Title:
Genetic features of persistently antimicrobial drug-susceptible extraintestinal pathogenic *Escherichia coli* pandemic sequence type 95

Yuan Hu Allegretti, Reina Yamaji, Sheila Adams-Sapper, Lee W Riley

Keywords: *Escherichia coli*, ExPEC, ST95, drug resistance, persistently drug-susceptible, microbial genome-wide association study

Abstract
Extraintestinal pathogenic *Escherichia coli* (ExPEC) belonging to multilocus sequence type 95 (ST95) is one of the most widespread ExPEC lineages associated with bloodstream infections (BSI) and urinary tract infections (UTI). In contrast to other widespread ExPEC sequence types, a large proportion ST95 strains remains susceptible to all antimicrobial agents used to treat BSI or UTI. We aimed to identify genomic features of ST95 associated with persistent drug susceptibility. We conducted a genome-wide association study of 80 ExPEC ST95 isolates from patients with BSI or UTI in Northern California, and 1669 ST95 isolates deposited in the Enterobase database. Of the total of 1749 ST95 isolates, we compared whole-genome sequences of 887 drug-susceptible strains and 862 strains resistant to one or more drugs (defined genotypically as strains harboring drug-resistance genes annotated in the ResFinder database) to identify genetic features associated with strains devoid of drug-resistance genes. By a genome-wide association study of 553 UTI and BSI ST95 isolates, we found 44 accessory genes significantly associated with drug susceptibility, six of which encoded hypothetical proteins. Fifteen of these were not found in any of the WGSs of ST131 ExPEC strains, which are frequently multidrug-resistant, and eight of these genes were annotated to encode transporter or transfer systems. These findings highlight the potential mechanisms by which ST95 strains may resist the acquisition of mobile DNA elements carrying drug-resistance genes.

Importance
Despite the increasing prevalence of antibiotic resistance in *Escherichia coli* strains that cause common infections such as urinary tract and bloodstream infections, a major pandemic lineage of extraintestinal pathogenic *E. coli* ST95 has remained susceptible to most drugs used to treat these infections. We compared genomes of 1749 ST95 isolates to identify genetic features that may explain why most strains of ST95 resist becoming drug-resistant. Identification of such genomic features could contribute to the development of novel strategies to prevent the spread of antibiotic-resistant genes and devise new measures to control antibiotic-resistant infections.

1. Introduction

Antimicrobial drug resistance is one of the most serious global public health concerns of our time, threatening the effective prevention and treatment of infectious diseases (1, 2). In particular, several drug-resistant Gram-negative bacterial species belonging to Enterobacteriaceae have come to be designated by the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) to be "urgent-threat" or "priority 1, critical" pathogens in need of new treatment and research (2, 3). Among Enterobacteriaceae, *Escherichia coli* (*E. coli*) is the most frequent cause of common extraintestinal infections such as urinary tract infections (UTI) and bloodstream infections (BSI) (4). They are referred to as extraintestinal pathogenic *E. coli* (ExPEC) (5). The prevalence and incidence of infections caused by ExPEC strains resistant to a wide spectrum of antimicrobial agents, including extended-spectrum beta-lactam (ESBL) drugs, fluoroquinolones, trimethoprim-sulfamethoxazole, and aminoglycosides have been rapidly increasing worldwide (6–12).

Epidemiological studies have shown that large proportions of community and healthcare-associated UTI and BSI reported globally are caused by ExPEC strains belonging to a limited set of lineages as defined by multilocus sequence typing (MLST) (13, 14). These include sequence types 131 (ST131), ST95, ST69, ST73, and ST393, which are responsible for nearly half of all *E. coli* UTIs or BSIs reported from many regions of the world (15). Most of these lineages except for ST95 have a high frequency of drug resistance. In particular, a large proportion of ST131, the most commonly reported ExPEC lineage worldwide, is multidrug-resistant (16–18).

Interestingly, despite its high prevalence, several recent studies have reported that ST95 strains remain susceptible to most of the antimicrobial drugs used to treat UTI and BSI (14, 19–21). We previously reported that the carriage of a UTI89-like plasmid called pUTI89* and *fimH* type is associated with the pan-susceptibility phenotype of ST95 (20, 21). However, the particular genomic features associated with this persistently drug-susceptible phenotype remained unclear. Here, we compared whole-genome sequences (WGS) of susceptible ST95 strains to those of ST95 strains resistant to at least one drug to identify genetic features associated with persistent drug susceptibility. We hypothesize that persistently susceptible ST95 strains restrict the acquisition or maintenance of mobile DNA elements that carry drug-resistance genes, but they nevertheless may have a compensatory fitness advantage that determines their successful global dissemination.

2. Results

2.1 Frequency and type of drug resistance genes among ST95 strains

We performed WGS analysis of 36 ST95 isolates that were previously collected from UTI patients in Northern California (14) and compared the sequences to those of 44 BSI ST95 isolates from Northern California (22) we previously reported, and 1669 ST95 WGSs deposited in the Enterobase database (23). We queried the carriage of recognized drug resistance genes against ResFinder and found that 837 (50%) of 1669 Enterobase ST95 isolates and 50 (62%) of 80 Northern California isolates did not contain any recognized resistance genes. Among the remaining 862 ST95 isolates, 112 recognized drug-resistant genes for 12 different classes of antibiotics were identified (Supplementary Table 2). Of these, 25 genes for 7 different classes of antibiotics were present in 80 clinical ST95 isolates from Northern California (Table 1). These classes of antimicrobial agents included aminoglycosides, beta-lactams, phenicols, trimethoprim, sulfonamide, tetracycline, and macrolides. Genes associated with fosfomycin, lincosamide, colistin, quinolone, and rifampicin resistance were not observed among any of the ST95 isolates from Northern California. In the period when the total annual submission of ST95 isolates to Enterobase was more than 50 (2010-2017), the prevalence of drug-resistant genes has remained relatively constant (33%-59%). When stratified by BSI or UTI, 86 (38%) of 225 BSI isolates and 132 (53%) of 248 UTI isolates were drug-resistant based on the ResFinder database.

Among the sequences from the 80 Northern California isolates, we found extended-spectrum beta-lactamase (ESBL) genes $bla_{CTX\text{-}M\text{-}14}$, $bla_{CTX\text{-}M\text{-}15}$, $bla_{TEM\text{-}1a}$, $bla_{TEM\text{-}1b}$, and $bla_{TEM\text{-}1c}$ in 1 (1.3 %), 1 (1.3 %), 3 (3.9 %), 21 (27 %), and 1 (1.3 %) of 80 isolates, respectively. The same set of ESBL genes was present in 29 (1.7 %), 44 (2.6 %), 18 (1.1 %), 335 (20.1 %), and 148 (8.9 %) of 1669 isolates among the sequences deposited in the Enterobase database, respectively. The temporal distribution of drug resistance genes in all the 1749 ST95 isolates is shown in Table 2.

2.2 Pan-genome analysis

A total of 35,134 genes constituted the pan-genome of 1749 ST95 WGSs. Of 35,134 genes, 3,688 (10.5%) were shared among more than 95% of the isolates (core genes) and 31,446 (89.5%) were distributed among subsets of the isolates (accessory genes). Of the latter, 29,522 genes were found in <15% of the isolates (cloud genes). Of the cloud genes, 13,970 (40%) genes were unique to only one isolate.

2.3 Phylogenetic and principal component analyses

Maximum-likelihood phylogenies of the SNP alignment of core genomes and the presence and absence of accessory genomes were obtained from FastTree (24) with the GTR model (Figure 1). Clustering based on *fimH* types was observed but we did not observe a strong correlation of drug resistance with phylogenetic trees based on the SNPs of core genomes and the presence and absence of accessory genes.

Principal component analysis (PCA) was also conducted to visually assess variables accounting for the variance. Cluster patterns based on the variance of SNPs and accessory genes were associated with the carriage of plasmid pUTI89* and with *fimH* types 18, 27, 30, 41, and 54 in core genomes and accessory genomes, respectively (Figure 2). However, the cluster patterns were not associated with drug resistance or pan-susceptibility.

2.4 Genes associated with drug-susceptible ST95 strains

We conducted a genome-wide analysis of ST95 genomes with 261 UTI ST95 isolates and 292 BSI ST95 isolates (80 from Northern California and 473 from the Enterobase ST95 collection) based on the metadata information. We limited the genome-wide association analysis to human BSI and UTI isolates (553 isolates in total) to control for disease status. The other isolates included those from nonhuman animals, the environment, and human non-urine or blood sources. The output of SNP-sites and Roary was used as variables. A logistic multivariate regression model adjusted for the carriage of pUTI89* and *fimH* types was conducted to identify genomic features significantly associated with drug susceptibility. We found 44 accessory genes significantly associated with strains that carried no recognized drug-resistance genes included in the ResFinder database (Table 3). We found 13 SNPs on core genes and 145 accessory genes associated with drug resistance (Supplementary Table 3). The correlation coefficient matrix of these significantly associated genes is shown in Figure 3.

Of the 44 genes associated with drug susceptible strains, 8 genes were annotated in the Prokka as *alkA* (DNA-3-methyladenine glycosylase 2), *cia* (colicin), *dinG*_1 (3'-5' exonuclease), *livK* (ABC transporter substrate-binding protein LivK), *maX_3* (PTS sugar transporter subunit IIA), *mltC* (membrane-bound lytic murein transglycosylase MltC), *traD* (type IV conjugative transfer system coupling protein TraD), and *umuD_2* (translesion error-prone DNA polymerase V autoproteolytic subunit). The other 36 genes were annotated by Prokka as encoding hypothetical proteins of unknown function. The sequences of 36 hypothetical proteins were sought by BLAST and information on protein function was obtained. Of these, 30 proteins were annotated in the NCBI database. Eight genes were annotated to encode transfer system or transporter-related proteins and 5 genes were annotated as phage genes. The functions of 44 genes are shown in Table 3.

2.5 Distribution of genes associated with pan-susceptible ST95 strains

Of 44 accessory genes significantly associated with drug susceptibility, 13 were found on pUTI89*, 19 were on non-pUTI89* plasmids, and 15 were on the chromosome (Table 3). Three genes were found both on pUTI89* plasmid and chromosomes.

We also sought these 44 gene sequences in ExPEC strains belonging to ST131, which are largely resistant to one or more antimicrobial agents. Fifteen of 44 genes did not match with any of the ST131 WGSs deposited in the NCBI database as of July 28, 2021 (Table 3). The other 29 genes had a match ranging from 1 to 28 from 7458 *Escherichia coli* O25b:H4-ST131 WGSs. NCBI has 127 complete whole genome sequences of *Escherichia coli* O25b:H4-ST131 and 7331 plasmid or short genome sequences of *Escherichia coli* O25b:H4-ST131 deposited. Of these sequences, 4 complete chromosome and 30 plasmid sequences carried a range of 1 to 14 genes associated with the pan-susceptible ST95 strains (Table 4).

3. Discussion

We identified genomic features associated with drug susceptibility of ExPEC ST95 lineage isolated from BSI and UTI patients, as well as other sources. We analyzed 1669 ST95 isolates deposited in Enterobase (as of March 12, 2020) and 80 ST95 isolates from UTI or BSI patients in Northern California. We also conducted a genome-wide association study with 553 ST95 BSI

and UTI human isolates. The sources of other isolates submitted to Enterobase include soil, water, vegetables, animals other than humans, and human samples other than urine or blood (e.g., stool or wound). Our previous study of WGSs of 86 isolates of ST95 from BSI patients in the United States found that the carriage of UTI89 like plasmid (pUTI89*) and *fimH6* was significantly associated with drug susceptibility (20). We also reported that in Northern California, ST95 strains remained persistently susceptible to most of the drugs we tested despite being one of the most prevalent ExPEC STs isolated from patients with community-acquired UTI (14). We hypothesized that ST95 strains restrict the acquisition of mobile DNA elements carrying drug-resistance genes and yet maintain a competitive advantage without becoming drug-resistant among the human intestinal microbiota.

Of the 1749 ST95 WGSs, 862 (49%) carried one or more drug-resistance genes. Among these, 94 (36%) of 261 UTI ST95 isolates and 154 (53%) of 292 BSI ST95 isolates carried drug-resistance genes, respectively. We observed a lower prevalence of drug-resistant ST95 isolates in Northern California compared to those in the Enterobase database (38% and 50%, respectively). This may be due to a submission bias to Enterobase—for example, drug-resistant ST95 may have been more likely to be sequenced and deposited in Enterobase. The Northern California isolates represented population-based samples (21, 25).

Principal component analysis and phylogenetic analysis based on the SNPs on core genes and the presence and absence of accessory genes were conducted to identify potential confounding factors associated with clustering. In both analyses, carriage of pUTI89* and *fimH* types were strongly associated with clustering. This suggests that the carriage of pUTI89* and the *fimH* types explain the majority of genomic variation we extracted from the pangenome analysis of the ST95 WGSs. This is consistent with our previous findings (20), and also offers more evidence for *fimH* typing as a useful method for genotyping ExPEC (26).

We used multivariate logistic models to identify genomic features significantly associated with drug pan-susceptibility while controlling for *fimH* types, and the carriage of pUTI89*. We limited the whole genome sequence data to human BSI or UTI isolates (553 isolates in total) to control for disease status. We found 44 accessory genes associated with drug susceptibility. We did not identify any SNPs in the core genes significantly associated with drug-susceptible strains. Of these 44 genes, 15 were not found in any of 7458 *E. coli* ST131 WGSs deposited in the NCBI nucleotide database. ST131 strains are frequently associated with multidrug resistance (15, 27), and thus these 15 ST95 genes not found in ST131 may include genetic features that confer upon ST95 strains the ability to resist becoming drug-resistant. The other 29 genes showed 1 to 28 matches with the ST131 genome sequences, but these matches represented less than 1% of the ST131 WGSs, whereas they represented 10-87% of the ST95 WGSs (Table 3).

Annotation of these 15 genes not found on any of the ST131 in the NCBI nucleotide database revealed that two genes were annotated as PTS sugar transporter subunit IIA and IIB, respectively. Previous reports revealed that PTS sugar transporters are involved in the penetration of polar antibiotics (28, 29) Three genes were associated with phage. The remaining functions were colicin, malate dehydrogenase, DNA breaking-rejoining protein, surface exclusion protein, and peptidase M14. Five were hypothetical proteins, functions unknown.

Acquisition and maintenance of accessory genes by *E. coli* must provide a population of *E. coli* strains an adaptive advantage in an ecological niche. ST95 strains may have evolved compensatory niche-adaptive mechanisms to remain viable in a variety of ecological niches (human and other animal intestines, environment) despite remaining susceptible to antibiotics. One such mechanism is the role of amino acid transporters reported to contribute to metabolic adaptation of uropathogenic *E. coli* (UPEC) in the urinary tract (30).

A limitation of this study is that the drug resistance phenotype of 1669 ST95 WGS obtained from the Enterobase database was predicted from the presence of drug resistance genes annotated in ResFinder and not by any experimental phenotyping tests. Although genotypic resistance has been shown to have 100% sensitivity and specificity in predicting phenotypic resistance in *E. coli* by some studies (31, 32), we cannot rule out the possibility of misclassifying drug resistance in this study. Nevertheless, we found strong associations of multiple accessory genes with the absence of recognized drug-resistance genes in ST95 strains.

Another limitation is the lack of consistent meta-data in the Enterobase database. Although the Enterobase database had data on 1669 ST95 isolates, the majority of them lacked information on the source of ST95 strains, which led to the reduction of sample size for the genome-wide association study (n=553). Lastly, submission bias of ST95 to Enterobase could have affected the prevalence of drug-resistant strains in our collection. The prevalence of drug-resistant ST95 strains in this collection may over-estimate the actual prevalence of drug-resistant ST95 in the human population.

In conclusion, we identified several genes associated with drug pan-susceptibility in ExPEC ST95 strains, which suggests that the ST95 lineage may have evolved to compensate for its susceptibility to antibiotics by acquiring unique fitness genes that enable a subpopulation of them to survive in multiple environmental niches without having to gain drug resistance. Experimental confirmation of these findings is needed to support this proposal, which is feasible with the relatively small number of unique genes in pan-susceptible ST95 strains we identified.

4. Materials and Methods

4.1 Sample collection in Northern California

As previously described by Yamaji et al. (14) and Adams-Sapper et al. (21), we consecutively collected urine samples from patients with symptoms of UTI attending our university outpatient health service between September 2016 and May 2017, and all BSI isolates from inpatients admitted to San Francisco General Hospital (SFGH) between July 2007 and September 2010. All *E. coli* isolates were genotyped by multilocus sequence typing (MLST) based on the seven-gene scheme described at the PubMLST website (https://pubmlst.org/bigsdb?db⫽pubmlst_mlst_seqdef) (33). All *E. coli* subtyped as ST95 were selected for WGS analysis. We analyzed 36 ST95 isolates from patients with UTI and 44 ST95 isolates from patients with BSI.

4.2 Short read genome sequencing

Genomic DNA was extracted from all strains with the Qiagen blood and tissue DNeasy kit. Library preparation for the MiSeq platform following a standard protocol for Illumina-compatible libraries was conducted at Functional Genomics Laboratory in Berkeley. Final

libraries were sequenced via a 300-bp paired-end run on a MiSeq instrument with V3 chemistry and standard Illumina analysis software at The California Institute for Quantitative Biosciences at UC Berkeley (QB3-Berkeley).

4.3 Genome assembly

MiSeq reads were screened and trimmed based on length and quality with BBDUK version 1.0 under the default setting (http://jgi.doe.gov/data-and-tools/bb-tools/). The trimming process also removed residual adapter sequences. Quality control of individual FASTQ files was conducted with FastQC (34) and all of the newly sequenced WGSdata passed the quality check. *De novo* assembly of trimmed paired reads for the 80 libraries was performed with Unicycler version 0.4.8 under the default setting (35). The number of reads used in each assembly was sufficient to give a minimum of 25-fold coverage, averaged across all contigs. A maximum of 45-fold coverage was used. Contigs that were less than 500bp in length and with less than 80% high-quality base calls were eliminated from subsequent analysis.

4.4 Complementary dataset

We obtained 1669 assembled whole genomes of ST95 and associated metadata from Enterobase (23) accessed on March 12, 2020. Geographical and collection information of these whole genomes of ST95 is shown in Supplementary Table 1.

4.5 Bioinformatics analysis

Using the ResFinder database, we identified antimicrobial drug resistance (AMR) genes among the 80 clinical ST95 isolates' genomes (UTI and BSI) and among the 1669 assembled genomes from the Enterobase database (accessed 1 April 2020). We defined drug-resistant strains as those having at least one recognized drug resistance gene, and pan-susceptible strains as those not having any recognized drug resistance gene. Plasmid incompatibility group and replicon types were identified with the PlasmidFinder database (accessed April 2020). *FimH* types were identified with the FimTyper database (accessed April 2020). All identifications were done with Abricate version 1.0.1(36) with a 95% identity threshold across the reference sequences. Annotation was performed on all the assembled genomes with Prokka version 1.14.0 (37). Genes annotated as encoding hypothetical proteins were further analyzed on BLAST (accessed April 2021) (38). A pangenome of the entire data set was constructed with Roary version 3.13.0 (39) with a 95% identity cutoff. Here, the core genomes were defined as genes present in more than 95% of the collection of the isolates, and accessory genomes were defined as genes present in less than or equal to 95% of the collection. Of the accessory genomes, genes present in less than 15% of the collection were defined as cloud genes.

A concatenated core CDS alignment was made from the Roary output, and we extracted single nucleotide polymorphism (SNP) information with snp-sites with the default option (40). Phylogenetic trees were constructed with FastTree (24) with maximum likelihood method with the GTR model based on the SNP alignment and the presence and absence of accessory genes from the Roary output. Visualization was done with iToL version 6.1.2 (http://itol.embl.de).

4.6 Statistical analysis

All statistical analyses were conducted with R version 3.6.3 (41). We used the prcomp function for principal component analysis. The association of a specific gene with drug resistance was tested with multivariate analysis with the glm function adjusted for the carriage of UTI89 like

plasmid and *fimH* types. For the regression analysis, only ST95 isolates from BSI and UTI patients were used to control for any other unmeasured potential confounding. We identified 473 BSI and UTI isolates based on the metadata information of Enterobase assessed on March 12, 2020. Combined with the 80 isolates from Northern California, 553 ST95 whole genome sequence data were used for this analysis. Relationships between each set of significantly associated genetic features were assessed with correlation coefficients. All graphs were plotted with ggplot2 (42).

4.7 Distribution of genes associated with pan-susceptibility

Once we identified ExPEC ST95 genes significantly associated with susceptibility to antimicrobial drugs, we analyzed the distribution and location of the gene in other ExPEC STs by BLAST (accessed June 2021) (38). We targeted ST131 because of their high frequency of drug resistance and the number of WGSs in the database (17, 43). We sought for these ST95 genes on (1) the ST131 chromosome, plasmid, or phage genes and (2) ST131 whole genome sequences deposited on NCBI database.

4.8 Accession numbers

The whole-genome shotgun sequence results described here have been deposited in DDBJ/ENA/GenBank under the accession numbers shown in Supplementary Table 4. Genome sequences have been deposited in NCBI BioProject under accession number PRJNA763994.

Acknowledgement

We thank Holly Berryman Stern and the other staff members of the Tang Center of University of California, Berkeley (UC Berkeley)-affiliated health care service for their time and support of this project. This work was supported by Centers for Disease Control and Prevention investments to combat antibiotic resistance under BAA 200-2016-91939. The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Figures

Figure 1

Maximum-likelihood phylogeny of 1774 ST95 isolates constructed with FastTree and visualized with iTOL

A: Phylogeny based on the core genome

B: Phylogeny based on presence and absence of accessory genome

Figure 2

Principle component analysis of 1774 ST95 isolates based on (A) core genome and (B) presence and absence of accessory genome

Figure 3

The correlation coefficient matrix of ST95 genes significantly associated with ST95 strains lacking drug-resistance genes.

Tables

Table 1

Frequency of antimicrobial resistance genes among *E. coli* isolates belonging to ST95 registered in Enterobase database and obtained in Northern California respectively, identified from ResFinder

Antibiotics class	Resistance gene name	Enterobase Total (%)	Northern California Total (%)
Pan susceptible		837 (50)	50 (62)
Aminoglycoside	*aac(3)-IId_1*	36 (2.16)	2 (2.56)
	aadA2_1	39 (2.34)	2 (2.56)
	aadA5_1	33 (1.98)	4 (5.13)
	ant(2'')-Ia_1	2 (0.12)	2 (2.56)
	ant(3'')-Ia_1	38 (2.28)	1 (1.28)
	aph(3'')-Ib_5	306 (18.33)	6 (7.69)
	aph(6)-Id_1	254 (15.22)	7 (8.97)
Beta-lactam	*blaCTX-M-14_1*	29 (1.74)	1 (1.28)
	blaCTX-M-15_1	44 (2.64)	1 (1.28)
	blaTEM-1A_1	18 (1.08)	3 (3.85)
	blaTEM-1B_1	335 (20.07)	21 (26.92)
	blaTEM-1C_1	148 (8.87)	1 (1.28)
Phenicol	*catA1_1*	16 (0.96)	2 (2.56)
	floR_2	17 (1.02)	1 (1.28)
Trimethoprim	*dfrA12_8*	33 (1.98)	2 (2.56)
	dfrA17_1	63 (3.77)	5 (6.41)
	dfrA5_1	153 (9.17)	2 (2.56)
Sulphonamide	*sul1_2*	1 (0.06)	1 (1.28)
	sul1_5	107 (6.41)	5 (6.41)
	sul2_2	278 (16.66)	4 (5.13)
	sul2_3	114 (6.83)	3 (3.85)
Tetracycline	*tet(A)_6*	211 (12.64)	8 (10.26)
	tet(B)_1	1 (0.06)	1 (1.28)
	tet(B)_2	107 (6.41)	5 (6.41)
Macrolide	*mph(A)_2*	42 (2.52)	3 (3.85)

Table 2

Frequency of drug resistance among ST95 isolates registered in the Enterobase database between 1947 and 2019 and ST95 isolates from Northern California between 2013 and 2017, by year, stratified by isolation information (BSI or UTI)

Year	Total			DR			Beta-lactam			Aminoglycoside			Tetracycline			Sulphonamide			Trimethoprim			Phenicol			Macrolide			Quinolone			Fosfomycin			Lincosamide			Colistin		
	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI	Total	UTI	BSI
1947	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1972	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1976	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1977	1		1	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1981	1		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1982	2		2	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1984	7	1	5	3	0	3	2	0	2	2	0	2	0	0	0	3	0	3	2	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1985	16		15	3		3	2		2	1		1	1		1	2		2	1		1	1		1	0		0	0		0	0		0	0		0	0		0	0	
1986	4	1	2	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1987	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1988	1			0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1991	1			1		1	0		0	0		0	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1992	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1993	1			0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1995	5	1	4	0		0	0		0	0		0	0		0	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1996	8	6	1	1	1	1	1		1	1		1	0		0	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
1998	11	3	3	3	1	2	0		0	0		0	1		0	0		0	0		0	0		0	1		1	0		0	0		0	0		0	0		0	0	
1999	4	1		0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
2000	6			0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
2001	22	10	10	15	6	7	14	5	7	8	2	6	4	2	2	9	3	5	5	2	3	6	2	4	9	4	5	9	0	9	0	0	0	0	0	0	0	0	0	0
2002	3			0			0			0			0			0			0			0			0			0			0			0			0			0	
2003	28		23	11		10	10		9	6		6	2		2	7		7	4		4	1		1	0		0	0		0	0		0	0		0	0		0	0	
2004	14		11	3		3	1		1	3		3	1		1	3		3	1		1	1		1	0		0	0		0	0		0	0		0	0		0	0	
2005	5		3	1		1	0		0	0		0	0		0	1		1	0		0	0		0	0		0	0		0	0		0	0		0	0		0	0	
2006	10	5	1	5	2	0	4	1	0	5	2	0	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2007	5		2	2		1	1		1	2		1	1		1	1		0	1		1	1		1	0		0	0		0	0		0	0		0	0		0	0	
2008	20		14	11		9	6		4	9		4	4		4	8		3	7		7	6		6	2		2	1		1	0		0	0		0	0		0	0	
2009	10		7	9		6	7		6	4		3	6		6	5		5	4		4	2		2	1		1	1		1	0		0	0		0	0		0	0	
2010	66	9	18	36	3	11	29	3	10	19	3	6	14	2	2	15	3	5	15	3	5	4	2	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2011	52	34	9	17	9	3	13	7	3	12	6	3	5	3	0	12	6	3	10	6	3	10	6	2	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2012	76	23	19	33	7	9	23	7	5	18	5	3	11	0	4	16	4	2	14	1	6	14	4	2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2013	68	39	19	26	8	13	20	7	9	10	1	5	7	1	4	11	1	6	9	2	6	9	2	0	2	2	3	1	1	1	0	0	0	0	0	0	0	0	0	0
2014	132	39	37	78	15	22	33	13	18	29	9	19	33	6	11	47	9	15	22	6	14	7	2	5	4	2	2	2	1	0	1	0	0	0	0	0	1	0	1	1
2015	138	32	32	71	22	26	51	16	23	42	6	23	31	10	15	41	7	21	32	7	17	1	0	1	5	1	2	4	0	4	0	0	0	1	0	1	1	0	1	1
2016	62	3	1	31	3	0	25	3	0	12	1	0	11	0	0	15	1	0	11	2	0	1	0	0	4	1	0	5	1	0	1	0	0	0	0	0	1	0	0	0
2017	91	5	2	44	3	1	37	3	1	14	0	0	13	1	1	21	0	0	18	2	0	3	0	0	3	0	0	3	0	0	1	0	0	0	0	0	3	0	0	0
2018	41			13			9			6			7			5			4			0			2			1			0			0			0			0	
2020	21			6			5			5			7			5			4			0			1			1			0			0			0			0	
Total*	946	217	235	407	60	127	293	67	98	211	36	90	165	25	48	235	36	83	159	31	65	56	4	14	27	5	8	20	2	8	5	0	2	1	0	1	6	0	1	2
Total**	1669	225	248	832	90	132	595	72	102	443	37	92	335	25	49	496	37	84	334	220	155	66	4	14	47	5	8	34	2	8	10	0	2	1	0	1	6	0	1	2
in CA (%)	80	36	44	30	38	8	22	22	50	27	34	6	17	5	14	10	23	3	8	10	23	9	11	1	3	1	3	2	5	3	7	0	0	0	0	0	0	0	0	0

*isolates with collection year information

**includes Enterobase isolates without collection year information

DR: drug resistant, CA: California

Table 3

Genes significantly associated with ST95 strains lacking drug-resistance genes. Annotation, location, and odds ratio with 95% confidence interval are shown for UTI and BSI isolates. These ST95 genes that had a matching sequence in *Escherichia coli* O25b:H4-ST131 WGSs from NCBI database are shown. Prevalence of these genes among 1774 ST95 collection is shown.

			UTI		BSI			
Gene/SNP	Annotation	location	OR [95% CI]	P-value	OR [95% CI]	P-value	ST131	ST95 (%)
Pan-Susceptible								
alkA	DNA-3-methyladenine glycosylase 2	chromosome	5.5 [1.7-20.4]	0.007	2.5 [1-6.6]	0.046	5	1420 (80)
cia	colicin	plasmid	4.8 [1.8-13.9]	0.003	2.6 [1.1-6.2]	0.023	0	318 (18)
dinG_1	3'-5' exonuclease	plasmid	5.3 [2-15.4]	0.001	4.7 [1.9-12.3]	0.001	1	311 (18)
livK	leucine-specific binding protein precursor, high-affinity branched-chain amino acid ABC transporter substrate-binding protein LivK	chromosome	3.8 [1.2-13.5]	0.03	4 [1.5-12]	0.008	5	1546 (87)
manX_3	PTS sugar transporter subunit IIA	chromosome	3.8 [1.2-13.5]	0.03	3.9 [1.5-11.8]	0.009	0	1547 (87)
mltC	membrane-bound lytic murein transglycosylase MltC	chromosome	2.1 [1-4.4]	0.042	2.1 [1-4.4]	0.043	1	1336 (75)
traD	type IV conjugative transfer system coupling protein TraD	plasmid	4.1 [1.9-9.2]	0	2.2 [1.2-4.3]	0.014	1	760 (43)
umuD_2	translesion error-prone DNA polymerase V autoproteolytic subunit	pUTI89	11.2 [3.6-38.5]	0	4.2 [1.6-10.9]	0.003	1	371 (21)
group_406	conjugal transfer mating pair stabilization protein TraG	plasmid	6.1 [2.5-16.1]	0	4.5 [2.2-9.7]	0	22	379 (21)
group_1304	hypothetical protein	pUTI89	43.5 [7.9-1000]	0	9.7 [2.5-62.5]	0.004	28	196 (11)
group_2195	DUF4113 domain-containing protein	plasmid	12 [4-40]	0	6.7 [1.9-24.4]	0.003	0	462 (26)
group_2212	DNA repair protein UmuC	pUTI89	18.9 [5.5-90.9]	0	5.8 [2-19.6]	0.002	2	264 (15)
group_2341	incFII family plasmid replication initiator RepA	plasmid	3.9 [1.6-10.2]	0.004	3.4 [1.5-7.5]	0.003	25	377 (21)
group_2492	conjugal transfer pilus acetylase TraX	plasmid	4.9 [2-12.7]	0.001	6.8 [3.3-15.2]	0	25	423 (24)
group_2899	phage integrase family protein	pUTI89	21.7 [6.2-83.3]	0	6.7 [1.9-24.4]	0.003	28	352 (20)
group_3175	DUF945 domain-containing protein	pUTI89	6.6 [1.7-43.5]	0.017	4.1 [1.1-19.6]	0.048	23	191 (11)
group_3361	transglycosylase SLT domain-containing protein	pUTI89	6.6 [1.9-30.3]	0.006	3.3 [1.1-11.2]	0.044	23	218 (12)
group_3450	fertility inhibition protein FinO	plasmid	3.4 [1.4-8.8]	0.009	3 [1.4-6.4]	0.005	25	384 (22)
group_3636	type IV conjugative transfer system protein TraL	pUTI89	6.2 [1.6-41.7]	0.022	5.6 [1.3-38.5]	0.036	22	169 (10)
group_4084	phage tail protein	chromosome	6.3 [1.2-50]	0.041	4.2 [1.4-13.3]	0.012	1	311 (18)
group_4555	TrbJ conjugal transfer protein	plasmid	6.4 [2.5-17.5]	0	4.5 [2.1-9.7]	0	12	395 (22)
group_5037	IS21 family transposase	pUTI89	23.3 [4.3-500]	0.003	9.3 [2.3-62.5]	0.005	5	191 (11)
group_5504	nuclease-like protein	plasmid	4.8 [2-12.2]	0.001	4.7 [2.4-9.6]	0	9	495 (28)
group_6014	IncF conjugal transfer surface exclusion protein TraT	plasmid	4 [1.7-9.4]	0.001	4.5 [2.2-9.3]	0	23	446 (25)
group_6737	phage minor tail protein G	chromosome	2.6 [1.2-5.9]	0.015	2.5 [1.2-5.3]	0.018	0	795 (45)
group_6764	hypothetical protein	plasmid	4.8 [1.6-17.5]	0.008	4 [1.3-15.6]	0.027	28	187 (11)
group_7168	tail fiber domain-containing protein	chromosome	6.3 [1.2-50]	0.041	4.3 [1.4-15.2]	0.016	1	305 (17)
group_8375	ISL3-like element ISEc53 family transposase	pUTI89, chromosome	16.4 [4.1-111.1]	0	8.9 [2.5-41.7]	0.002	4	238 (13)
group_8556	phage portal protein	chromosome	2.6 [1.2-5.9]	0.015	2.5 [1.2-5.4]	0.014	0	603 (34)
group_9933	DinI-like family protein	pUTI89, chromosome	10.8 [3.3-38.5]	0	4.8 [1.7-14.1]	0.003	1	373 (21)
group_9939	malate dehydrogenase	plasmid	3 [1.3-7.2]	0.011	4.2 [2-9.3]	0	0	400 (23)
group_9942	hypothetical protein	plasmid	4.1 [1.8-10]	0.001	3.8 [1.9-7.8]	0	0	459 (26)
group_9945	hypothetical protein	plasmid	4.1 [1.8-10.2]	0.001	3.3 [1.5-7.5]	0.003	0	382 (22)
group_10528	ATP-binding protein	pUTI89	19.6 [3.7-333.3]	0.005	16.1 [3-333.3]	0.009	1	183 (10)
group_11645	DNA breaking-rejoining protein	chromosome	2.6 [1.2-5.9]	0.015	2.5 [1.2-5.3]	0.018	0	610 (34)
group_12425	PAS domain-containing protein	pUTI89	4.5 [1.4-17.2]	0.016	9.2 [2.2-62.5]	0.007	7	226 (13)
group_13166	PTS sugar transporter subunit IIB	chromosome	3.8 [1.2-13.5]	0.03	3.7 [1.5-10.5]	0.009	0	1545 (87)
group_13675	ClbS/DfsB family four-helix bundle protein	pUTI89, chromosome	10 [2.9-47.6]	0.001	9.3 [2.7-43.5]	0.001	4	246 (14)
group_14237	surface exclusion protein	plasmid	3.1 [1.4-7]	0.005	4.9 [2.5-10.1]	0	0	528 (30)
group_14238	phage tail assembly protein T	chromosome	2.6 [1.2-6]	0.017	2.3 [1.1-5.2]	0.033	0	542 (31)
group_14249	hypothetical protein	plasmid	3.7 [1.6-8.8]	0.002	4.5 [2.2-9.2]	0	0	485 (27)
group_15802	replication regulatory protein RepA	plasmid	3.2 [1.4-7.7]	0.009	5.2 [2.6-10.8]	0	4	522 (29)
group_17896	DUF2190 family protein	chromosome	2.6 [1.2-5.9]	0.015	2.5 [1.2-5.4]	0.014	0	597 (34)
group_20939	peptidase M14	plasmid	3.2 [1.4-7.7]	0.006	4.5 [2.2-9.2]	0	0	485 (27)

Table 4

NCBI accession number and the number of matching regions of ST131 whole genome sequence that carries one or more genes of 44 ST95 genes found to be associated with ST95 strains lacking drug-resistance genes.

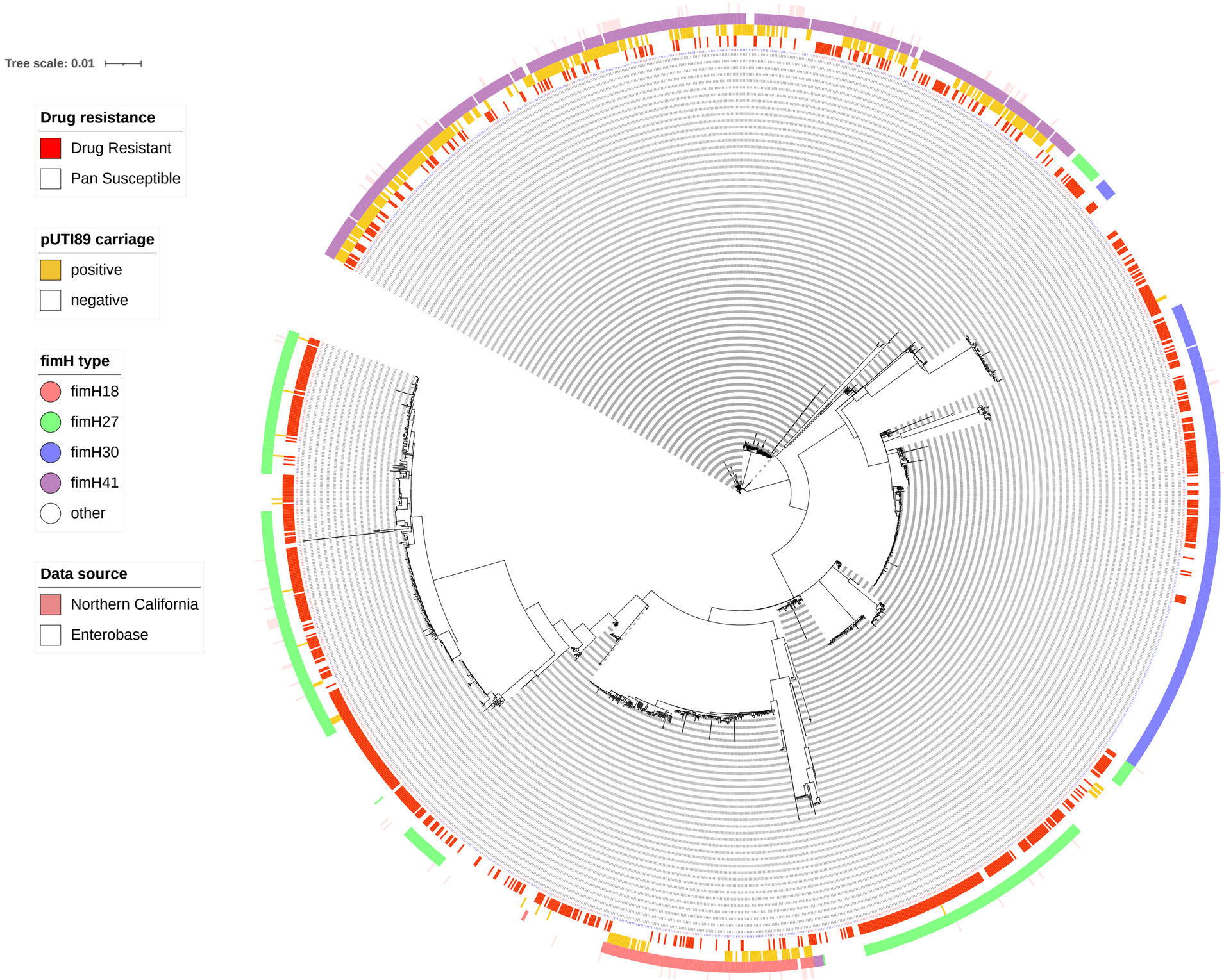
Location	NCBI accession number	Number of genes found
ST131 chromosome	CP063774	11
	CP051615	5
	CP051609	5
	HG941718	9
ST131 plasmid	AP018456	12
	AP018457	11
	AP018458	12
	AP019526	9
	CP063775	3
	CP051616	12
	CP051610	9
	CP015086	1
	HG941719	13
	JX077110	7
	MK295817	9
	MK295818	11
	MK295819	11
	MK295820	12
	MK295821	14
	MK295822	5
	MK295823	14
	MK295824	11
	MK295825	12
	MK295826	6
	MK295827	7
	MK295828	12
	MK295829	12
	MK295830	12
	MK295831	14
	MK295832	11
	MK295833	14
	MK295834	12
	MW646300	13
	MW646306	7

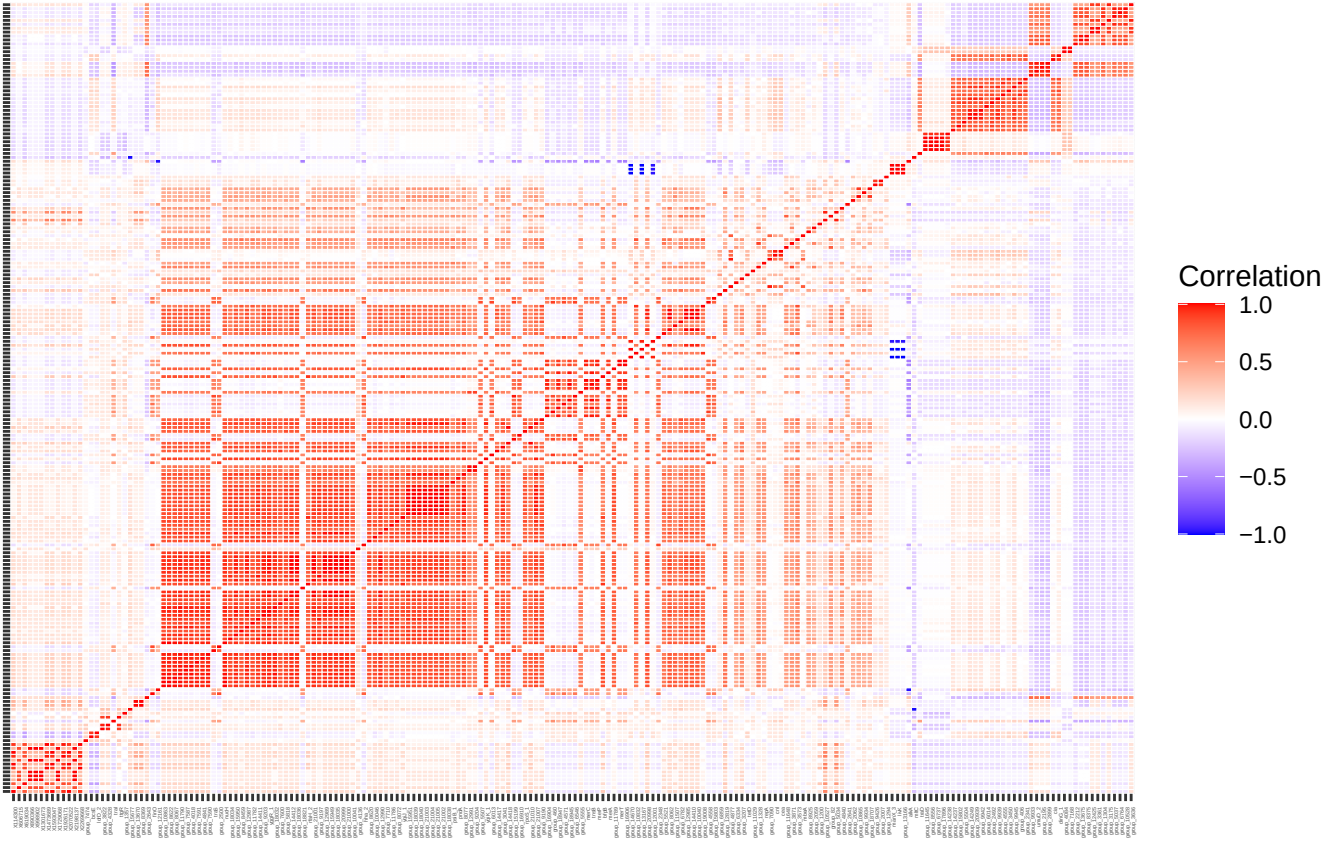
References

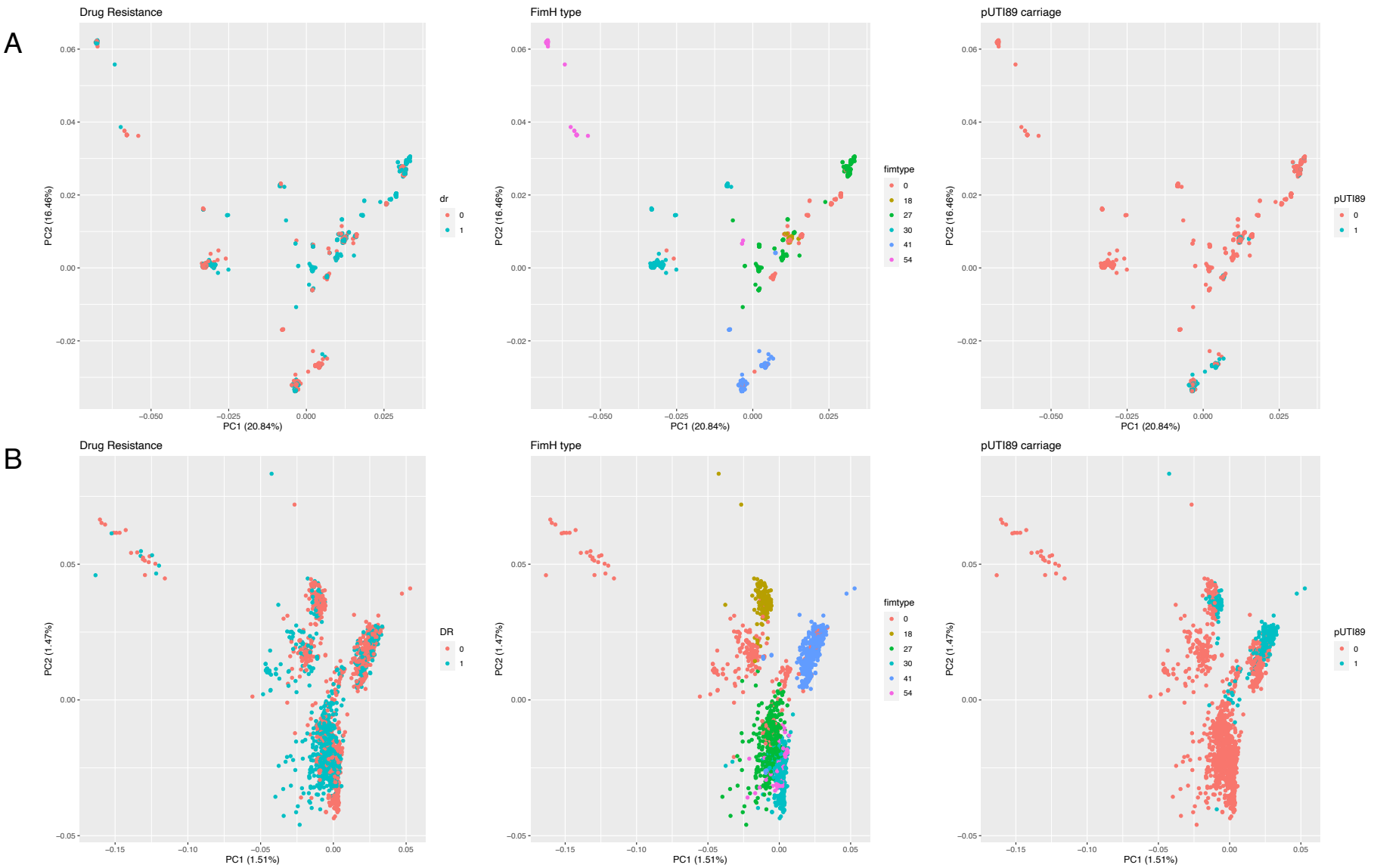
1. World Health Organization (WHO). 2015. Global Action Plan on Antimicrobial Resistance https://doi.org/10.1128/microbe.10.354.1.
2. World Health Organization (WHO). 2020. Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report.
3. Centers for Disease Control and Prevention. 2019. Antibiotic resistance threats in the United States.
4. WHO (World Health Organisation). 2014. ANTIMICROBIAL RESISTANCE Global Report on Surveillance.
5. Johnson JR, Russo TA. 2002. Extraintestinal pathogenic Escherichia coli: "The other bad E coli." J Lab Clin Med 139:155–162.
6. Thaden JT, Jr VGF, Sexton DJ, Anderson DJ. 2016. Increasing Incidence of Extended-Spectrum β-Lactamase- Producing Escherichia coli in Community Hospitals throughout the Southeastern United States. Infect Control Hosp Epidemiol 37:49–54.
7. Allocati N, Masulli M, Alexeyev MF, Di Ilio C. 2013. Escherichia coli in Europe: An overview. Int J Environ Res Public Health 10:6235–6254.
8. Yassin AK, Gong J, Kelly P, Lu G, Guardabassi L, Wei L, Han X, Qiu H, Price S, Cheng D, Wang C. 2017. Antimicrobial resistance in clinical Escherichia coli isolates from poultry and livestock, China. PLoS One 12:1–8.
9. Borges CA, Tarlton NJ, Riley LW. 2019. Escherichia coli from Commercial Broiler and Backyard Chickens Share Sequence Types, Antimicrobial Resistance Profiles, and Resistance Genes with Human Extraintestinal Pathogenic Escherichia coli. Foodborne Pathog Dis 16:813–822.
10. Y. M, T. N, M. T, T. K, M. Y, M. N, S. T, S. I, O. Y, T. H, M. S, H. W, T. S, A. H, T. H, N. F, T. K, Y. Y, K. U, G. Y, K. M, S. Y. 2017. Population structure of Japanese extraintestinal pathogenic Escherichia coli and its relationship with antimicrobial resistance. J Antimicrob Chemother 72:1040–1049.
11. Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, Pitout JDD. 2019. Global Extraintestinal Pathogenic Escherichia coli (ExPEC) Lineages 32:1–25.
12. Hu Y, Matsui Y, Riley LW. 2020. Risk factors for fecal carriage of drug-resistant Escherichia coli : a systematic review and meta-analysis. Antimicrob Resist Infect Control 3:1–12.
13. Matsui Y, Hu Y, Rubin J, de Assis RS, Suh J, Riley LW. 2020. Multilocus sequence typing of Escherichia coli isolates from urinary tract infection patients and from fecal samples of healthy subjects in a college community. Microbiologyopen 9:1225–1233.
14. Yamaji R, Rubin J, Thys E, Friedman CR, Riley LW. 2018. Persistent pandemic lineages of uropathogenic Escherichia coli in a college community from 1999 to 2017. J Clin Microbiol 56:1–10.
15. Riley LW. 2014. Pandemic lineages of extraintestinal pathogenic Escherichia coli. Clin Microbiol Infec 20:380–390.
16. Li D, Reid CJ, Kudinha T, Jarocki VM, Djordjevic SP. 2020. Genomic analysis of trimethoprim-resistant extraintestinal pathogenic Escherichia coli and recurrent urinary tract infections. Microb Genomics 6:1–19.
17. Ludden C, Decano AG, Jamrozy D, Pickard D, Morris D, Parkhill J, Peacock SJ, Cormican M, Downing T. 2020. Genomic surveillance of escherichia coli ST131 identifies local expansion and serial replacement of subclones. Microb Genomics 6.

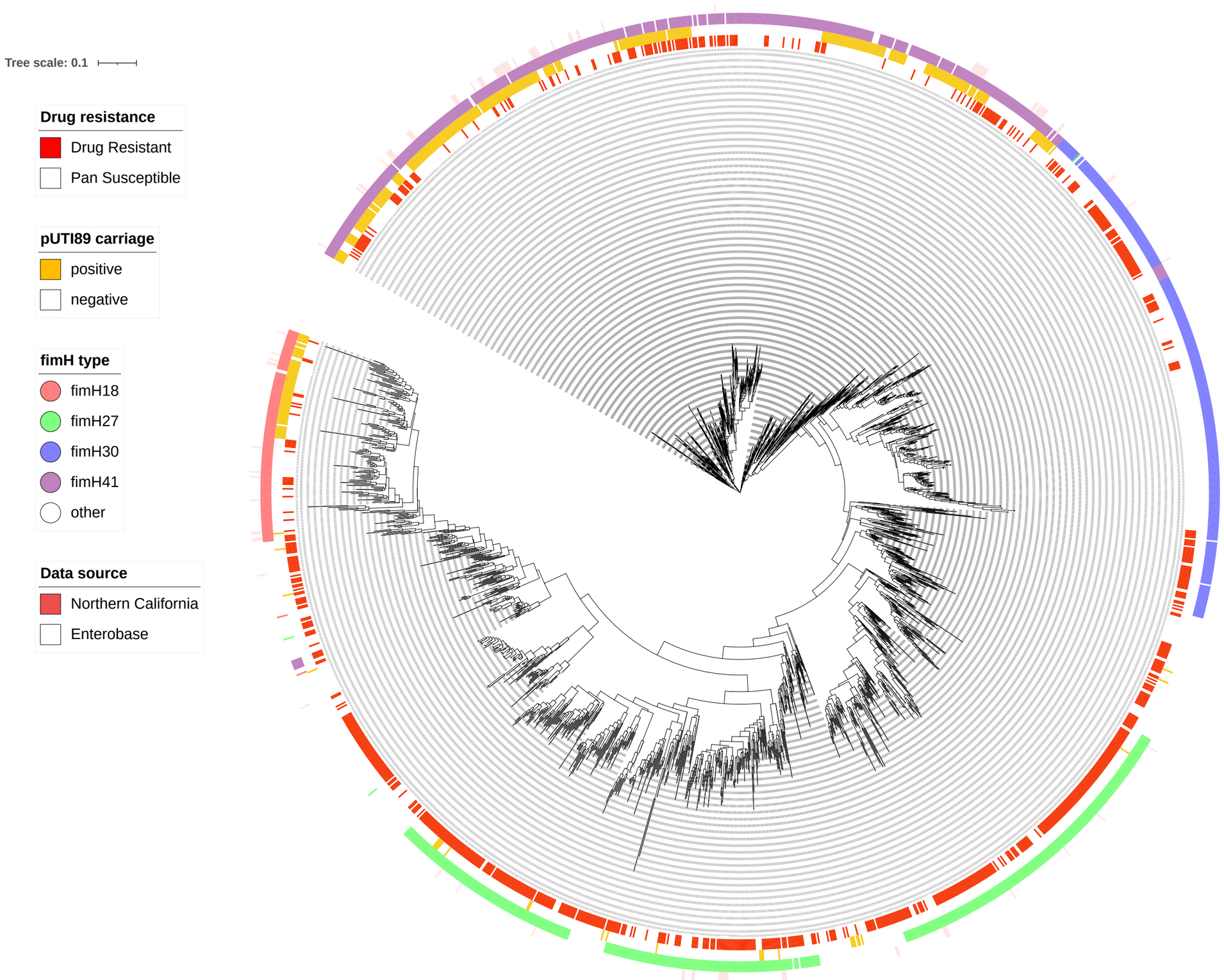
18. Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, Martin V, Peacock SJ, Parkhill J. 2017. Systematic longitudinal survey of invasive Escherichia coli in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. Genome Res 27:1437–1449.
19. McNally A, Kallonen T, Connor C, Abudahab K, Aanensen DM, Horner C, Peacock SJ, Parkhill J, Croucher NJ, Corander J. 2019. Diversification of colonization factors in a multidrug-resistant escherichia coli lineage evolving under negative frequency- dependent selection. MBio 10:1–19.
20. Stephens CM, Adams-sapper S, Sekhon M, Johnson JR, Riley LW. 2017. Genomic Analysis of Factors Associated with Low Prevalence of Antibiotic Resistance in Extraintestinal Pathogenic Escherichia coli Sequence Type 95 Strains. mSphere 2:e00390-16.
21. Adams-Sapper S, Diep BA, Perdreau-Remington F, Riley LW. 2013. Clonal composition and community clustering of drug-susceptible and -resistant Escherichia coli isolates from bloodstream infections. Antimicrob Agents Chemother 57:490–497.
22. Stephens C, Adams-Sapper S, Sekhon M, Johnson J, Riley L. Genomic analysis of Escherichia coli ST95 isolates from bloodstream infections reveals multiple sub-lineages with distinct plasmids and antibiotic resistance gene content. https://doi.org/10.1128/EC.4.3.604.
23. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Achtman M. 2020. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res 30:138–152.
24. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. PLoS One 5.
25. Yamaji R, Friedman CR, Rubin J, Suh J, Thys E, McDermott P, Hung-Fan M, Riley LW. 2018. A Population-Based Surveillance Study of Shared Genotypes of Escherichia coli Isolates from Retail Meat and Suspected Cases of Urinary Tract Infections. mSphere 3:e00179-18.
26. Dias RCS, Moreira BM, Riley LW. 2010. Use of fimH single-nucleotide polymorphisms for strain typing of clinical isolates of Escherichia coli for epidemiologic investigation. J Clin Microbiol 48:483–488.
27. Johnson JR, Johnston B, Clabots C, Kuskowski MA, Castanheira M. 2010. Escherichia coli Sequence Type ST131 as the Major Cause of Serious Multidrug-Resistant E. coli Infections in the United States . Clin Infect Dis 51:286–294.
28. Jeckelmann JM, Erni B. 2020. Transporters of glucose and other carbohydrates in bacteria. Pflugers Arch Eur J Physiol 472:1129–1153.
29. Garcia De Gonzalo C V., Denham EL, Mars RAT, Stülke J, Van Der Donk WA, Van Dijl JM. 2015. The phosphoenolpyruvate: Sugar phosphotransferase system is involved in sensitivity to the glucosylated bacteriocin sublancin. Antimicrob Agents Chemother 59:6844–6854.
30. Mann R, Mediati DG, Duggin IG, Harry EJ, Bottomley AL. 2017. Metabolic adaptations of Uropathogenic E. coli in the urinary tract. Front Cell Infect Microbiol 7.
31. Su M, Satola SW, Read TD. 2019. Genome-based prediction of bacterial antibiotic resistance. J Clin Microbiol 57.
32. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agersø Y, Lund O, Larsen MV, Aarestrup FM. 2013. Genotyping using whole-genome sequencing is a realistic alternative to

surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother 68:771–777.
33. Tartof SY, Solberg OD, Manges AR, Riley LW. 2005. Analysis of a Uropathogenic Escherichia coli Clonal Group by Multilocus Sequence Typing. J Clin Microbiol 43:5860–5864.
34. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.
35. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:1–22.
36. Seemann T. 2014. ABRicate.
37. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
38. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:1–9.
39. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.
40. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb genomics 2:e000056.
41. R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria.
42. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
43. Totsika M, Beatson SA, Sarkar S, Phan MD, Petty NK, Bachmann N, Szubert M, Sidjabat HE, Paterson DL, Upton M, Schembri MA. 2011. Insights into a multidrug resistant Escherichia coli pathogen of the globally disseminated ST131 lineage: Genome analysis and virulence mechanisms. PLoS One 6.


Tree scale: 0.01
Drug resistance
Drug Resistant
Pan Susceptible
pUTI89 carriage
positive
negative
fimH type
fimH18
fimH27
fimH30
fimH41
other
Data source
Northern California
Enterobase


Correlation
1.0
0.5
0.0
−0.5
−1.0


A
Drug Resistance
FimH type
pUTI89 carriage
PC2 (16.48%)
PC1 (20.84%)
dr
fimtype
pUTI89
0
1
18
27
30
41
54
B
PC2 (1.47%)
PC1 (1.51%)
DR


Tree scale: 0.1
Drug resistance
Drug Resistant
Pan Susceptible
pUTI89 carriage
positive
negative
fimH type
fimH18
fimH27
fimH30
fimH41
other
Data source
Northern California
Enterobase