First Indian report on Genome-wide Comparison of Multidrug-Resistant *Escherichia coli* from Blood Stream Infections

[1]Naveen Kumar Devanga Ragupathi, [1*]Balaji Veeraraghavan, [1]Dhiviya Prabaa Muthuirulandi Sethuvel, [1]Shalini Anandan, [1]Karthick Vasudevan, [1]Ayyan Raj Neeravi, [1]Jones Lionel Kumar Daniel, [2]Sowmya Sathyendra, [2]Ramya Iyadurai, [3,4]Ankur Mutreja

[1]Department of Clinical Microbiology, Christian Medical College, Vellore 632004, India
[2]Department of Medicine, Christian Medical College, Vellore 632004, India
[3]Department of Medicine, Addenbrookes Hospital, University of Cambridge, Cambridge, United Kingdom
[4]Wellcome Sanger Institute (WTSI), Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom

*Corresponding author:
Dr. V. Balaji
Professor
Department of Clinical Microbiology
Christian Medical College
Vellore – 632004
E-mail: vbalaji@cmcvellore.ac.in
Ph: +91 9442210555

Abstract:

Background

Multidrug-resistant (MDR) *E. coli* with extended-spectrum β-lactamases (ESBLs) is becoming endemic in health care settings around the world. Baseline data on virulence and AMR of specific lineages of *E. coli* circulating in developing countries like India is currently lacking.

Methods

Whole-genome sequencing was performed for 60 MDR *E. coli* isolates. Genome-wide analysis was performed at single nucleotide polymorphism (SNP) level resolution to identify the relation between the isolates in context of time, virulence and AMR determinants possessed.

Results

Genome comparison revealed the presence of ST-131 global MDR and ST410 as emerging-MDR clades of *E. coli*. AMR gene profile for cephalosporin and carbapenem resistance differed between the clades. Genotypes $bla_{\text{CTX-M-15}}$ and $bla_{\text{NDM-5}}$ were common among cephalosporinases and carbapenemases, respectively. For aminoglycoside resistance, *rmtB* was positive for 31.7% of the isolates, of which 30% were co-harboring carbapenemases. Further, the FimH types and virulence gene profile positively correlated with the SNP based phylogeny, which also revealed the evolution of MDR clones among the study population with temporal accumulation of SNPs. The predominant clone was ST167 (bla_{NDM} lineage) followed by ST405 (global clone ST131 equivalent) and ST410 (fast spreading high risk clone).

Conclusions

This is the first report on the whole genome analysis of MDR *E. coli* lineages circulating in India. Data from this study will provide public health agencies a baseline portfolio of AMR and virulence in pathogenic *E. coli* in the region.

Keywords: *E. coli*; ST131; ST167; ST410; high risk clone; SNP phylogeny

Introduction

Escherichia coli is the leading cause of bloodstream infections (BSIs) caused by Gram-negative bacteria [1] and other common infections including urinary tract infections (UTIs). As an important commensal component of the biosphere, *E. coli* is colonizes the lower gut of animals and humans and gets released in the environment as a norm allowing widespread dissemination.

Virulence of *E. coli* is driven by multiple factors that involves adhesins, toxins, siderophores, lipopolysaccharide (LPS), capsule, and invasins [2]. It was reported that a large proportion of MDR *E. coli* carried by people are acquired via food, especially from farm animals [3]. Although most of the MDR *E. coli* are reported to be community acquired rather than healthcare [4,5], recently MDR *E. coli*, which produce extended-spectrum β-lactamases (ESBL) have been endemic in health care settings.

Among MDR *E. coli*, AMR mediated by ESBL is mainly due to the *bla*CTX-M family, particularly *bla*CTX-M-15 and 14, compared to the less frequent observation of *bla*SHV and *bla*OXA families [6-8]. Carbapenem resistance in *E. coli* was mostly reported to be mediated by *bla*OXA-48 [9], *bla*NDM and *bla*VIM [10]. Also, AMR among *E. coli* is increasingly reported for fluoroquinolones and third- and fourth-generation cephalosporins. Sequence type 131 (ST131) predominates globally among such MDR *E. coli* strains [11].

Our study was aimed at identifying the virulence and AMR genetic determinants predominant in MDR *E. coli*. We constructed core genome phylogeny using high quality SNP information to infer the genome wide association of virulence and resistance attributes among the sequenced samples.

Materials and Methods

Isolates and identification

A total of 99257 specimens were screened from BSI during the year 2015 to 2016 from the patients attending Christian Medical College, Vellore, India. Isolation and identification of the organism were carried out using a standard protocol [12]. Of the 1100 *E. coli* positives, 10% were observed resistant to carbapenems, of which 60 MDR isolates were selected in random for further characterization.

Antimicrobial susceptibility testing (AST)

Disc diffusion

AST testing was carried out using the Kirby-Bauer disk diffusion method. The antimicrobial agents tested were Amikacin (30 µg), netilmicin (30 µg), gentamycin (10 µg), chloramphenicol (30 µg), ciprofloxacin (5 µg), cefotaxime (30 µg), cefoxitin (30 µg), ceftazidime (30 µg), cefpodoxime (10 µg), piperacilllin-tazobactam (100/10 µg), cefoperazone-sulbactam (75/30), imipenem (10 µg) and meropenem (10 µg), tigecycline (15 µg) and tetracycline (30 µg) according to guidelines suggested by CLSI M100-S27, 2017. Quality control strains (*K. pneumoniae* ATCC 700603, *P. aeruginosa* ATCC 27853 and *E. coli* ATCC 25922) were used in all batches as recommended by the Clinical and Laboratory Standards Institute.

Minimum Inhibitory Concentration (MIC) for colistin

Colistin MIC for the studied isolates was determined by broth microdilution and interpreted using CLSI 2017 breakpoint recommendations. *mcr-1* positive *E. coli* with the expected range 4 – 8 µg/ml, *E. coli* ATCC 25922 (0.25 – 2 µg/ml) and *P. aeruginosa* ATCC 27853 (0.5 – 4 µg/ml) were used as quality control (QC) strains for colistin MIC determination.

Next generation sequencing and genome assembly

Genomic DNA was extracted using a QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany). Whole genome sequencing (WGS) was performed using an Ion Torrent^{TM} Personal Genome Machine^{TM} (PGM) sequencer (Life Technologies, Carlsbad, CA) with 400-bp read chemistry according to the manufacturer’s instructions. Data were assembled with reference *E. coli* strain (NC000913) using Assembler SPAdes v.5.0.0.0 embedded in Torrent Suite Server v.5.0.3.

Genome annotation

The assembled sequence was annotated using PATRIC, the bacterial bioinformatics database and analysis resource (http://www.patricbrc.org), and NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP, http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html). Downstream analysis was performed using the CGE server (http://www.cbs.dtu.dk/services) and PATRIC. The resistance gene profile was analysed using ResFinder 2.1 from the CGE server (https://cge.cbs.dtu.dk//services/ResFinder/). The sequences were also screened for antimicrobial resistance genes in the Antibiotic Resistance Genes Database (ARDB) and Comprehensive Antibiotic Resistance Database (CARD) through PATRIC. Virulence genes from the genomes were identified using VirulenceFinder 2.0 (https://cge.cbs.dtu.dk/services/VirulenceFinder/). Serotype of the isolates were identified using SerotypeFinder 1.1 (https://cge.cbs.dtu.dk/services/SerotypeFinder/).

Genome based MLST analysis

Sequence types (STs) were analysed by MLST 1.8 (MultiLocus Sequence Typing) (https://cge.cbs.dtu.dk//services/MLST/). To visualize the possible evolutionary relationships between isolates, STs of the study isolates and the globally reported strains were computed using PHYLOViZ software v2.0 based on goeBURST algorithm. The study used Warwick database for all sequence based MLST analysis of *E. coli*.

Genome comparison analyses

Gview, interactive genome viewer was used to compare the annotated *E. coli* genome arrangements with the reference *E. coli* K12 genome (NC_000913) [13]. Core genome analysis was performed using Roary: the Pan Genome Pipeline v3.11.2 from Sanger Institute [14]. Further, the tree file was visualised and analysed in iTOL v4 (https://itol.embl.de/).

This Whole Genome Shotgun project has been deposited at GenBank under the accession numbers PVPX00000000, PVPW00000000, PVPV00000000, PVPU00000000, PVPT00000000, PVPS00000000, PVPR00000000, PVPQ00000000, PVPP00000000, PVPO00000000, PVPN00000000, PVPM00000000, PVPL00000000, PVPK00000000, PVPJ00000000, PVPI00000000, PVPH00000000, PVOS00000000, PVPG00000000, PVPF00000000, PVPE00000000, PVPD00000000, PVPC00000000, PVPB00000000, PVPA00000000, PVOZ00000000, PVOY00000000, PVOX00000000, PVOW00000000, PVOV00000000, PVOT00000000, RCAC00000000, RCAE00000000, RCAI00000000, RCAH00000000, RCAJ00000000, RCAD00000000, RCAG00000000, RCAF00000000, RCAL00000000, RCAN00000000, RCAM00000000, RCAK00000000, SAZJ00000000, SAZP00000000, SAZU00000000, SAZV00000000, SAZF00000000, SAZG00000000, SAZH00000000, SAZI00000000, SAZK00000000, SAZL00000000, SAZM00000000, SAZN00000000, SAZO00000000, SAZQ00000000, SAZR00000000, SAZS00000000, SAZT00000000. The version described in this manuscript is version 1.

Results

Antimicrobial susceptibility

All 60 *E. coli* isolates were resistant to carbapenems, quinolones, cephalosporins and beta-lactamase inhibitors (Table S1). Whereas all the isolates were susceptible to colistin except B7532 and B9021, which exhibited an MIC of 32 µg/ml.

Whole genome sequence analysis

Phylogeny of MDR E. coli

MLSTFinder revealed the different sequence types of the isolates. The study isolates belonged to 6 clonal complexes with 14 different sequence types. Few of the sequence types were observed to share same founder types revealing the evolution of these strains. CC10 and CC 405 were the two major CCs observed with ST-167, ST-410 and ST-405 as the common STs. Interestingly, nine isolates belonging to CC/ST-131 were identified, of which, all were of H-30 clade, except the isolate BA9313 (H-24). goeBURST analysis revealed the relation between sequence types observed within same clonal complex (Figure 1).

Figure 1: A) goeBURST image of *E. coli* showing clonal complex 10 with STs 44, 167, 410, 448, 617, 1702 and 2851, B) goeBURST image of *E. coli* showing clonal complex/ ST-131 (a), and CC405 (b) with STs 405 and 5954.

E. coli genome comparison

Whole genome composition of 60 MDR *E. coli* was compared with the *E. coli* K-12 reference genome which shoes the region of differences between these genomes (Figure S1). Totally 2518792 SNPs were identified in all analyzed genomes. On minimum 5957 and maximum 74713 SNPs were identified in each of the 60 MDR *E. coli* genomes when compared to the reference genome.

Figure S1: Circular genome plot comparing 60 MDR *E. coli* thereby showing differences in genome composition in comparison to the reference genome NC000913 *E. coli*.

Core vs pan genome

Comparison between the core and pan genomes of 60 MDR *E. coli* isolates revealed 2258 core genes across all 60 isolates among the 17944 total gene clusters. This includes 600 soft core genes in 57 to 59 isolates, 3984 shell genes in 9 to 57 isolates and 11102 genes in less than 9 isolates (Figure 4). Further the pan genome matrix of the isolates depict the similarity between core genes and pan genes. Low similarity was observed across the genomes for the pan genes with high gene numbers and vice versa for the core genes (Figure 2).

Figure 2: Pan genome vs core genome comparison depicting high number of pan genomes and low number of conserved genes A), Break-up of core genes, soft core genes and pan genes B) among the 60 MDR *E. coli* isolates from BSI

Figure 3: Pan genome matrix identified with the core gene sets of 60 MDR *E. coli* isolates depicting the important antimicrobial resistance (*bla*NDM-5, *bla*CTX-M-15, *bla*OXA-10, *mcr-1.1*) and virulence genes (invG, fimH, fliC2, intA) respectively.

Serotype prediction

Based on the whole genome data, serotype of the MDR *E. coli* isolates were detected. O102:H6 was the commonest serotype (18.75%) followed by O89:H9 (15.6%), O8:H9

(12.5%), O89:H5 (9.4%) and other serotypes. Few isolates had partial alignment of O or H antigen genes with the reference strains which further needs to be confirmed by additional tests.

Genetic virulence factors of MDR E. coli

Three common virulence gene profiles were observed among the isolates as follows, i) *iss*, *capU*, *gad*, ii) *ipfA*, and iii) *eilA*, *gad*, *air*. The FimH virulence typing revealed the types 5, 24, 27, 28, 30, 35, 54, 191, 54-like in comparison to the database.

Comparison of virulence and clonal traits

The virulence gene profiles of the 60 isolates were then compared to the FimH virulence types, MLST sequence type and SNP phylogeny. The *E. coli* isolates showed distinct groups including ST-131; H-30 clade based on the virulence genes identified (Figure 6). The sequence types were linked to the groups of same virulence gene profile and FimH virulence types.

Figure 4: SNP phylogeny based comparison of genetic virulence traits observed in MDR *E. coli* strains exhibiting, sequence types, virulence gene profile, O and H antigens and Fim-H types.

Antimicrobial resistance genetic determinants

ResFinder revealed the presence of multiple AMR genes in each of the MDR *E. coli* (Figure 7). Aminoglycoside and beta lactam resistance genes were seen high compared to other groups. Among the aminoglycoside resistance genes, *aadA5* and *aac(6')lb-cr* were higher followed by *aadA2* and *rmtB*. While among beta lactamases, *bla*CTX-M-15 was higher followed by *bla*NDM-5, *bla*OXA-1 and *bla*TEM-1B. Similarly, most of the isolates harboured *mphA*, *catB4*, *sul1*, *tetB*, *dfrA17* and *dfrA12*. Interestingly, two isolates, B7532 and B9021 carried *mcr-1.1* responsible for plasmid-mediated colistin resistance. The two isolates also showed phenotypic resistance to colistin with high MIC (>32 µg/ml).

Figure 5: Antimicrobial resistance genes observed in MDR *E. coli* compared to the SNP based phylogeny. Depicting prevalence of *bla*NDM-5 among carbapenemases and *bla*CTX-M-15 among cephalosporinases.

Discussion

The increasing usage of third-generation β-lactams and β-lactam inhibitors was accompanied with increases in prevalence of the MDR phenotype among *E. coli*. A study from India with invasive *E. coli* isolates between 2014 and 2016 reported susceptibility to cefoxitin 53%, ceftazidime 33%, cefotaxime 26%, ceftriaxone 25%, cefepime 29%, piperacillin tazobactam 66%, imipenem and meropenem 89%, aztreonam 36%, ciprofloxacin 19%, levofloxavin 23%, and amikacin 91% [15]. Among these, about 64% of *E. coli* were found to be ESBL producers.

The increasing frequency of antimicrobial resistance in clinical *E. coli* isolates was known to be associated with genetic determinants such as $bla_{\text{CTX-M}}$, bla_{NDM}, and *mcr* genes. In this study, multiple AMR genes for beta lactams, carbapenems, fluoroquinolones, tetracycline, aminoglycosides and colistin were identified. The presence of genotypic AMR genes correlated well with phenotypic expression for beta lactams, carbapenems, fluoroquinolones and tetracycline. Plasmids IncFII majorly carried AMR genes $bla_{\text{CTX-M-15}}$, $bla_{\text{NDM-5}}$, *aadA2,* *rmtB* , *sul1, drfA12*, *erm*(B) and *tetA*, while IncFI plasmids carried mostly *aadA5, sul2, dfrA17, mph*(A) and *tetB* genes [16].

The MDR *E. coli* isolates upon SNP based phylogeny grouped to four major clades, ST167, ST410, ST405 and ST131. Among the isolates, for carbapenem resistance, *bla*NDM-5 was common in ST131, ST405 and ST410 clades. While among ST167, *bla*$_{\text{NDM-4, -5}}$ and $_{\text{-7}}$ were seen. Whereas, previous reports identified *bla*$_{\text{NDM-1}}$ as common among *E. coli* though the sample size was lesser [17], whereas in China, *bla*$_{\text{NDM-1}}$ and *bla*$_{\text{NDM-5}}$ were seen in equal numbers [18].

bla$_{\text{OXA-48}}$ type carbapenemases are the most commonly reported in *E. coli* [19] followed by *bla*$_{\text{NDM}}$ [20], *bla*$_{\text{IMP}}$ [21] and *bla*$_{\text{KPC}}$ [22] type carbapenemases have also been determined. Studies report occurrence of *bla*$_{\text{OXA-48}}$ from as low as 3% to 22% [23, 9]. A report from India has shown, *bla*$_{\text{NDM}}$ was common among carbapenemases in *E. coli* (70%), followed by *bla*$_{\text{OXA-48}}$ (24%) and *bla*$_{\text{VIM}}$ (17%). *bla*$_{\text{NDM}}$ alone is 48%, *bla*$_{\text{OXA-48}}$ alone is 19%, *bla*$_{\text{NDM}}$+*bla*$_{\text{OXA-48}}$ is 5%, *bla*$_{\text{NDM}}$+*bla*$_{\text{VIM}}$ is 17% [24]. In our study, the combinations observed are, *bla*$_{\text{NDM-1}}$ - 1, *bla*$_{\text{NDM-4}}$ - 1, *bla*$_{\text{NDM-5}}$ – 6, *bla*$_{\text{NDM-7}}$ – 1, *bla*$_{\text{OXA-1}}$ - 21, bla$_{\text{OXA-181}}$ - 2, *bla*$_{\text{NDM-5}}$+*bla*$_{\text{OXA}}$ - 16, *bla*$_{\text{NDM-5}}$+*bla*$_{\text{OXA-181}}$ – 1, *bla*$_{\text{NDM-1}}$+*bla*$_{\text{OXA}}$ – 2. This revealed that *bla*$_{\text{OXA-1}}$ was predominant followed by *bla*$_{\text{NDM}}$ in carbapenem resistant *E. coli* in our setting, though *bla*$_{\text{OXA-181}}$ was rare.

In addition, aminoglycoside resistance had been an important concern among Gram-negatives. Acquired 16S-RMTases were known to confer extremely high level of aminoglycoside resistance, due to which key aminoglycosides including gentamicin, tobramycin, and amikacin are ineffective against carbapenem resistant strains [25].

Accordingly, plazomicin, a new aminoglycoside agent identified to combat against carbapenem-resistant Enterobacteriaceae, was found inactive if the isolates co-produced 16S-RMTases [26]. In this study, about 90% of the RMTase positive *E. coli* co-harbored carbapenemases, which further adds to the burden of carbapenem resistance. Similarly, Taylor et al. [27] and Poirel et al. [28] had reported 83.1% and 45.4% co-occurrence of carbapenemases in 16S RMTase producing Enterobacteriaceae, respectively.

For cephalosporin resistance, the isolates of ST131 and ST405 clades carried *bla*$_{\text{CTX-M-15}}$, whereas ST167 and ST410 carried *bla*CTX-M-15 and *bla*$_{\text{CMY}}$ genes. Previously 54.34% *bla*$_{\text{CTX-M}}$ was reported in ESBL positive isolates from India [29]. Recently, plasmid-mediated colistin resistance is increasingly reported from *E. coli* [30-32]. This study observed two isolates (B7532, B9021) with *mcr-1.1* expressing high MIC of >32 μg/ml to colistin. Though the isolates were from 2007, no records of *mcr* in *E. coli* were observed after 2007.

The antimicrobial susceptibility of *E. coli* has been shown to vary geographically [33]. Among the different clonal groups observed, *E. coli* ST131 was most commonly associated with community acquired infection, which recently were highly associated with healthcare settings. ST131 was reported as the predominant lineage carrying *bla*$_{\text{CTX-M-15}}$ and other ESBLs. Most of the MDR *E. coli* carrying *bla*$_{\text{CTX-M-15}}$ from different countries in Europe and North America were homogenously grouped into the *E. coli* O25:H4-ST131 [6,35-36]. In this study, 87% of the isolates carried *bla*$_{\text{CTX-M-15}}$, among various STs, where only nine isolates belong to ST131. Among the observed STs in this study, *bla*$_{\text{CTX-M-15}}$ was previously reported for its strong association among ST167, ST617, ST405, ST410, ST131 and ST361 [36].

Though ST131 clones were predominantly reported worldwide, the STs observed in this study were known for its distinct lineages. ST167 was previously reported for its ability to carry *bla*$_{\mathrm{NDM}}$ genes in China [37-39]. ST405 had been known as another global clonal group similar to ST131 [40]. ST405 lineage was reported to carry *bla*$_{\mathrm{NDM}}$ genes in hospital settings [41] and also reported to carry *bla*KPC-2 [42]. In addition, ST405 was reported as a lineage carrying fluoroquinolone resistance in Japan [40]. Recently, ST410 was reported as a possible international high risk clone with B2/H24R, B3/H24Rx, and B4/H24RxC AMR clades. B3/H24Rx was reported to be evolved by acquisition of the *bla*$_{\mathrm{CTX\text{-}M\text{-}15}}$ and an IncFII plasmid. B4/H24RxC emerged by acquiring IncX3 plasmid with *bla*$_{\mathrm{OXA\text{-}181}}$ known for carbapenem resistance, which further acquired *bla*$_{\mathrm{NDM\text{-}5}}$, on a conserved IncFII plasmid [43]. In this study, all ST410 isolates ($n = 7$) harboured *bla*$_{\mathrm{CTX\text{-}M\text{-}15}}$ gene, while only B25762, BV643, B12243 and B32605 had IncFII plasmids and *bla*$_{\mathrm{NDM\text{-}5}}$ (B3/H24RxC). B12243, in addition harboured IncX3 with *bla*$_{\mathrm{OXA\text{-}181}}$ (B4/H24RxC), while BA22372 and BA9615 had only *bla*$_{\mathrm{OXA\text{-}181}}$ in IncX3 plasmid (B4/H24RxC).

Virulence genes observed among the *E. coli* isolates varied according to the different clades observed. The comparison of the virulence gene type with SNP based phylogeny revealed the acquaintance and deletion of virulence genes. Genes *iss*, *capU* and *gad* were observed in ST167 clade. ST131 possessed *iha*, *sat*, *cnfl* and *senB* in addition to *iss* and *gad*. ST131 had lost the *capU* genes. Further, ST405 clade also lost *iss* and gained *eilA* and *air* genes with FimH type 29. Few isolates of ST405 retained *iha* and *sat* genes belonging to FimH 27 type within ST405. This was followed by ST410 (FimH 24) that predominantly had *ipfA* gene and lost all other genes, except few. Overall, gad gene served as backbone for ST167, 131 and ST405 clades, while *ipfA* for ST410. Similar studies comparing the evolution of virulence pattern with phylogeny is still lacking.

FimH had been reported as a major candidate for the development of a vaccine against pathogenic *E. coli* [44] which was responsible for mannose-sensitive bacterial adhesion [45]. Though high nucleotide conservation of >98% was observed in *fimH* alleles, minor sequence differences were reported to correlate with differential binding and adhesion phenotypes [45].

Fim-H types correlated with the STs observed in the study isolates. FimH 27 and 29 type grouped with ST405 clade. The ST405 clade with 27 and 29 fimH types would have possibly emerged from the ST131 clade with Fim-H type 24 and 30. However, another clade with FimH 24 belonged to ST410 with CC10. Similar study comparing the MLST and fim-H types were previously reported by Dreux et al. [46] in Adherent-Invasive *E. coli*.

Moreover, concatenated SNPs based phylogeny revealed higher discrimination between the clinical MDR *E. coli* isolates and disclosed the evolutionary pattern with accumulated SNPs. Among the study isolates, even within the clones, evidence of evolution is seen with the difference in root to tip. This suggests continuous evolution leading to diversified MDR *E. coli* strains in India.

Conclusions:

To the best of our knowledge, this is the first report on SNP phylogeny in comparison with AMR and virulence traits in *E. coli* in India. The study revealed the prevalence of *bla*$_{\text{NDM-5}}$ among the clades ST131, ST405 and ST410 clades. *bla*$_{\text{CTX-M-15}}$ was responsible for cephalosporin resistance in ST131 and ST405 clades whereas, ST167 and ST410 carried both *bla*$_{\text{CTX-M-15}}$ and *bla*$_{\text{CMY}}$ genes. For aminoglycoside resistance, *rmtB* was positive for 31.7% of the isolates, of which 30% were co-harboring carbapenemases. The FimH types and virulence gene profile positively correlated with the SNP based phylogeny. However the predominant ST131 epidemic clone was smaller the study population while ST167 and ST405 clones with multiple AMR genes were predominant. Isolates with *iss*, *capU* and *gad* virulence genes were the major type. Moreover, SNP based phylogeny revealed the evolution of the MDR clones among the study population with the accumulation of SNPs, which suggests continuous molecular surveillance to understand the spread of MDR clones in India.

References:

1. Kennedy KJ, Roberts JL, Collignon PJ. Escherichia coli bacteraemia in Canberra: incidence and clinical features. Med J Aust. 2008;188:209-13.
2. López-Banda DA, Carrillo-Casas EM, Leyva-Leyva M, Orozco-Hoyuela G, Manjarrez-Hernández ÁH, Arroyo-Escalante S, et al. Identification of Virulence Factors Genes in Escherichia coli Isolates from Women with Urinary Tract Infection in Mexico. BioMed Research International, Volume 2014, Article ID 959206, 10 pages. http://dx.doi.org/10.1155/2014/959206
3. Johnson JR, Sannes MR, Croy C, Johnston B, Clabots C, Kuskowski MA, et al. Antimicrobial drug-resistant *Escherichia coli* from humans and poultry products, Minnesota and Wisconsin, 2002–2004. Emerg Infect Dis. 2007;13:838-46.
4. Mesa RJ, Blanc V, Blanch AR, Cortés P, González JJ, Lavilla S *et al*. Extended-spectrum β-lactamase-producing Enterobacteriaceae in different environments (humans, food, animal farms, and sewage). J Antimicrob Chemother. 2006;58:211-5.
5. Laupland KB, Church DL, Vidakovich J, Mucenski M, Pitout JD. Community-onset extended-spectrum β-lactamase (ESBL)-producing *Escherichia coli*: importance of international travel. J Infect. 2008;57:441-448.

6. Nicolas-Chanoine MH, Blanco J, Leflon-Guibout V, Demarty R, Alonso MP, Canica MM. Intercontinental emergence of *Escherichia coli* clone O25:H4–ST131 producing CTX-M-15. J Antimicrob Chemother. 2008;61:273–81.
7. Lau SH, Kaufmann ME, Livermore DM, Woodford N, Willshaw GA, Cheasty T. UK epidemic *Escherichia coli* strains A–E, with CTX-M-15 β-lactamase, all belong to the international O25:H4–ST131 clone. J Antimicrob Chemother. 2008;62:1241–4.
8. Alhashash F, Weston V, Diggle M, McNally A. Multidrug-Resistant Escherichia coli Bacteremia. Emerg Infect Dis. 2013;19(10):1699-701.
9. Candan ED, Aksöz N. *Escherichia Coli*: Characteristics of Carbapenem Resistance and Virulence Factors. Braz arch biol technol. 2017;60.
10. S Nagaraj, SP Chandran, P Shamanna, R Macaden. Carbapenem resistance among *Escherichia coli* and *Klebsiella pneumoniae* in a tertiary care hospital in south India. Indian J Med Microbiol. 2012;30(1):93-5.
11. Kanamori H1,2, Parobek CM3, Juliano JJ4, Johnson JR5,6, Johnston BD5,6, Johnson TJ7, Weber DJ4,2, Rutala WA4,2, Anderson DJ. Genomic Analysis of Multidrug-Resistant Escherichia coli from North Carolina Community Hospitals: Ongoing Circulation of CTX-M-Producing ST131-H30Rx and ST131-H30R1 Strains. Antimicrob Agents Chemother. 2017;61(8):pii: e00912-17.
12. Abbott SL, Cheung WKW, Janda JM. The genus Aeromonas: biochemical characteristics, atypical reactions, and phenotypic identification schemes. J Clin Microbiol. 2003;41:2348–57.
13. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. PLoS ONE 2014;9(8):e104984.

14. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.

15. Veeraraghavan B, Jesudason MR, Prakasah JAJ, Anandan S, Sahni RD, Pragasam AK, Bakthavatchalam YD, Selvakumar RJ, Dhole TN, Rodrigues C, Roy I, Joshi S, Chaudhuri BN, Chitnis DS. Antimicrobial susceptibility profiles of gram-negative bacteria causing infections collected across India during 2014–2016: Study for monitoring antimicrobial resistance trend report. Indian J Med Microbiol. 2018;36:32-6.

16. Ragupathi NKD, Bakthavatchalam YD, Mathur P, Pragasam AK, Walia K, Ohri VC, Veeraraghavan B. Plasmid profiles among some ESKAPE pathogens in a tertiary care centre in south India. Indian J Med Res. 2019;149(2):222-31.

17. Ho PL, Cheung YY, Wang Y, Lo WU, Lai EL, Chow KH, et al. Characterization of carbapenem-resistant Escherichia coli and Klebsiella pneumoniae from a healthcare region in Hong Kong. Eur J Clin Microbiol Infect Dis. 2016;35(3):379-85.

18. Liang WJ, Liu HY, Duan GC, Zhao YX, Chen SY, Yang HY, et al. Emergence and mechanism of carbapenem-resistant *Escherichia coli* in Henan, China, 2014. J Infect Public Health. 2018;11(3):347-51.

19. Poirel L, Yilmaz M, Istanbullu A, Arslan F, Mert A. Spread of NDM-1-producing Enterobacteriaceae in a neonatal intensive care unit, Istanbul, Turkey. Antimicrob Agents Chemother. 2014;58:2929-33.

20. Mushtaq S, Irfan S, Sarma JB, Doumith M, Pike R, Pitout J et al. Phylogenetic diversity of Escherichia coli strains producing NDM-type carbapenemases. J Antimicrob Chemother. 2011;66:2002-5.

21. Aktas Z, Satana D, Kayacan C, Ozbek B, Gurler N, Somer A et al. Carbapenem resistance in Turkey: Repeat report on OXA-48 in Klebsiella pneumoniae and first report on IMP-1 beta-lactamase in Escherichia coli. Afr J Microbiol Res. 2012;6:3874-8.

22. Morris D, Boyle F, Ludden C, Condon I, Hale J, O'Connell N et al. Production of KPC-2 Carbapenemase by an Escherichia coli Clinical Isolate Belonging to the International ST131 Clone. Antimicrob Agents Chemother. 2011;55:4935-6.

23. Budak S, Oncul O, Aktas Z, Acar A, Ozyurt M, Turhan V, et al. The determination of carbapenem resistance in Escherichia coli and Pneumoniae isolates related to nosocomial infections and the evaluation of risk factors. Southeast Asian J Trop Med Public Health. 2014;45(1):113-22.

24. Sharma A, Bakthavatchalam YD, Gopi R, Anandan S, Verghese VP, Veeraraghavan B. Mechanisms of Carbapenem Resistance in K. pneumoniae and E. coli from Bloodstream Infections in India. J Infect Dis Ther. 2016;4:293.

25. Doi Y, Wachino JI, Arakawa Y. Aminoglycoside Resistance: The Emergence of Acquired 16S Ribosomal RNA Methyltransferases. Infect Dis Clin North Am. 2016; 30(2): 523–37.

26. Yu F, Wang L, Pan J, Yao D, Chen C, Zhu T, et al. Prevalence of 16S rRNA methylase genes in *Klebsiella pneumoniae* isolates from a Chinese teaching hospital: coexistence of *rmtB* and *armA* genes in the same isolate. Diagn Microbiol Infect Dis. 2009;64:57–63.

27. Taylor E, Sriskandan S, Woodford N, Hopkins KL. High prevalence of 16S rRNA methyltransferases among carbapenemase-producing Enterobacteriaceae in the UK and Ireland. Int J Antimicrob Agents. 2018; 52(2):278-82.

28. Poirel L, Goutines J, Aires-de-Sousa M, Nordmann P. High Rate of Association of 16S rRNA Methylases and Carbapenemases in Enterobacteriaceae Recovered from Hospitalized Children in Angola. Antimicrob Agents Chemother. 2018;62:e00021-18.
29. Ravikant, Kumar P, Ranotkar S, Zutshi S, Lahkar M, Phukan C, et al. Prevalence and identification of extended spectrum β-lactamases (ESBL) in *Escherichia coli* isolated from a tertiary care hospital in North-East India. Ind J Exp Biol. 2016;54:108-114.
30. Elnahriry SS, Khalifa HO, Soliman AM, Ahmed AM, Hussein AM, Shimamoto T, Shimamoto T. Emergence of Plasmid-Mediated Colistin Resistance Gene mcr-1 in a Clinical Escherichia coli Isolate from Egypt. Antimicrob Agents Chemother. 2016;60 (5):3249-50.
31. Ghafur A, Shankar C, GnanaSoundari P, Venkatesan M, Mani D, Thirunarayanan MA, Veeraraghavan B. Detection of chromosomal and plasmid-mediated mechanisms of colistin resistance in Escherichia coli and Klebsiella pneumoniae from Indian food samples. J Glob Antimicrob Resist. 2019;16:48-52.
32. Yang F, Shen C, Zheng X, Liu Y, El-Sayed Ahmed MAE, Zhao Z, et al. Plasmid-mediated colistin resistance gene mcr-1 in *Escherichia coli* and *Klebsiella pneumoniae* isolated from market retail fruits in Guangzhou, China. Infect Drug Resist. 2019;12:385-9.
33. Ikram R, Psutka R, Carter A, Priest P. An outbreak of multi-drug resistant Escherichia coli urinary tract infection in an elderly population: a case-control study of risk factors. BMC Infect Dis. 2015;15:224.
34. Severin JA, Mertaniasih NM, Kuntaman K, Lestari ES, Purwanta M, Lemmens-Den Toom N, et al. Molecular characterization of extended-spectrum β-lactamases in clinical *Escherichia coli* and *Klebsiella pneumoniae* isolates from Surabaya, Indonesia. J Antimicrob Chemother. 2010;65:465-9.

35. Ewers C, Grobbel M, Stamm I, Kopp PA, Diehl I, Semmler T, et al. Emergence of human pandemic O25:H4-ST131 CTX-M-15 extended-spectrum-β-lactamase-producing *Escherichia coli* among companion animals. J Antimicrob Chemother. 2010;65:651-60.

36. Brolund A, Edquist PJ, Makitalo B, Olsson-Liljequist B, Soderblom T, Wisell KT, et al. Epidemiology of extended-spectrum b-lactamase-producing *Escherichia coli* in Sweden 2007–2011. Clin Microbiol Infect. 2014;20:O344–O352.

37. Yang P, Xie Y, Feng P, Zong Z. *bla*NDM-5 Carried by an IncX3 Plasmid in *Escherichia coli* Sequence Type 167. Antimicrob Agents Chemother. 2014;58(12):7548–52.

38. Feng Y, Yang P, Xie Y, Wang X, McNally A, Zong Z. *Escherichia coli* of sequence type 3835 carrying *bla*NDM-1, *bla*CTX-M-15, *bla*CMY-42 and *bla*SHV-12. Sci Rep. 2015;5:12275.

39. Shen P, Yi M, Fu Y, Ruan Z, Du X, Yu Y, et al. Detection of an *Escherichia coli* Sequence Type 167 Strain with Two Tandem Copies of *bla*NDM-1 in the Chromosome. J Clin Microbiol. 2017;55:199–205.

40. 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-_-Lactamase-Producing *Escherichia coli* Sequence Type 131 (ST131) and ST405 Clonal Groups. Antimicrob Agents Chemother. 2013;57(10):4736–42.

41. Zhang X, Feng Y, Zhou W, McNally A, Zong Z. Cryptic transmission of ST405 *Escherichia coli* carrying *bla*NDM-4 in hospital. Sci Rep. 2018;8:390.

42. Cai JC, Zhang R, Hu YY, Zhou HW, Chen GX. Emergence of *Escherichia coli* Sequence Type 131 Isolates Producing KPC-2 Carbapenemase in China. Antimicrob Agents Chemother. 2014;58(2):1146–52.

43. Roer L, Overballe-Petersen S, Hansen F, Schønning K, Wang M, Røder BL, et al. *Escherichia coli* Sequence Type 410 Is Causing New International High-Risk Clones. mSphere. 2018;3(4):pii: e00337-18.

44. Tchesnokova V, Aprikian P, Kisiela D, Gowey S, Korotkova N, Thomas W *et al.* Type 1 Fimbrial Adhesin FimH Elicits an Immune Response That Enhances Cell Adhesion of *Escherichia coli.* Infect. Immun. 2011;79(10):3895–3904.

45. Bouckaert J, Mackenzie J, de Paz JL, Chipwaza B, Choudhury D, Zavialov A *et al.* The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of Escherichia coli pathotypes. Mol Microbiol. 2006;61(6):1556–1568.

46. Dreux N, Denizot J, Martinez-Medina M, Mellmann A, Billig M, Kisiela D *et al.* Point Mutations in FimH Adhesin of Crohn's Disease-Associated Adherent-Invasive *Escherichia coli* Enhance Intestinal Inflammatory Response. PLoS Pathog. 2013;9(1):e1003141.

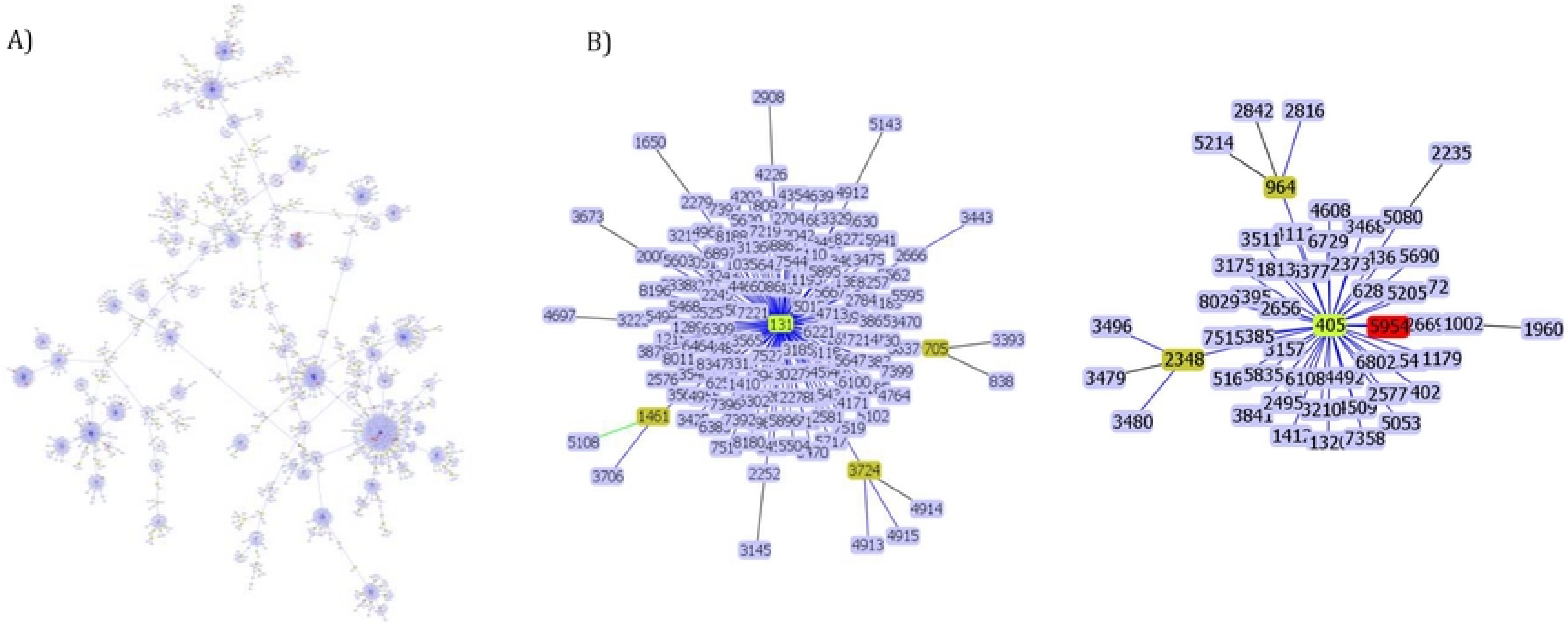
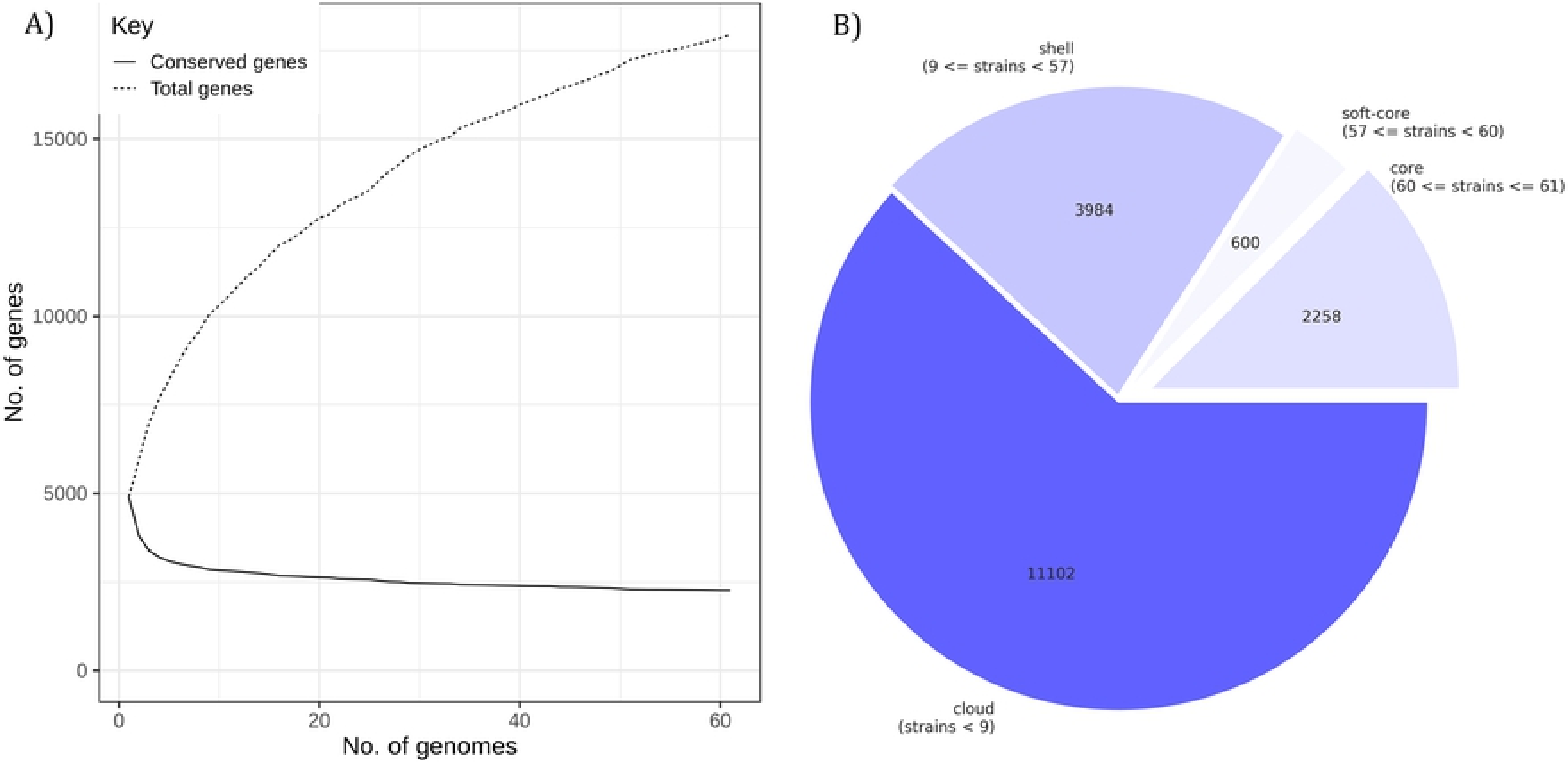

A)
B)
2908
1650
5143
4226
3673
2279
4912
3443
2666
8196
4697
131
705
3393
838
1461
5108
3706
3724
2252
4914
4915
4913
3145
2842
2816
5214
2235
964
4608
5080
3511
6729
3175
5690
8029
405
5954
1002
1960
3496
2348
3479
3480
7515385
3157
6802
1179
2577
402
5053
7358
3841

Figure 1



A)
Key
Conserved genes
Total genes
No. of genes
15000
10000
5000
0
0
20
40
60
No. of genomes
B)
shell
(9 <= strains < 57)
soft-core
(57 <= strains < 60)
core
(60 <= strains <= 61)
cloud
(strains < 9)
3984
600
2258
11102

Figure 2

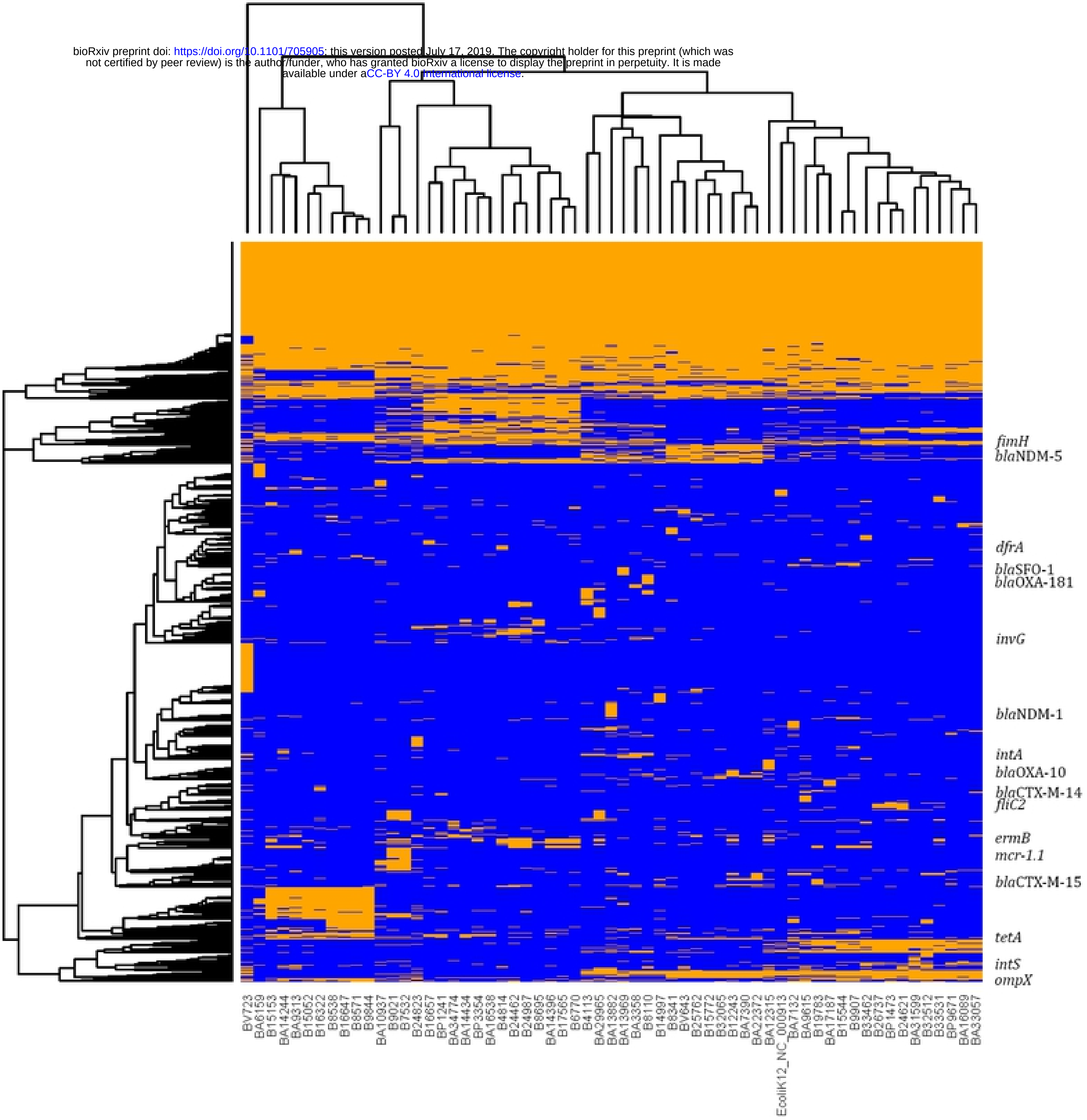

fimH
blaNDM-5
dfrA
blaSFO-1
blaOXA-181
invG
blaNDM-1
intA
blaOXA-10
blaCTX-M-14
fliC2
ermB
mcr-1.1
blaCTX-M-15
tetA
intS
ompX
BV723
BA6159
B15153
BA14244
BA9313
B15052
B16322
B8538
B16647
B8571
B9844
BA10937
B9021
B7532
B24823
B16657
BP1241
BA34774
BA14434
BP3354
BA16538
B4814
B24462
B24987
B8695
BA14396
B17565
B6770
B4113
BA29965
BA13882
BA13969
BA3358
B8110
B14997
B8341
BV643
B25762
B15772
B32065
B12243
BA7390
BA22372
BA12315
EcoliK12_NC_000913
BA7132
BA9615
B19783
BA17187
B15544
B9907
B33462
B26737
BP1473
B24621
BA31599
B32512
B33531
BP9671
BA16089
BA33057

Figure 3

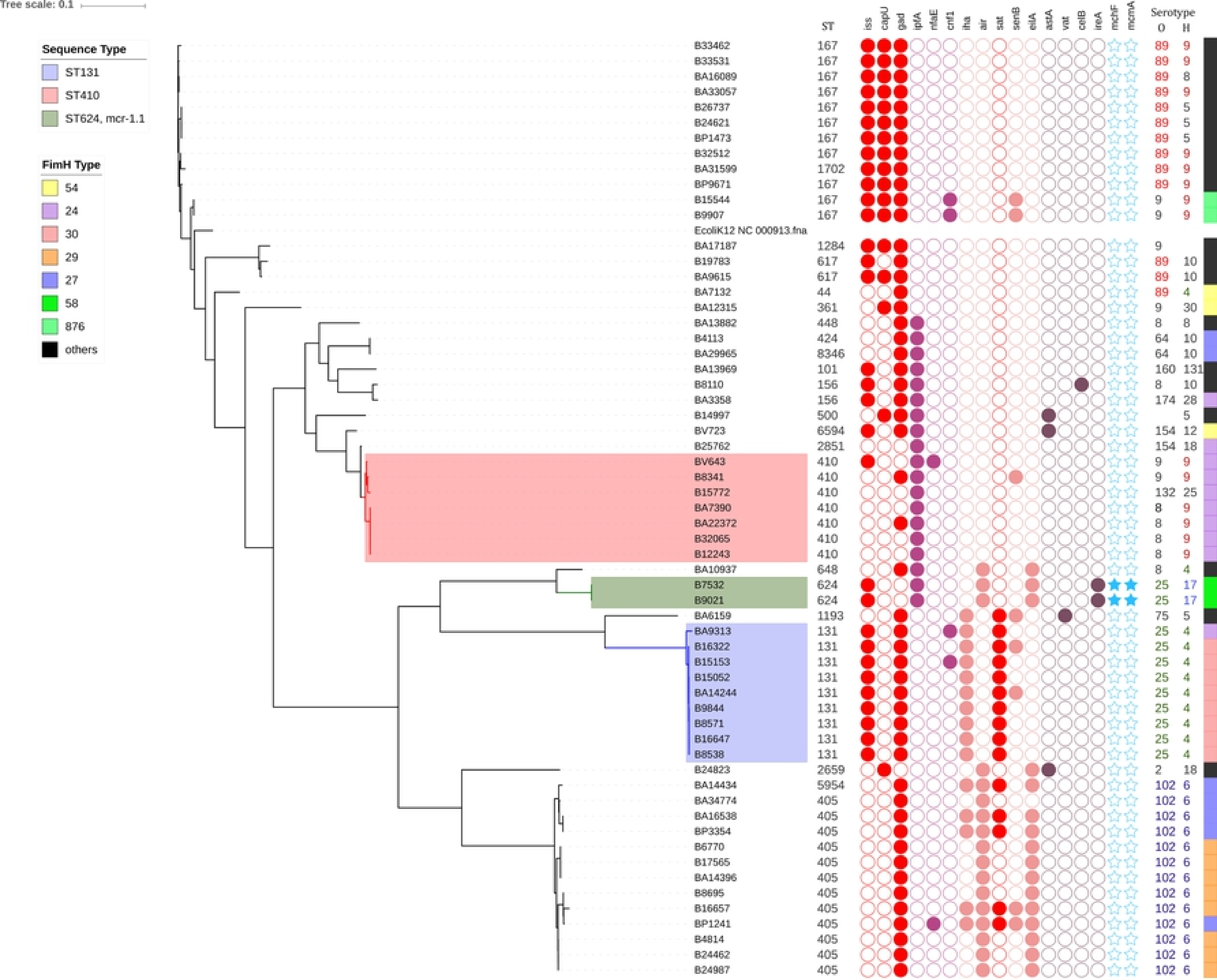

Tree scale: 0.1
Sequence Type
ST131
ST410
ST624, mcr-1.1
FimH Type
54
24
30
29
27
58
876
others
ST
iss
capU
gad
iutA
nfaE
cnf1
iha
air
sat
senB
eilA
astA
vat
celB
ireA
mchF
mcmA
Serotype
O
H
B33462 167 89 9
B33531 167 89 9
BA16089 167 89 8
BA33057 167 89 9
B26737 167 89 5
B24621 167 89 5
BP1473 167 89 5
B32512 167 89 9
BA31599 1702 89 9
BP9671 167 89 9
B15544 167 9 9
B9907 167 9 9
EcoliK12 NC 000913.fna
BA17187 1284 9
B19783 617 89 10
BA9615 617 89 10
BA7132 44 89 4
BA12315 361 9 30
BA13882 448 8 8
B4113 424 64 10
BA29965 8346 64 10
BA13969 101 160 131
B8110 156 8 10
BA3358 156 174 28
B14997 500 5
BV723 6594 154 12
B25762 2851 154 18
BV643 410 9 9
B8341 410 9 9
B15772 410 132 25
BA7390 410 8 9
BA22372 410 8 9
B32065 410 8 9
B12243 410 8 9
BA10937 648 8 4
B7532 624 25 17
B9021 624 25 17
BA6159 1193 75 5
BA9313 131 25 4
B16322 131 25 4
B15153 131 25 4
B15052 131 25 4
BA14244 131 25 4
B9844 131 25 4
B8571 131 25 4
B16647 131 25 4
B8538 131 25 4
B24823 2659 2 18
BA14434 5954 102 6
BA34774 405 102 6
BA16538 405 102 6
BP3354 405 102 6
B6770 405 102 6
B17565 405 102 6
BA14396 405 102 6
B8695 405 102 6
B16657 405 102 6
BP1241 405 102 6
B4814 405 102 6
B24462 405 102 6
B24987 405 102 6

Figure 4

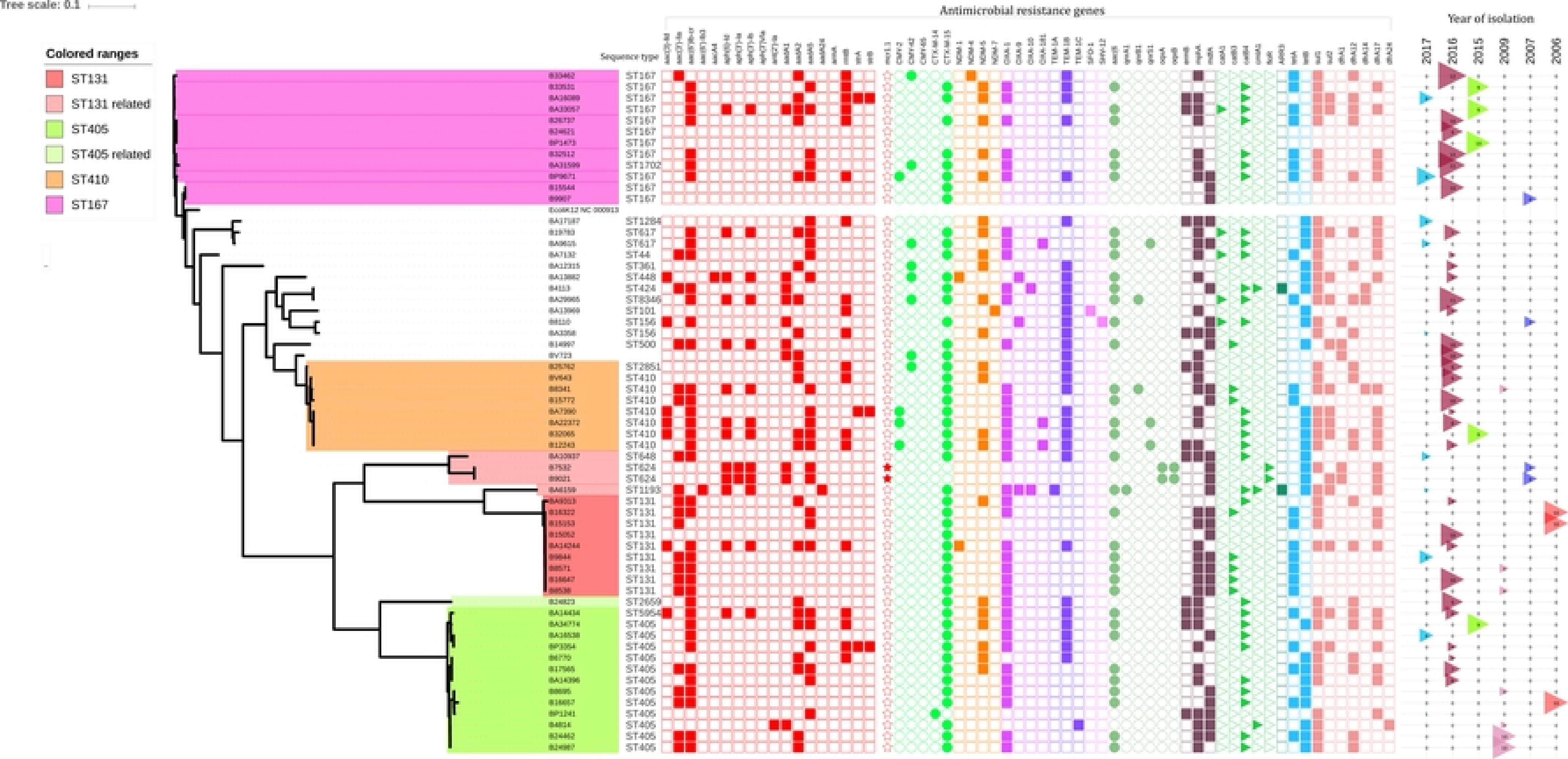

Tree scale: 0.1
Colored ranges
ST131
ST131 related
ST405
ST405 related
ST410
ST167
Sequence type
Antimicrobial resistance genes
Year of isolation
2017
2016
2015
2009
2007
2006
ST167
ST167
ST167
ST167
ST167
ST167
ST167
ST167
ST1702
ST167
ST167
ST167
ST1284
ST617
ST617
ST44
ST361
ST448
ST424
ST8346
ST101
ST156
ST156
ST500
ST2851
ST410
ST410
ST410
ST410
ST410
ST410
ST410
ST648
ST624
ST624
ST1193
ST131
ST131
ST131
ST131
ST131
ST131
ST131
ST131
ST131
ST2659
ST5954
ST405
ST405
ST405
ST405
ST405
ST405
ST405
ST405
ST405
ST405
ST405
ST405

Figure 5