The C1-M27 clade lacked the first part of the transfer regions (*tra)*. Some regions common to both C2/*H*30Rx and C1/*H*30R clades are present, but the presence or absence of other regions are divergent even within the same clade. The presence of resistance genes is also shown in Technical Appendix Figure 5.



Technical Appendix Figure 5. Plasmid replicons, plasmid addiction systems, and antimicrobial resistance genes of extraintestinal pathogenic *Escherichia coli*. Black indicates presence of each gene. Gray area of Tn*2* column indicates truncated Tn*2*. Results of statistical tests for gene prevalence comparison between

clades are shown at the bottom rows; black indicates high prevalence of the former clade and red indicates high prevalence of the latter clade. F1:A2:B20 IncF plasmids were prevalent in the C1/*H*30R clade while F2:A1:B- plasmids were prevalent in the C2/*H*30Rx clade. Three CTX-M-14–producing C2/*H*30Rx isolates had mixture types of replicons from the CTX-M-15–producing C2/*H*30Rx and CTX-M-14–produing C1/*H*30R isolates. Only C2/*H*30Rx isolates had *vagCD* plasmid addiction system, *aac(6')-Ib-cr-bla*$_{OXA-1}$-Δ*catB3* resistance gene set. C2/*H*30Rx isolates more frequently had *vagCD* and *hok/sok* plasmid addiction systems than the C1/*H*30R isolates. *srnBC* plasmid addiction system and *sul2-strA-strB* resistance gene set originally found in RSF1010 plasmid were more frequently found in the C1/*H*30R isolates than the C2/*H*30Rx isolates. None of the C1-M27 clade isolates had Tn*2* (*bla*$_{TEM-1}$). Class 1 integron In*54* (*dfrA17-aadA5*) was more frequently found in the C1-M27 clade isolates than the C2/*H*30Rx isolates. CTX-M-14–producing C1/*H*30R isolates more frequently had *aac(3)-IId* than the C1-M27 clade or C2/*H*30Rx isolates. Two C1-M27 isolates carried *bla*$_{NDM-1}$ on IncN2 plasmid backbone and ΔIS*Aba125*-IS*Ec33*-Δ*ISAba125-bla*$_{NDM-1}$-*ble*$_{MBL}$-Δ*trpF*-IS*Sen4*-Tn*5403* structure and 1 CTX-M-14–producing C1/*H*30R isolate had *bla*$_{KPC-2}$ in IS*Kpn27*-Δ*bla*$_{TEM-1}$-*bla*$_{KPC-2}$-Δ*traN-korC-klcA* structure.