Virulence and resistance properties of E. coli isolated from urine samples of hospitalized patients in Rio de Janeiro, Brazil - The role of mobile genetic elements

SOLIDNESS Working Group; da Cruz Campos, Ana Carolina; Couto, Natacha; Lucas da Silva Andrade, Nathália; Friedrich, Alex W; de Paula Rosa, Ana Cláudia; Vieira Damasco, Paulo; Chlebowicz-Fliss, Monika A; Rossen, John W A

Published in:
International journal of medical microbiology

DOI:
10.1016/j.ijmm.2020.151453

Publication date:
2020

International Journal of Medical Microbiology 310 (2020) 151453

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

Virulence and resistance properties of *E. coli* isolated from urine samples of hospitalized patients in Rio de Janeiro, Brazil – The role of mobile genetic elements

Ana Carolina da Cruz Campos [a,b,*], Natacha Couto [b], Nathália Lucas da Silva Andrade [a], Alex W. Friedrich [b], Ana Cláudia de Paula Rosa [a], Paulo Vieira Damasco [c,d], Monika A. Chlebowicz-Fliss [b], John W.A. Rossen [b,e,f], the SOLIDNESS Working Group [1]

[a] Universidade do Estado do Rio de Janeiro, Faculdade de Ciências Médicas, Departamento de Microbiologia, Imunologia e Parasitologia, Brasil. Blv. 28 de Setembro, 87 fundos 3° andar, 20550-170, Vila Isabel, Rio de Janeiro, Brazil
[b] University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and Infection Prevention, Hanzeplein 1, 9713GZ, Groningen, The Netherlands
[c] Universidade Federal do Estado do Rio de Janeiro, Departamento de Medicina Geral, Rua Mariz e Barros, 775, Tijuca, Rio de Janeiro, 22270-004, Brazil
[d] Universidade do Estado do Rio de Janeiro, Departamento de Doenças Infecciosas e Parasitárias Brasil. Blv 28 de Setembro, 87 fundos 3° andar, 20550-170, Vila Isabel, Rio de Janeiro, Brazil
[e] ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD), Basel, Switzerland
[f] University of Utah School of Medicine, Department of Pathology, 30 N. 1900 E, Salt Lake City, UT, 84132, USA

ARTICLE INFO

Keywords:
Plasmids
Pathogenicity islands
Genomic islands
E. coli
ST131
UTIs

ABSTRACT

Extraintestinal pathogenic *E. coli* (ExPEC) is the most frequent etiological agent of urinary tract infections (UTIs). Particular evolutionary successful lineages are associated with severe UTIs and higher incidences of multidrug resistance. Most of the resistance genes are acquired by horizontal transfer of plasmids and other mobile genetic elements (MGEs), and this process has been associated with the successful dissemination of particular lineages. Here, we identified the presence of MGEs and their role in virulence and resistance profiles of isolates obtained from the urine of hospitalized patients in Brazil. Isolates belonging to the successful evolutionary lineages of sequence type (ST) 131, ST405, and ST648 were found to be multidrug-resistant, while those belonging to ST69 and ST73 were often not. Among the ST131, ST405, and ST648 isolates with a resistant phenotype, a high number of mainly IncFII plasmids was identified. The plasmids contained resistance cassettes, and these were also found within phage-related sequences and the chromosome of the isolates. The resistance cassettes were found to harbor several resistance genes, including $bla_{CTX\text{-}M\text{-}15}$. In addition, in ST131 isolates, diverse pathogenicity islands similar to those found in highly virulent ST73 isolates were detected. Also, a new genomic island associated with several virulence genes was identified in ST69 and ST131 isolates. In addition, several other MGEs present in the ST131 reference strain EC958 were identified in our isolates, most of them exclusively in ST131 isolates. In contrast, genomic islands present in this reference strain were only partially present or completely absent in our ST131 isolates. Of all isolates studied, ST73 and ST131 isolates had the most similar virulence profile. Overall, no clear association was found between the presence of specific MGEs and virulence profiles. Furthermore, the interplay between virulence and resistance by acquiring MGEs seemed to be lineage

Abbreviations: ATCC, American type culture collection; BRIG, blast ring image generator; BR, Broad-Range; CGE, center for genomic epidemiology; DNA, deoxyribonucleic acid; dsDNA, double-stranded Deoxyribonucleic acid; ESBL, extended-spectrum beta-lactamases; GI, genomic islands; HS, high sensitivity; HPI, high-pathogenicity island; ST, sequence type; LB, Luria Bertani; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight analyzer; MDR, multidrug resistance; MGE, mobile genetic elements; PAI, pathogenicity islands; PAIDB, pathogenicity islands database; PBS, phosphate-buffered saline; pMLST, plasmid multilocus sequence typing; TSA, trypticase soy agar; UTIs, urinary tract infections; VFDB, virulence factor database.

* Corresponding author at: Department of Medical Microbiology and Infection Prevention, University of Groningen, University Medical Center Groningen, EB80 Hanzeplein 1, 9713 GZ, Groningen, The Netherlands.

E-mail addresses: anabio86@gmail.com (A.C. da Cruz Campos), n.monge.gomes.do.couto@umcg.nl (N. Couto), nathalu84@yahoo.com (N. Lucas da Silva Andrade), alex.friedrich@umcg.nl (A.W. Friedrich), anarosa2004@gmail.com (A.C. de Paula Rosa), paulovieiradamasco@gmail.com (P. Vieira Damasco), m.a.chlebowicz@umcg.nl (M.A. Chlebowicz-Fliss).

[1] A list of members appears in the Supplementary Note.

https://doi.org/10.1016/j.ijmm.2020.151453
Received 15 January 2020; Received in revised form 15 August 2020; Accepted 25 September 2020
Available online 28 September 2020

dependent. Although the acquisition of IncF plasmids, specific PAIs, GIs, and other MGEs seemed to be involved in the success of some lineages, it cannot explain the success of different lineages, also indicating other (host) factors are involved in this process. Nevertheless, the detection, identification, and surveillance of lineage-specific MGEs may be useful to monitor (new) emerging clones.

1. Introduction

Extraintestinal pathogenic *Escherichia coli* (ExPEC), particularly uropathogenic *E. coli* (UPEC), is an important etiological agent of urinary tract infections (UTIs) (Calhau et al., 2015). Their treatment became more complicated as regular antibiotic regimens fail due to the increase of antibiotic resistance, mainly to cephalosporins, aminoglycosides, and fluoroquinolones (Zowawi et al., 2015). This worldwide increase is associated with specific evolutionary successful lineages, such as the often multidrug-resistant (MDR), sequence type (ST) 131 clone (Mathers et al., 2015; Price et al., 2013). Isolates of this lineage often belong to phylogenetic group B2 and serotype O25b:H4 and possess a variable combination and number of virulence genes, which may be responsible for their virulence potential and successful global spread (Dahbi et al., 2013; Peirano and Pitout, 2014).

Within the ST131 lineage, some sublineages are often classified according to their adhesive subunit of type 1 fimbriae (*fimH*) sequence, like ST131 *fimH*30. *E. coli* belonging to this sublineage are usually referred to as H30-R when being resistant to fluoroquinolones and H30-Rx when they are also CTX-M-15-producing (Dahbi et al., 2014). MDR-ExPEC clones are often associated with plasmid-based Extended-Spectrum Beta-Lactamase (ESBL) genes, such as the frequently detected CTX-M-15 gene (Johnson et al., 2012). Recently, the classification of ST131 isolates in virotypes according to the presence of specific virulence genes was introduced, virotype C being the most prevalent among ESBL-producing ST131 *E. coli* (Miyoshi-akiyama et al., 2016). Other evolutionary successful lineages, as ST69, ST73, ST405, and ST648, have been associated with an enhanced ability to colonize and persist in the host. Like ST131, also ST405 and ST648 are considered high-risk lineages as they are associated with the spread of resistance genes.

Evolutionary successful lineages contain a variety of virulence genes such as adhesins, siderophores, and toxins that are important to overcome the host defense, for colonization and in further steps during the infection cycle (Gazal et al., 2015; Guidoni et al., 2006; Koga et al., 2014; Zaw et al., 2013). Most of these virulence genes are present on pathogenicity islands (PAIs) (Gazal et al., 2015), mobile genetic elements (MGEs) that can be transmitted horizontally even between species (Johnson and Stell, 2000). PAIs are a subset of genomic islands (GIs) and are characterized by the presence of virulence genes, a biased G + C content and codon usage, carriage of mobile sequence elements and are flanked by direct repeats and tRNA-encoding genes, which together define the area where insertion of foreign DNA can occur. MGEs are not only associated with virulence but also with resistance profiles. Among the different MGEs, plasmids are considered to be the most important for the spread of antibiotic resistance (Hall et al., 2017; Song et al., 2018), and they play an essential role in the evolution of bacterial clones (Downing, 2015; Lyimo et al., 2016). The acquirement of plasmids may also increase the virulence of bacteria (Koga et al., 2014). Indeed, there is an association between high-risk lineages and IncF plasmids, particularly those containing replicon types FIA and FII. Therefore, we characterized the presence of MGEs in *E. coli* isolates from successful evolutionary lineages isolated from the urine of hospitalized patients in Rio de Janeiro, Brazil. Furthermore, we aimed at characterizing the role of these MGEs in the antimicrobial resistance and virulence properties of ST131 and non-ST131 isolates.

2. Material and methods

2.1. Bacterial isolates, whole-genome sequencing, assembly and annotation

E. coli isolates used in this study were collected from urine samples of hospitalized patients from four different hospitals located in Rio de Janeiro, Brazil (Campos et al., 2018). Antibiotic susceptibility testing was performed using the VITEK-2 (bioMérieux, Marcy L'Etoile, France), and isolates resistant to three or more antibiotic classes were classified as multidrug-resistant (Campos et al., 2018). After the identification using MALDI-TOF, DNA was extracted using the Ultraclean Microbial DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, US). Libraries were prepared using the Nextera XT library preparation kit with the Nextera XT v2 index kit and sequenced on a MiSeq sequencer, using the MiSeq reagent kit v2 generating 250-bp paired-end reads according to the manufacturer's protocols (Illumina, San Diego, CA, USA). Quality trimming of reads was performed with CLC Genomics Workbench v10.0.1 (Qiagen, CLC bio A/S, Aarhus, Denmark) using a minimum Phred (Q) score of 28. *De novo* assembly was performed using CLC Genomics Workbench v10.0.1 (Qiagen, CLC bio A/S, Aarhus, Denmark) using default settings and optimal word sizes based on the maximum N50 value (the largest scaffold length, N, such that 50 % of the assembled genome size is contained in scaffolds with a length of at least N). Annotation was performed by uploading the assembled genomes onto the RAST server version 2.0 (Aziz et al., 2008). The phylogenetic types were determined as described by Clermont et al. (2013). The sequences of all isolates are available in the ENA database (project number: PRJEB23420).

2.2. Long-read whole genome sequencing

To characterized the newly identified GI, named GI—II, in more detail, we also sequenced ST131 isolates 2724, 3218, 5770D and 5848, and ST69 isolates 108, 605, 2441, 2445 and 4953 using long-read sequencing on a MinION device (Oxford Nanopore Technologies [ONT], Oxford, United Kingdom). For this, total DNA was extracted using the DNeasy Ultraclean Microbial kit (Qiagen, Hilden, Germany). The DNA concentration and purity were measured using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the Qubit double-stranded DNA (dsDNA) BR assay kits (Life Technologies, Carlsbad, CA, USA). The library of the strain 5848 was prepared using the 2D sequencing kit (SQK-) while the other strains were made using the Rapid Sequencing kit (SQK-RAD004) according to the manufacturer's instructions (Oxford Nanopore, Oxford, UK). The final library was assessed with the Genomic DNA ScreenTape assay with the 2200 TapeStation system (Agilent Technologies, Waldbronn, Germany) before loading it onto FLO-MIN106 R9.4 flow cells. For the next steps, default settings for the different software tools were used unless indicated otherwise. Base-calling was performed using either Albacore v2.0 or Guppy v3.2.2 (ONT). The quality of the data was analyzed through Poretools v0.6.0 (Loman and Quinlan, 2014). A hybrid assembly using Illumina short reads and ONT long reads was performed using Unicycler v0.4.1 (Wick et al., 2017). Bandage v0.8.1 (Wick et al., 2015) was used to visualize the assembly graphs.

2.3. Plasmid identification and analysis of virulence and resistance genes

Fasta files were uploaded to the PlasmidFinder (v2.0) tool (CGE)

(Carattoli et al., 2014) to identify the plasmids' replicon sequence and incompatibility types. Replicon sequences with at least 90 % coverage and 99 % identity were considered to be present in the isolates. Plasmid multi-locus sequence typing (pMLST) was performed by uploading the fasta files to the pMLST (v2.0) tool (CGE) (Carattoli et al., 2014). We also predicted the plasmid-derived sequences by uploading the fasta file to the online mlplasmids tool (Arredondo-Alonso et al., 2018) and MOB-suite tool (Robertson and Nash, 2018). The contigs with a score higher than 0.60 using mlplasmids and the contigs identified as plasmid-derived sequences by the MOB-suite software were further analyzed using the Artemis tools and the Basic Local Alignment Search Tool (BLAST, v.2.7.1 NCBI). Fasta files containing just the predicted sequences were constructed and uploaded to the VirulenceFinder tool (v.2.0) (CGE) (Joensen et al., 2015) to identify the virulence genes, present in the plasmid-derived sequences. Also, to identify the resistance genes present on plasmids, these sequences were uploaded to the ResFinder tool (3.1) (CGE) (Zankari et al., 2012), and heat-maps were generated using the GraphPrism software (v.7.04) (GraphPad Software, La Jolla, US). Resistance cassettes were identified by blasting the sequence data of the isolates against the NCBI database, and finally, the alignments were performed using ACT (Carver et al., 2005).

2.4. Detection of genomic and pathogenicity islands

The PAIs were identified by aligning the isolate sequences against a PAI database containing complete sequences of PAIs commonly found in *E. coli*, i.e., PAI I_{536}, PAI II_{536}, PAI III_{536}, PAI IV_{536} and PAI I_{CFT073} downloaded from the NCBI database and PAI sequences available in the PAIDB (Yoon et al., 2005) (see supplementary data S.1 for reference sequences). Identification of the in reference strain EC958 present PAIs, GIs, prophages, HPIs, ROD insertion elements, *ratA*-like toxin, Flag-2 lateral flagellar locus, and Type VI secretion system, in our isolates, was performed using the Basic Local Alignment Search Tool (BLAST) v.2.7.1. The elements with coverage between 40 %–89 % and > 90 % identity were classified as partially present, while the elements with > 90 % coverage and > 90 % identity were classified as present. Elements with < 40 % coverage and/or < 90 % identity were considered absent. The alignment of the new GI named GI-II in this study with GI CP023826 was visualized using the EasyFig tool (Sullivan et al., 2011).

2.5. Statistical tests

The Fisher's exact test was used to evaluate the association between the presence of IncF plasmids and the multidrug-resistant profile using GraphPad Prism v.7.0.3 (GraphPad Software, La Jolla, US). Values of $P < 0.05$ were considered statistically significant.

3. Results

3.1. Bacterial isolates

In total, 49 *E. coli* isolates belonging to phylogenetic groups B2 and D obtained from urine samples of hospitalized patients admitted to four different hospitals in Rio de Janeiro were analyzed in this study. The presence of MGEs was determined in 26 isolates belonging to ST131, of which 24 belonged to serotype O25:H4 and the other two to serotype O16:H5. Based on *fimH* typing, the majority of ST131 isolates were *fimH*30 (n = 22), while two isolates were *fimH*22, and another two isolates were *fimH*41. The ST131 isolates presented different virotypes, namely A, B, C, C2, C3, and D. The non-ST131 isolates used in this study belonged to ST69 (n = 9), ST73 (n = 4), ST405 (n = 4) and ST648 (n = 6). ST69 and ST73 isolates were not associated with MDR, whereas ST405 and ST648 isolates often were MDR (including ESBL-producing and carbapenem-producing ones). The isolates investigated in this study belonged to phylogenetic group B2 (n = 36) and D (n = 13), as we have shown before (Campos et al., 2018). For the analyses of the PAIs, we included 18 additional isolates belonging to phylogenetic groups A and B1 and collected in the same period (Supplementary data S.2).

3.2. Plasmid analyses

IncF plasmids were present in all ST131 isolates, the majority of these plasmids containing the IncFII (n = 21) and/or the IncFIA (n = 9) replicon. pMLST results revealed a high diversity of pMLST profiles, especially among plasmids found in *H*30-ST131 isolates. In the *H*30-ST131 isolate susceptible to all antibiotics tested, an F1:A2:B20 profile was found. In the *H*30-R-ST131 isolates resistant to fluoroquinolones, F1:A2:B20 (n = 4), F2:A1:B- (n = 2), F1:A2:B- (n = 1), F-:A2:B20 (n = 1) and F18:A6:B8(n = 10) pMLST profiles were identified. Besides, in the *H*30-R-ST131 isolates also resistant to carbapenems, one isolate had an F1:A2:B20 profile, and for two isolates, the F allele was not identified. Therefore, these were classified as F-:A2:B20. Among the highly resistant sub-lineage *H*30-Rx-ST131, two pMLST profiles, F2:A-:B- (n = 5) and F31:A4:B1 (n = 4), were found in isolates of virotype A and C, respectively. The two *H*22-ST131 and two *H*41-ST131 isolates had the same pMLST profiles, i.e., F2:A-:B1 and F29:A-:B10, respectively, profiles that were not found in *H*30-ST131 isolates. Overall our results show a link between the resistance phenotype and the plasmid type (Fig. 1 and Table 1). The pMLST profiles found in other ST isolates were different from the ones in ST131 isolates, except for the F29:A:B10 plasmid that was present in two ST131 isolates and three ST69 isolates and F31:A4:B1 present in four ST131 and one ST69 isolates. In addition, other plasmid profiles were found in the ST69 isolates, i.e., pMLST profiles F36:A4:B1 (n = 2), F1:A1:B66 (n = 1) and F2:A-:B10 (n = 1). For one ST69 isolate, no plasmid profile could be identified. Interestingly, the *H*41/ST131 sublineage is less associated with an MDR profile and has the same pMLST profile, F29:A:B10, as identified in three of our ST69 isolates, including the ESBL-producing isolate (605). IncF plasmids were not identified in ST73 isolates. The IncF plasmids in ST648 isolates were classified as F1:A1:B1 and in ST405 isolates as F2:A-:B10, except for one isolate that had an F2:A-:B- plasmid (see supplementary data S.3).

3.3. Identification of plasmid-based resistance gene

We analyzed the presence of plasmid-based resistance genes in our isolates. Our results identified 35 different resistance genes, including $bla_{CTX\text{-}M\text{-}15}$, $bla_{CTX\text{-}M\text{-}8}$, $bla_{CTX\text{-}M\text{-}2}$, $bla_{KPC\text{-}2}$, and *aac(6′)-Ib-cr* genes, which represented 68.6 % of all the resistance genes identified in our isolates. Most resistance genes (64.8 %) were found in ST131, ST405, and ST648 isolates, whereas fewer resistance genes were identified in the ST69 and ST73 isolates (Fig. 2). Among ST131 isolates, a higher number of resistance genes was identified in MDR and ESBL-producing isolates belonging to the A/*H*30 and C/*H*30 sublineages. We also found several resistance genes between IS elements, i.e., within a resistance cassette, in *H*30Rx-ST131, ST405, and ST648 isolates. A resistance cassette highly similar to the one present in the pEC958 plasmid found in the *E. coli* ST131 reference strain EC958 (NCBI: HG941719.1) was identified in different plasmids among genetically closely related A/*H*30Rx ST131 isolates (2102, 2206, 5420 and 6638; Fig. 3A). A similar resistance cassette was also identified among C/*H*30Rx ST131 isolates (5770D, 1710D and 9533D; Fig. 3B). However, two regions were found to be different. The IS26/$bla_{CTX\text{-}M15}$/Tn3 region present in pEC958 was replaced by other genes between IS26 elements in our plasmid. Besides, a region of 14Kb containing IS26 elements and resistance genes present in pEC958 was deleted from the plasmid and had most likely been integrated into the chromosome of our isolate (Fig. 3B and C). Interestingly, two resistance cassettes were identified on the chromosome of our MDR isolates (Figs. 3C and D). In ST648 isolates, we identified a chromosomal resistance cassette within phage-related sequences similar to one found in the FDAARGOS_497 reference strain (NCBI: CP033853.1; Fig. 3D). The other chromosomal resistance cassette was identified in the chromosome of *H*30Rx ST131 isolates (5770D, 1710D, 9533D),

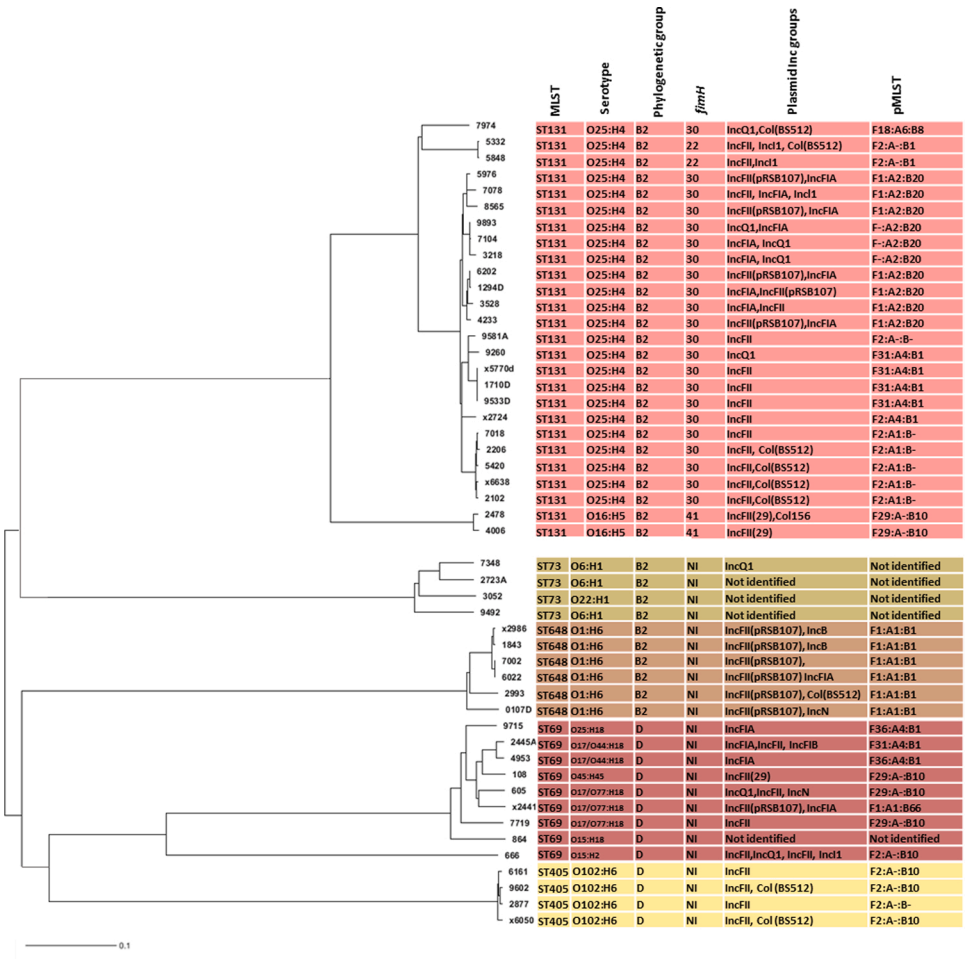


Fig. 1. NJ tree for all the 49 isolates with the distance based on cgMLST. The cgMLST scheme was based on 2764 target genes. The columns indicate the ST type, serotype, phylogenetic groups, the fimH type (only for ST131 isolates), the plasmid replicon type and the pMLST profile. In pink the ST131 isolates, in green the ST73 isolates, in brown the ST648 isolates, in dark pink the ST69 isolates and in yellow the ST405 isolates. Please note that for isolate 7348, the MLST type (ST73) could only be identified using the CGE webtool, not by Seqsphere. NI, not investigated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
Association between the pMLST, virotype and antibiotic resistance profiles among ST131 isolates.

Isolates[1]	MLST	Serotype	Phylogenetic group	*fimH*	Virotype	ESBL	MDR	PMQR		pMLST
								qnrB	*aac(6')-Ib-cr*	
Virotype A, ESBL, resistance to fluoroquinolones										
x6638	ST131	O25:H4	B2	*fimH30*	A	*blaCTX-M-15*	+[2]	–	+	F2:A1:B-
2102	ST131	O25:H4	B2	*fimH30*	A	*blaCTX-M-15*	+	–	+	F2:A1:B-
5420	ST131	O25:H4	B2	*fimH30*	A	*blaCTX-M-15*	+	–	+	F2:A1:B-
2206	ST131	O25:H4	B2	*fimH30*	A	*blaCTX-M-15*	+	–	+	F2:A1:B-
Virotype A, non-ESBL, sensitive to fluoroquinolones										
2478	ST131	O16:H5	B2	*fimH41*	A	–	–	–	–	F29:A-:B10
4006	ST131	O16:H5	B2	*fimH41*	A	–	–	–	–	F29:A-:B10
Virotype C, ESBL, resistance to fluoroquinolones										
1710D	ST131	O25:H4	B2	*fimH30*	C	*blaCTX-M-15*	+	–	+	F2:A4:B1
5770D	ST131	O25:H4	B2	*fimH30*	C	*blaCTX-M-15*	+	–	+	F2:A4:B1
9533D	ST131	O25:H4	B2	*fimH30*	C	*blaCTX-M-15*	+	–	+	F2:A4:B1
Virotype C, carbapenemase-producer, sensitive to fluoroquinolones										
7104	ST131	O25:H4	B2	*fimH30*	C2	*blaKPC-2*	+	–	–	F-:A2:B20
3218	ST131	O25:H4	B2	*fimH30*	C2	*blaKPC-2*	+	–	–	F-:A2:B20
9893	ST131	O25:H4	B2	*fimH30*	C3	*blaKPC-2*	+	–	–	F-:A2:B20
Virotype C, MDR, sensitive to fluoroquinolones										
8565	ST131	O25:H4	B2	*fimH30*	C3	–	+	–	–	F1:A2:B20
5976	ST131	O25:H4	B2	*fimH30*	C3	–	+	–	–	F1:A2:B20

[1] Isolates for which no association was found are not listed in this table.
[2] +, indicates the presence of the gene or characteristic; -, indicates the absence of the gene or characteristic; isolates are grouped according to their virotype, presence of the extended-spectrum beta-lactamase (ESBL) gene, and their multidrug-resistant (MDR), plasmid-mediated quinolone resistance (PMQR), and plasmid multi-locus sequence type (pMLST) profiles.

which was also closely related to a cassette from reference strain FDAARGOS_497 (Fig. 3C). The resistance cassette found in C/*H*30Rx ST131 isolates (5770D, 1710D and 9533D) was also identified in *H*30-R ST131 isolate (6202) (Fig. 3E), and in ST405 (6050) and ST69 (605) isolates. Thus, this resistance cassette was present in several isolates but on different plasmids. More precisely, in the A/*H*30Rx-ST131 isolates

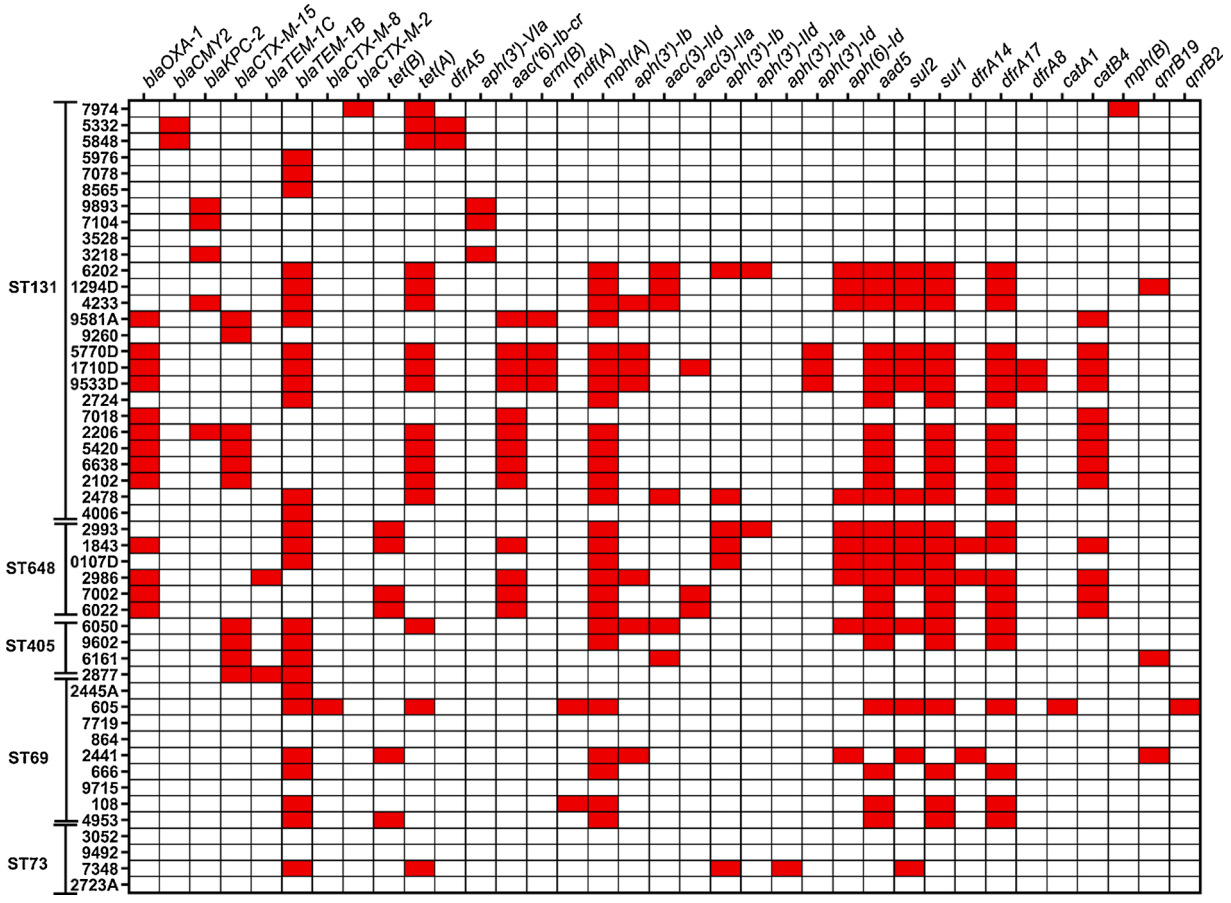


Fig. 2. Distribution of acquired resistance genes identified in the plasmid-derived regions. Red hits indicate the presence of a resistance gene with at least 90 % identity and 60 % coverage. The samples are organized by ST type. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

(5420, 2102 and 6638) it was identified in a plasmid similar to pEC958, while in the C/*H*30Rx-ST131 isolates (5770D, 9531D, 1710D) it was present in a plasmid more similar to pecAZ146 (NCBI:CP018990) (Fig. 4). Finally, we found a statistically significant association between the presence of IncF plasmids and an MDR profile ($p = 0.0142$) and an ESBL profile ($p = 0.0107$).

3.4. Identification of mobile genetic elements present in the ST131 reference strain EC958

The ST131 reference strain EC958 (NCBI: HG941719.1) was used to identify MGEs such as genomic islands (GI-*pheV*, GI-*selC,* GI-*leuX,* GI-*thrW),* phages (1–7), high pathogenicity islands (HPIs), and other insertion elements (ROD, Flag-2, type IV secretion, O-antigen, rat-like) associated with virulence, in our *E. coli* isolates. The results are shown in Fig. 5. The genomic island GI-*thrW* present in reference strain EC958 was highly conserved among ST131 isolates but was not present in any of the non-ST131 isolates. GI-*pheV* was just partially present in almost all ST131 and ST69 isolates, and completely absent in all other lineages. GI-*selC* was absent in almost all ST131 and non-ST131 isolates, except in strain 7018 (ST131), whereas the GI-*leuX* was completely present only in one *H*30 ST131 isolate (9533D) and partially present in the majority of the other ST131 isolates and among a few non-ST131 isolates (ST648, ST405, and ST69). Interestingly, an HPI (PAI IV_{536}) was present in all ST131 isolates. This HPI was also present in all ST73, ST405, and ST648 isolates, and the majority of ST69 (108, 605, 666, 864, 7719, and 9715) isolates. From the seven prophages identified in the EC958 strain, prophages 4 and 6 were partially present in the majority of ST131 isolates, while prophage 5 was completely present in almost all ST131 isolates, except for the *H*41-ST131 isolates. Prophage 1 was completely present in 38.46 % (10 isolates) of the ST131 and partially present in another 42.30 % (11 isolates) ST131 isolates. Prophage 2 was completely present in 88.46 % of our ST131 isolates. Prophage 3 was partially found in twelve (46.15 %) and completely present in thirteen (50 %) ST131 isolates. Finally, prophage 7 was present in seventeen isolates (65.38 %) and partially present in four (15.38 %) of the ST131 isolates.

In addition, ROD insertion elements present in EC958 were also identified among our ST131 isolates. ROD1 was present in all isolates, ROD2 was present or partially present in *H*30-ST131 and *H*22-ST131 isolates, and ROD3 was completely present in *H*30-ST131 and partially present in the other ST131 isolates. ROD insertion elements were not identified among non-ST131 isolates.

The O-antigen loci were partially identified in 38.46 % (10 isolates) ST131 isolates and partially identified in 34.61 % (9 isolates), while the *ratA*-like toxin was found in 96.15 % ST131 isolates. Flag-2 lateral flagellar locus was also conserved among ST131 being completely absent only in *H*41-ST131 isolates. Interestingly, Flag-2 was also present in five ST69 isolates (108, 2445A, 4953, 7719, and 9715). The type IV secretion locus was identified in four ST131 isolates (2206, x2724, 7018, and 8565) and just partially identified in the other ST131 isolates, not present in non-ST131 isolates. The capsular locus was present in *H*41-ST131 isolates and only partially present in most *H*30-ST131 and *H*22-ST131 isolates. Among non-ST131, the capsular locus was partially identified among ST45, ST73, ST69, ST93, ST405, ST453, and ST648 isolates (See Fig. 5; see supplementary S.4).

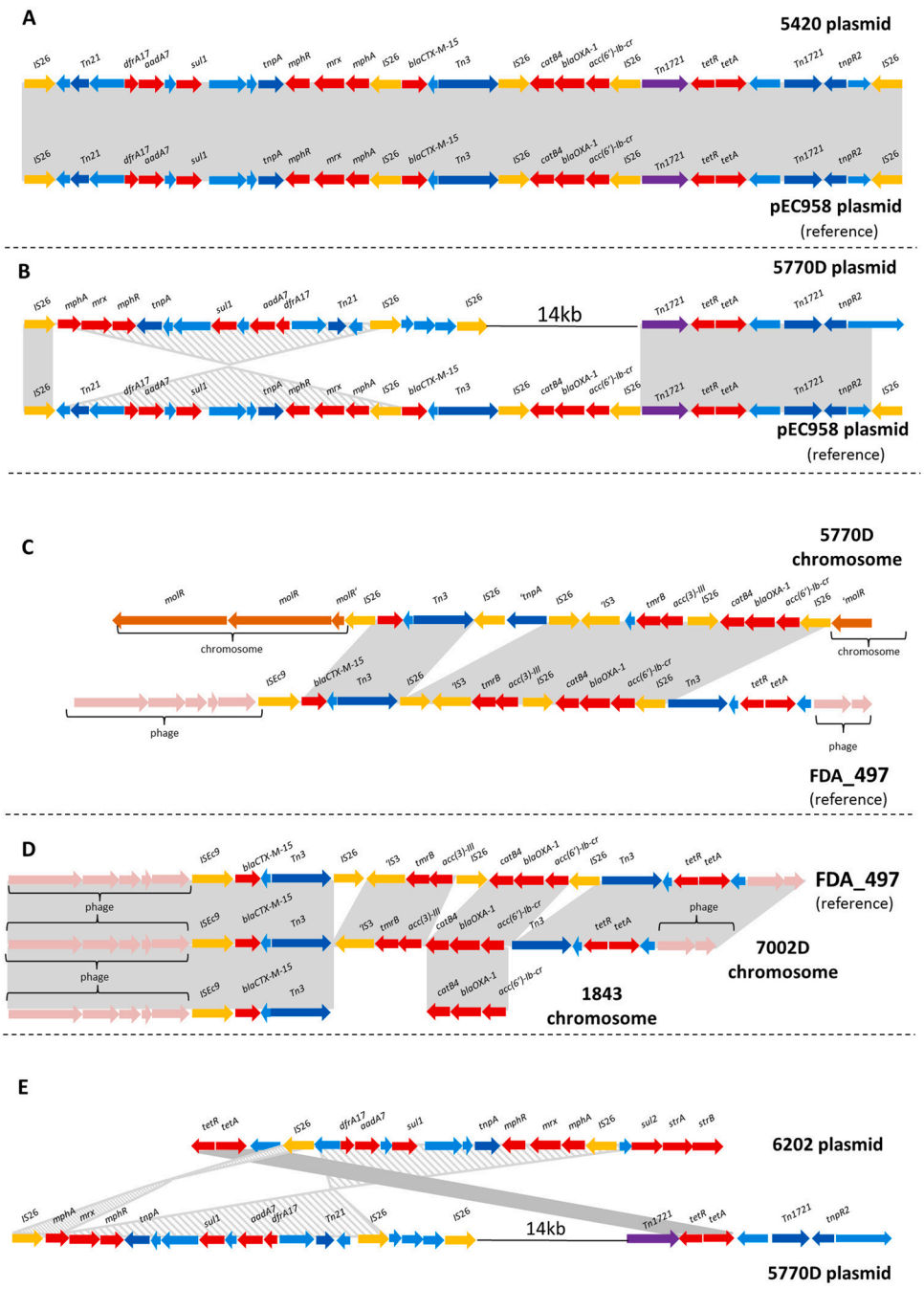


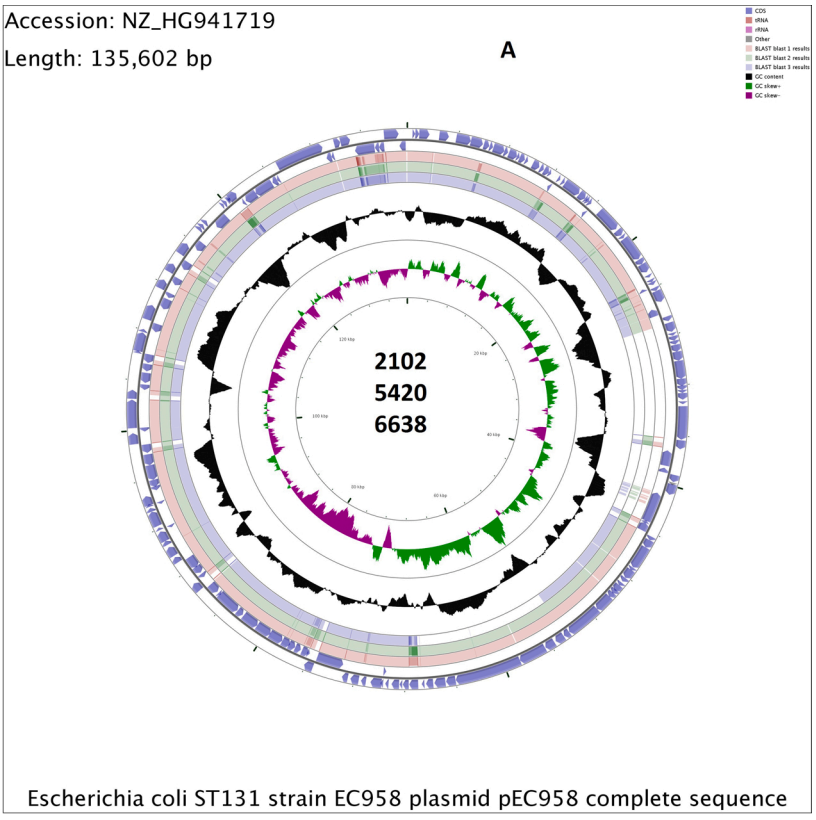
Fig. 3. Alignment of resistance cassettes present in plasmid and chromosome regions using hybrid assemblies of short and long-read sequences. Alignments of **A.** the resistance cassette present in the plasmid of *H*30Rx-ST131, virotype A isolate 5420 and the one found in plasmid pEC958 of the *Escherichia coli* ST131 EC958 reference strain; **B.** the resistance cassette present on the plasmid of 5770D isolate and the one found on plasmid pEC958 of the *Escherichia coli* ST131 EC958 reference; **C.** the resistance cassette found in the chromosome of the 5770D isolate and the one on the chromosome of the FDAARGOS_497 reference strain; **D.** the resistance cassette present in representative isolates of ST648, i.e., 7002D and 1843, and the resistance cassette present on the chromosome of the FDAARGOS_497 reference strain; **E.** the resistance cassette present in plasmids present in *H*30Rx-ST131 isolates 6202 and 5770D. The resistance genes are indicated in red, the IS elements in yellow, Tn elements in dark blue, the phage protein genes in light pink, and other genes present in the chromosome in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

3.5. Characterization of a new genomic island

We also identified a new genomic island (GI) GI-II similar to a GI present in an ST131 isolate from Sweden (reference number: CP023826) in our isolates. It appeared to be of the tRNAPhe-type, was 95900bp in size and contained P fimbriae genes associated with pyelonephritogenic *E. coli* (*papA, papB, papI, papX*), the antigen 43 (*agn43*) associated with biofilm-forming ability, toxins and antitoxin genes *yeeV* and *yeeU*, and genes associated with iron uptake (*iucA*, *iucC, iucD*, *iutA*, *tonB*). This GI was fully present in only two ST131 (7018 and 3218) and partially present in nineteen ST131 isolates, and in five ST69 isolates (108, 605, 2441, 2445A and 4953) (see supplementary data S.4 and Fig. 5 and 6).

3.6. Identification of PAIs

We also investigated the presence of other pathogenicity islands (PAIs) in our isolates (Fig. 7). In each isolate at least one PAI was identified, PAI I_{CFT073} being the most frequently identified one and completely present in 73.13 % of isolates tested and partially in 26.86 %. PAI II_{536} was just partially present in six isolates (5.9 %), and PAI-I_{536} present in only one isolate, so they were the least frequently identified ones. Although a high diversity in PAI profiles was observed, genetically closely related isolates had a more similar profile. Among ST131 isolates, the *H*30-Rx clade C (1710D, 9260, 5770D, 9581A, and 9533D) ones were the only isolates in which PAI II_{536} was at least partially present. The PAI-IV_{536} was identified in almost all ST131 isolates. The PAI-I_{4787} was identified in twenty (76.92 %) ST131 isolates. PAI-AGI3 was only partially present in four ST131 isolates, the *H*22-ST131



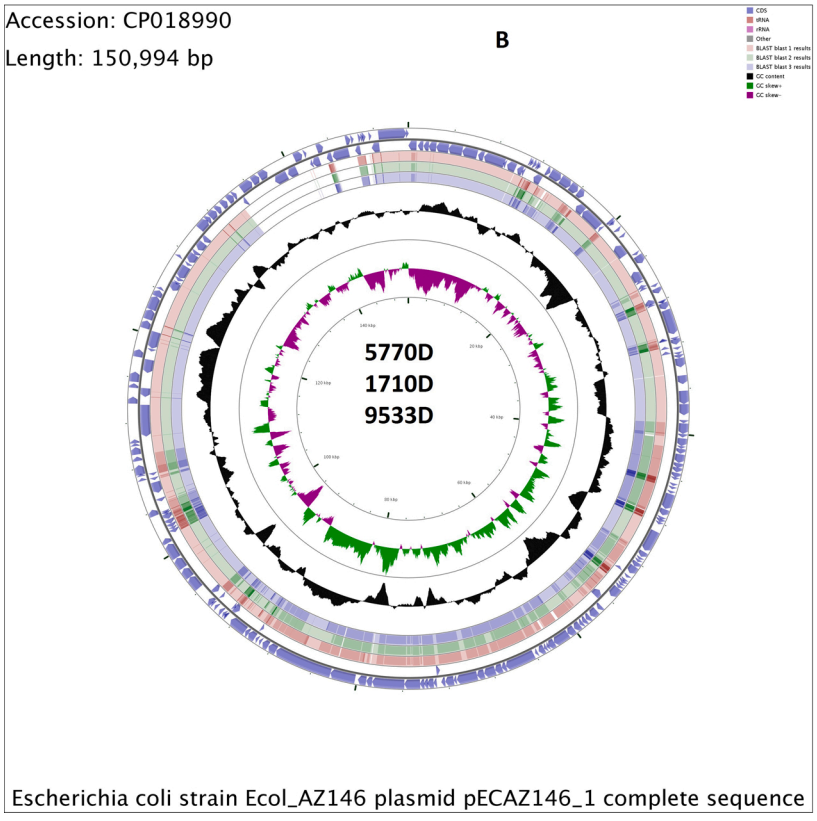


Fig. 4. Alignments of plasmids carrying the resistance cassettes identified in *H*30Rx-ST131 isolates. (A) Alignment between the pEC958 plasmid present in the ST131 *Escherichia coli* EC958 reference strain and the plasmids present in *H*30-Rx virotype A ST131 isolates 2102 (pink), 5420 (green) and 6638 (purple). (B) Alignment between the pECAZ146_1 plasmid present in theST131 *E. coli* Ecol_AZ146 and the plasmids present in *H*30-Rx virotype C ST131 isolates 5770D (pink), 1710D (green) and 9533D (purple). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

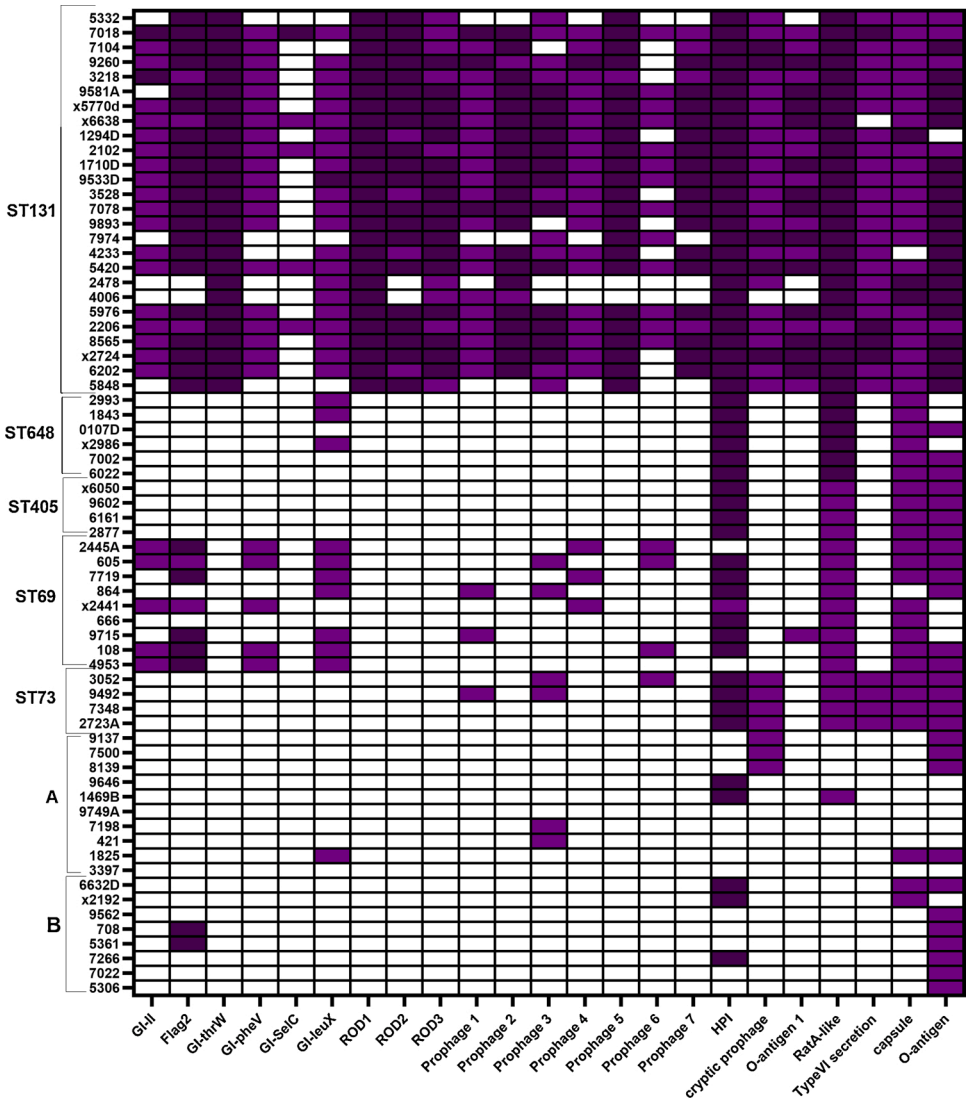


Fig. 5. MGEs identified in *E. coli* isolates.. Identifying the genomic islands GI-*thrW*, GI-*selC*, GI-*pheV*, GI-*leuX*, prophages and other MGEs present in the reference strain EC958.. The dark purple hits indicate the complete sequence of MGEs with > 90 % coverage and > 90 % identity, while the light purple hits indicate partial presence of MGEs with 40 %-89 % coverage and > 90 % identity. The white hits indicate absence of MGEs with < 40 % coverage and/or < 90 % identity. Isolates are grouped either by sequence type (ST) or their phylogenetic group (A or B). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

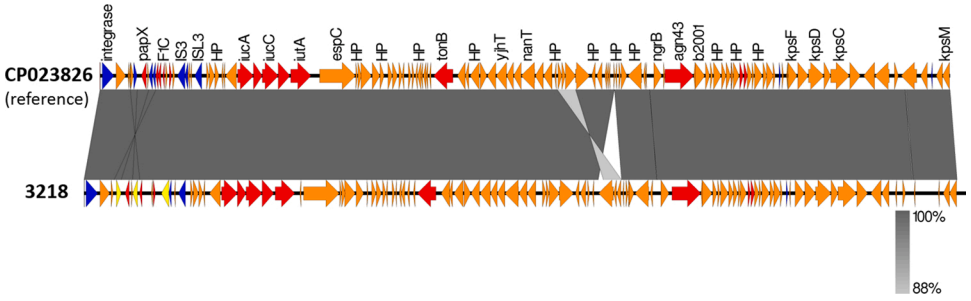


Fig. 6. Alignment of the genomic island (GI) GI-II present in our isolates and the GI found in a Swedish ST131 isolate using hybrid assemblies of short and long-read sequences. On top the Swedish GI (CP023826) and on the bottom, an ST131 representative isolate (3218). In red, the main virulence genes; in blue, the mobile genetic elements (IS, transposase, integrases); in yellow, the IS elements that are slightly different from those found in the Swedish GI; other coding sequences are indicated in orange and include those uncharacterized encoding proteins and genes not directly involved in resistance or virulence profiles. HP, hypothetical proteins. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

isolates (5332 and 5848) and *H*30-ST131 isolates (7018 and 7104). The ETT2 PAI, the espC PAI, and LEE were absent in all ST131 isolates (see supplementary material S.5). On the other hand, ETT2 PAI was present in the majority of ST69 (55.5 %), ST405 (100 %) and ST648 (66.6 %) isolates, while absent in the ST73 ones. The PAI I4787 was only partially present in a few ST648 and ST69 isolates (33.3 %). The ST131 *H*30-Rx clade C presented more similarities in PAIs distribution with ST73 than with other ST types. A different pattern of PAIs was found in the 49 isolates that belong to phylogenetic groups B2 and D, compared to the one found in isolates belonging to phylogenetic groups A and B1. In the latter isolates, the PAI II_{536} and PAI I_{536} were absent, PAI I_{CFT073} was only partially present (Fig. 7 and supplementary material S.5).

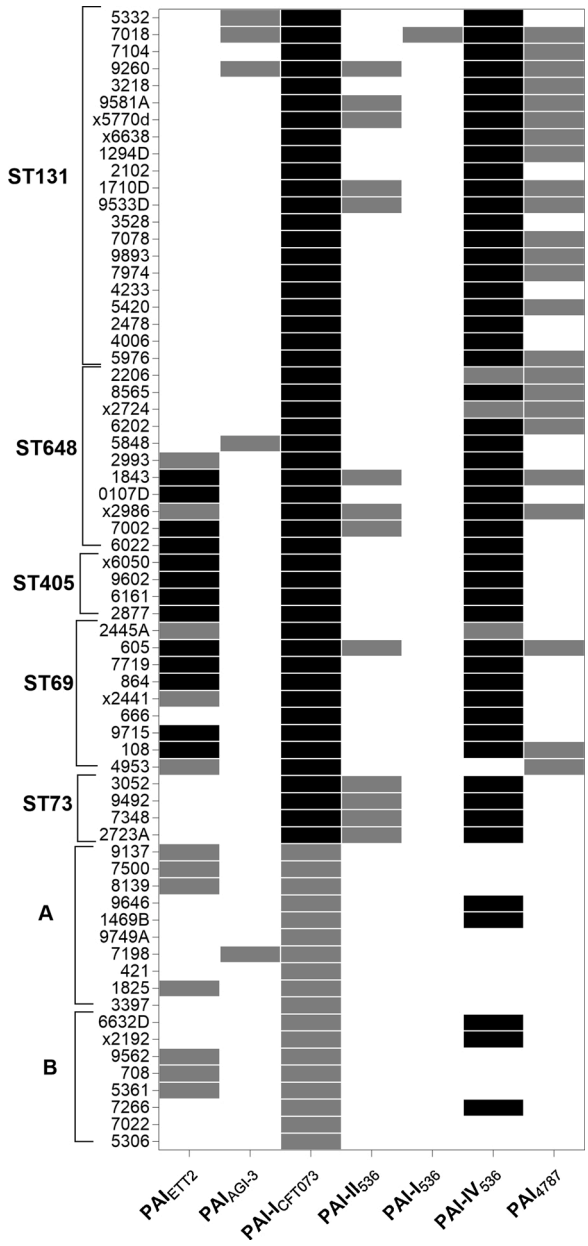


Fig. 7. Pathogenicity islands present in the isolates. The figure shows the most frequently found PAIs identified among different ST types. The black hits indicate the presence of a PAI, while the grey hits indicate that the PAI is partially present, and white hits indicate the absence of the PAI. The isolates belonging to phylogenetic groups B2 and D are indicating with their ST types, whereas isolates belonging to phylogenetic groups A and B1 are indicated by A and B, respectively.

3.7. Virulence genes and mobile genetic elements

We investigated the presence of 77 virulence genes associated with plasmid-based regions or pathogenicity islands present in our 49 isolates. Overall the number of virulence genes identified in plasmid-based regions was lower than the number of genes present in pathogenicity islands. The majority of identified genes on plasmids were associated with iron uptake, and a minor part was associated with toxin production and other virulence determinants. Detected genes include the aerobactin siderophore and the *iroE, iroD, iroB, iroN, iucA, iucB, iucC, iucD, iutA, sitA, sitB, sitC* and *sitD* genes, the gene encoding enterotoxin *senB*, the increased serum survival (*iss*) gene, as well as colicin and microcin associated genes (*mchF* and *cba*) that are involved in the killing of other bacteria. In general, the diversity of MGE-associated virulence genes was high within the ST types. Moreover, the ST648 and ST69 isolates presented a higher virulence score (VS) associated with MGEs, i.e., 44.5 and 36.4, respectively, compared to the isolates of other ST types including ST131 (VS = 24.1), ST73 (26.5) and ST405 (VS = 25.2) isolates. The ST648 and ST69 isolates had the same PAIs (PAI_{ETT2} and PAI I_{CFT073}), which carry genes of the system secretion type 3 (TSS3), pap cluster genes and hemolysin and iron-uptake system genes (Table 2). ST73 and C/*H*30-ST131 isolates (9581A, 9260, 5770D, 1710D and 9533D) had a very similar PAI-associated virulence gene profile. These C/H30-ST131 isolates were associated with an MDR profile and also had higher numbers of virulence genes associated with plasmids and PAIs than other ST131 isolates (Fig. 8). Also, the Fisher's exact test did not identify a significant association between the PAI profile and the presence of IncF plasmids in CTX-M-15-producing isolates.

Table 2
Comparison of the virulence score and presence of PAIs in ST131 and non-ST131 isolates.

	ST131	ST73	ST648	ST69	ST405
VS[a]	24.19	26.5	44.5	36.44	25.25
PAI_{ETT2}	0 (0)[b]	0 (0)	4 (66.6)	5 (55.5)	4 (100)
PAI_{AGI3}	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
PAI I_{CFT073}	26 (100)	4 (100)	6 (100)	9 (100)	4 (100)
PAI II_{536}	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
PAI I_{536}	0 (0)	0 (0)	0 (0)	0(0)	0 (0)
PAI IV_{536}	23 (88.46)	4 (100)	6 (100)	6 (66.6)	4 (100)
PAI_{4787}	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

[a] VS, virulence score, indicated as the mean number of virulence genes in each lineage. This table only lists PAIs of which the complete sequences were identified in the isolates.

[b] The number between brackets indicate the percentage of isolates that have the complete PAI.

4. Discussion

UTIs caused by MDR ExPEC are challenging to treat and manage. In this study, we investigated the presence of MGEs in *E. coli* isolates circulating in Brazil and their role in resistance and virulence. The most prevalent plasmids were of the IncF type and were mainly present in ST131, ST405, and ST648 isolates, known to be often MDR. Our results agree with previous studies in which MDR isolates, including ESBL-CTX-M-15-producing ones, were shown to be associated with IncF plasmids (Dolejska et al., 2012; Partridge et al., 2011). Indeed, a statistically significant association between MDR/ESBL profiles and the presence of IncF plasmids was found in our isolates. Although pMLST profiles were highly diverse in ST131 isolates, F1:A2:B20 and F2:A1:B- were the ones most often found in antibiotic-resistant ST131 isolates. In contrast, F29:A:B10 plasmids were present in two susceptible *H*41/ST131 isolates and three susceptible ST69 isolates. Our results are in agreement with previous studies showing that F29:A-:B10 plasmids were associated with *H*22, *H*41 or *H*30 susceptible (*H*30-S) ST131 isolates, while F1:A2:B20 and F2:A1:B- plasmids were associated with the resistant *H*30-R and *H*30-Rx ST131 isolates (Harris et al., 2017; Johnson et al., 2016; Kanamori et al., 2017). Notably, we found that ST131 isolates having the same virotype and similar antibiotic resistance had the same plasmid profile, indicating an association between the plasmid profile and specific sublineages, which reinforce the hypothesis that specific plasmids are associated with the ongoing evolution of ST131 sublineages.

Additionally, ST405 and ST648 isolates also carried several IncF plasmids, which may be associated, like in ST131, with the successful spread of these clones. The presence of the F1 and F2 alleles is a common feature seen in ST131, ST648, and ST405 lineages associated with an MDR profile. None of the ST73 isolates carried IncF plasmids, which could be the reason why they did not have an MDR phenotype.

Isolates from ST648, although genetically diverse, carried the same

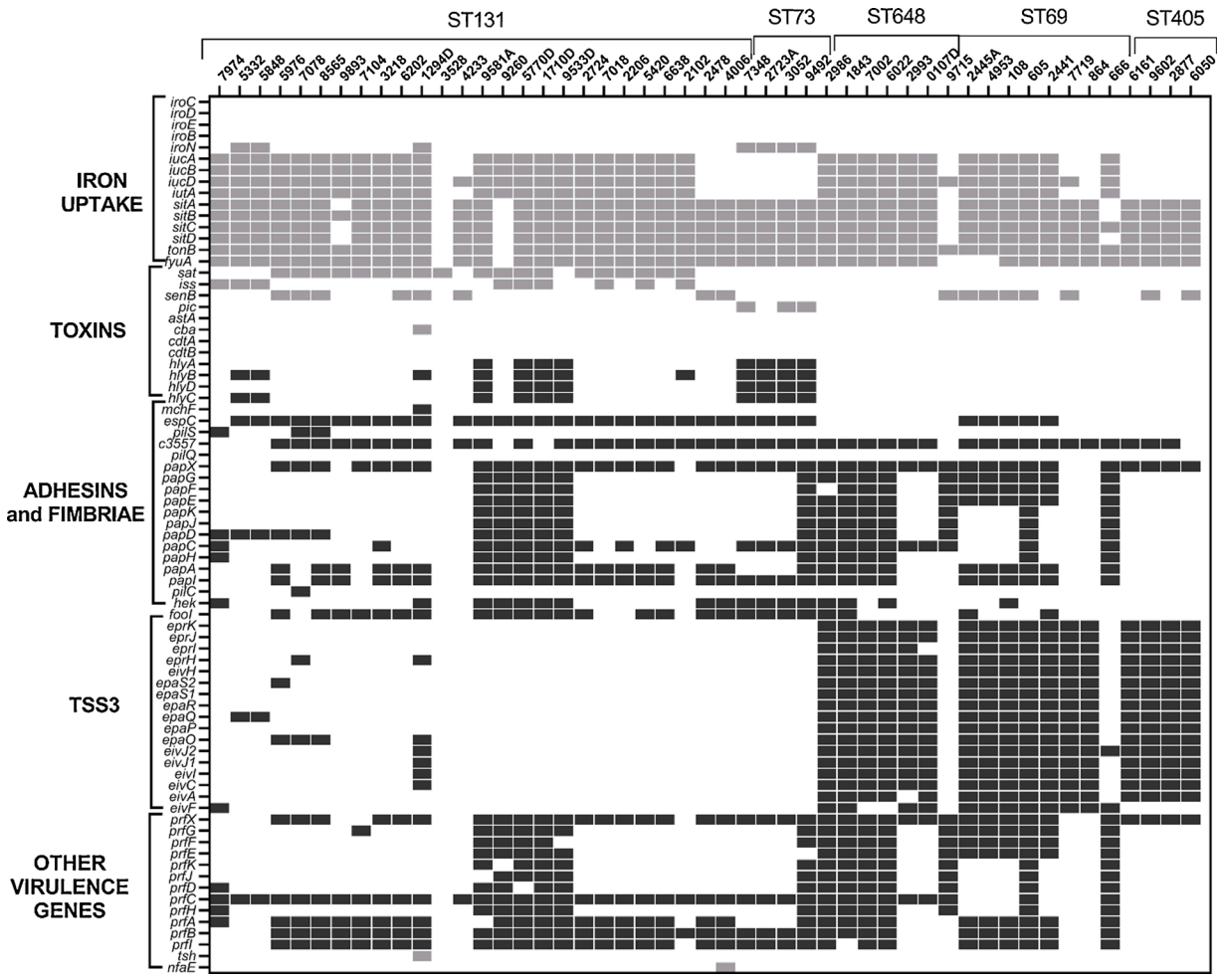


Fig. 8. Heat map of virulence genes present in mobile genetic elements identified in ST131, ST73, ST648, ST69 and ST405 isolates. Black hits indicate the genes present in pathogenicity islands, and the grey hits indicate the presence of virulence genes associated with plasmids. TSS3 indicates the genes associated with the system secretion type 3.

plasmid type, which could indicate that its successful spread is related to the presence of this particular IncF F1:A1:B1 plasmid. This same type of plasmid was identified in ST648 isolates from wastewater in Norway and hospitalized children in the US, and was associated with several resistance genes, including *bla*$_{CTX\text{-}M\text{-}15}$ (Paulshus et al., 2019; Weissman et al., 2013) that was not found in our plasmid. However, the *bla*$_{CTX\text{-}M\text{-}15}$ gene was located on the chromosome of our isolates. The presence of this gene has been associated before with the worldwide spread of ST131 bacteria (Bevan et al., 2017; Calhau et al., 2013) and may also be of benefit for bacteria of belonging to other STs.

We also identified the presence of resistance genes, including *bla*$_{CTX\text{-}M\text{-}15}$, *dfrA17,* and *aac(‘6)-Ib-cr* in the plasmid-derived regions. Most resistance genes were identified in the ST131, ST405, and ST648 lineages. This could be explained by the presence of resistance cassettes, defined as blocks of resistance genes that were identified within plasmids present in these isolates. In *H*30-Rx virotype A ST131 isolates, a cassette similar to the one found on plasmid pEC958 present in the ST131 *E. coli* ST131 reference strain (EC958) was detected. In virotype C ST131 isolates, resistance cassettes were found in plasmids and on the chromosome. In addition, the ST648 isolates presented a phage-related resistance cassette also present on the chromosome. The presence of a resistance cassette on the chromosome indicates a transposition event, which may have occurred to increase the stability of the *bla*$_{CTX\text{-}M\text{-}15}$ gene, and may have been facilitated by the presence of IS26 elements (Decano et al., 2019). This may also be the reason of the presence of phage-related resistance cassette present in the chromosome of our ST648 isolates. The presence of resistance cassettes associated with MGEs was already described in previous studies from different countries that showed the importance of IncF plasmids in resistance and dissemination of *bla*$_{CTX\text{-}M\text{-}15}$ and *aac(6′)Ib-cr* genes among high-risk *E. coli* lineages (Calhau et al., 2013; Dolejska et al., 2012; Matsumura et al., 2013). The observation that the same plasmid containing the resistance cassette was found in each sublineage may explain the association of *bla*$_{CTX\text{-}M\text{-}15}$ with sublineages *H*30-Rx ST131 (virotype A) and *H*30-Rx ST131 (virotype C). Also, in our study, resistance cassettes without the *bla*$_{CTX\text{-}M\text{-}15}$ gene were identified in some ST131 and ST405 isolates but not in ST73 isolates supporting the hypothesis that plasmid-mediated resistance is mainly associated with specific high-risk lineages.

Further, we investigated if the MGEs present in the highly virulent and resistant EC958 strain were also present in our isolates. Our study revealed that GI-*thrW* was the only highly conserved GI present among all our ST131 isolates. The genomic island GI-*leuX* was present in just one isolate and partially present in other ST131 isolates, whereas GI-*pheV* was only partially present among ST131 isolates. This is different from previous studies showing that GI-*leuX* and GI-*pheV* are considered essential for and are identified in nearly all *H*30-ST131 strains (Forde et al., 2014). GI-*leuX* and GI-*thrW* have been identified among other UPEC strains and are known to carry the siderophore receptor locus and Type I restriction/modification system structures also present in other B2 *E. coli* strains. Besides, the presence of complete sequences of these MGEs exclusively in ST131 isolates was already described previously (Forde et al., 2015; Ranjan et al., 2017), and may be explained by the

presence of specific methyltransferases associated with these GIs that can restrict the transfer of these mobile elements to other lineages. We also investigated the presence of prophages. We found that prophage 5 was present in all ST131 isolates. The prophage 7 was the more conserve among ST131 isolates. Prophages 1, 2, and 3 were mostly partially present in the ST131, while prophages 4 and 6 were entirely or partially absent in ST131 isolates. These prophages, particularly prophage 6, are partially present in several *H*30-ST131 reference strains. Also, the HPI was identified in almost all our ST131 isolates, except for two isolates where it was only partially present. In previous studies, the presence of HPI was already associated with phylogenetic group B2 isolates (Calhau et al., 2015; Cyoia et al., 2015). We also identified it in the majority of B2/ST131, B2/ST648, B2/ST73, D/ST405, and D/ST69 isolates (Calhau et al., 2015). Moreover, different from what has been reported previously, we did not find an association between the resistance to antibiotics and the presence of this HPI. Other genomic regions present in EC958 strain known to be variable between *E. coli* as RODs, the *ratA*-like toxin, and Flag-2 lateral flagellar locus were more frequently identified among ST131.

A new GI (named GI-II in this study) was at least partially identified in most ST131 and ST69 isolates and was completely present in two *H*30-ST131 isolates. This GI-II harbored virulence genes, including the P fimbriae, *papAG* operon, the iron uptake system, the *iutA/iucABCD* operon, the *tonB* gene, and the antigen 43, and was similar to a GI found in an isolate from Sweden (reference number: CP023826). Interestingly, one ST69 isolate (2441) that partially present this GI-II also had the same plasmid profile (F29:A-:B10) as *H*41-ST131 isolates. This ST69 isolate is also one of the two MDR ST69 isolates in our study.

Also, the presence of other PAIs was investigated. The complete sequence of PAI_{ETT2} was only identified among ST69, ST405 and ST648 isolates, while PAI I_{CFT073} and PAI IV_{536} were present in almost all isolates, which is consistent with a previous study showing a high prevalence of those latter two PAIs among ExPEC isolates in Brazil (Gazal et al., 2015). Different from previous studies (Calhau et al., 2013), PAI I_{536} and PAI II_{536} were at least partially identified among MDR *H*30-ST131 isolates in our study. Interestingly, these PAIs are more frequently reported to be present in pathogenic than in commensal *E. coli* (Östblom et al., 2011). The distribution of PAI among ST131 and non-ST131 isolates was highly diverse in isolates that belong to B2 and D phylogenetic groups, especially if you also consider isolates in which the PAI was partially present. Nevertheless, we observed that ST73 isolates had a similar profile as *H*30-Rx virotype C ST131 isolates. Most PAIs present in B2 and D isolates were complete or partially absent in most of the A and B1 isolates. Our results were similar to those previously reported (Bozcal et al., 2018; Sarowska et al., 2019), indicating that the PAIs distribution is more associated with the phylogenetic groups than with specific lineages.

We also investigated the presence of virulence genes associated with MGEs. In general, the number of plasmid-based virulence genes was lower than the number of virulence genes present in PAIs in our isolates. The majority of virulence genes identified on plasmid-derived regions were associated with the iron uptake system, adhesions, and toxins. Most of these genes encoded for aerobactin (*iutA/iucABCD*, *sitABCD*) and salmochelin (*iroBCDEN*). These virulence genes have been associated before with virulence in ExPEC (Sarowska et al., 2019) and the production of siderophores was reported previously to be essential for survival in the urinary tract, since the reuptake of iron is part of the host's defense against bacterial infections (Subashchandrabose and Mobley, 2015). Other virulence genes identified in plasmid-derived regions were enterotoxin (*senB)*, increased serum survival (*iss)*, colicin, and microcin (*mchF* and *cba*). The virulence score associated with PAIs and plasmids was lower in ST131 than in most of the non-ST131 isolates. Although, PAI I_{536} and PAI II_{536} were partially present in highly resistant and virulent *H*30-Rx-ST131 isolates there was no clear association between the presence of specific PAIs, the virulence score and antibiotic resistance.

On the other hand, the interplay between resistance and virulence that is observed in ST131 sublineages may be explained not by an increase in the number of virulence genes through acquisition of MGEs, e.g., PAIs, but by the acquisition of specific MGEs carrying virulence and resistance genes without increasing the bacterial fitness cost. The assumption that successful clones are characterized by MGEs that primarily contribute to pathogenicity and resistance but do not otherwise affect their biological fitness is a much-discussed aspect that requires further experimental verification. Nevertheless, it could explain the similarities in PAIs distribution between the *H*30-Rx ST131 and ST73 isolates, shown in our previous study, to have the highest virulence score (Campos et al., 2018). Likewise, GI-II is only present in ST131 and ST69 isolates. Therefore, we hypothesize that the *H*30-Rx ST131 sublineage may acquire MGEs having only genes necessary for virulence, such as the GI-II, that are already present in other lineages and at the same time have other MGEs that are exclusive for ST131 in order to improve the balance between resistance and virulence.

5. Conclusions

In summary, the presence of specific plasmids and resistance cassettes seems to be associated with successful lineages and may be essential for their emergence and evolution towards high-risk clones. Indeed, a clear association between the presence of plasmids with MDR and ESBL profiles was found in high-risk clones. On the other hand, the association between the presence of MGEs and virulence is more complex and involves different MGE types, plasmids, PAIs, and other GIs, all having important virulence factors. In addition, the relation between the acquisition of MGEs and the evolution of the bacteria seems to be lineage-specific, as is supported by the observation that particular MGEs were exclusively identified among *H*30-Rx ST131 isolates. Further studies on MGEs are required to increase our knowledge about how these MGEs affect the virulence and resistance of *E. coli* clones and eventually can lead to the emergence of specific clones. Such knowledge will be useful to identify potentially high virulent lineages in the future early.

Ethical approval and consent to participate

This study was approved by the Pedro Ernesto University Hospital ethical committee according and with Brazilian legislation and received this register number: CAAE number:45780215.8.0000.5259. All the participants signed the consentient term authorizing the use of bacterial isolates in this study.

Author's contributions

AC conceived, designed experiments, analyzed the results and drafted the manuscript. NC analyzed the results and critically revised the manuscript for valuable intellectual content. NA provided the samples, generated some of the data, and drafted some of the work. AR, AF, JR, MF and PD substantially contributed to the interpretation of the data for the work and critically revised the manuscript for relevant intellectual content.

Funding

The author has received financial support from the Abel Tasman Talent Program for biomedical research talent of the University Medical Center Groningen, which aims to support high-quality international research at the University of Groningen/University Medical Center Groningen, the Netherlands. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. This work was partially funded by the 7th JPIAMR call and ZonMw project number 549007002 named SOLIDNESS.

Acknowledgements

The authors would like to thank the Abel Tasman Talent Program and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) for their financial support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijmm.2020.151453.

References

Arredondo-Alonso, S., Rogers, M.R.C., Braat, J.C., Verschuuren, T.D., Top, J., Corander, J., Willems, R.J.L., Schürch, A.C., 2018. Mlplasmids: a user-friendly tool to predict plasmid-and chromosome-derived sequences for single species. Microb. Genom. 4 https://doi.org/10.1099/mgen.0.000224.

Aziz, R.K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9, 1–15. https://doi.org/10.1186/1471-2164-9-75.

Bevan, E.R., Jones, A.M., Hawkey, P.M., 2017. Global epidemiology of CTX-M b -lactamases : temporal and geographical shifts in genotype. J Antimicrob Chemother 1 72 (8), 2145–2155. https://doi.org/10.1093/jac/dkx146.

Bozcal, E., Eldem, V., Aydemir, S., Skurnik, M., 2018. The relationship between phylogenetic classification, virulence and antibiotic resistance of extraintestinal pathogenic *Escherichia coli* in İzmir province,Turkey. Peer J 24 (6). https://doi.org/10.7717/peerj.5470.

Calhau, V., Ribeiro, G., Mendonça, N., Jorge, G., Silva, D., 2013. Prevalent combination of virulence and plasmidic-encoded resistance in ST 131 *Escherichia coli* strains. Virulence 4, 726–729. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3925704/.

Calhau, V., Domingues, S., Ribeiro, G., Mendonça, N., Da Silva, G.J., 2015. Interplay between pathogenicity island carriage, resistance profile and plasmid acquisition in uropathogenic *Escherichia coli*. J. Med. Microbiol. 64, 828–835. https://doi.org/10.1099/jmm.0.000104.

Campos, A.C.C., Andrade, N.L., Ferdous, M., Chlebowicz, M.A., Santos, C.C., Correal, J.C. D., Lo Ten Foe, J.R., Rosa, A.C.P., Damasco, P.V., Friedrich, A.W., Rossen, J.W.A., 2018. Comprehensive Molecular Characterization of *Escherichia coli*Isolates from Urine Samples of Hospitalized Patients in Rio de Janeiro, Brazil. Front. Microbiol. https://doi.org/10.3389/fmicb.2018.00243.

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M.V., Lund, O., Villa, L., Aarestrup, F.M., Hasman, H., 2014. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. https://doi.org/10.1128/AAC.02412-14.

Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the artemis comparison tool. Bioinformatics 21, 3422–3423. https://doi.org/10.1093/bioinformatics/bti553.

Clermont, O., Christenson, J.K., Denamur, E., Gordon, D.M., 2013. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Enviromental Microbiology Reports 5, 58–65. https://doi.org/10.1111/1758-2229.12019.

Cyoia, P.S., Rodrigues, G.R., Nishio, E.K., Medeiros, L.P., Koga, V.L., Pereira, A.P.D., Vespero, E.C., Houle, S., Dozois, C.M., Nakazato, G., Kobayashi, R.K.T., 2015. Presence of virulence genes and pathogenicity islands in extraintestinal pathogenic *Escherichia coli* isolates from Brazil. J. Infect. Dev. 9, 1068–1075. https://doi.org/10.3855/jidc.6683.

Dahbi, G., Mora, A., López, C., Alonso, M.P., Mamani, R., Marzoa, J., Coira, A., García-Garrote, F., Pita, J.M., Velasco, D., Herrera, A., Viso, S., Blanco, J.E., Blanco, M., Blanco, J., 2013. Emergence of new variants of ST131 clonal group among extraintestinal pathogenic *Escherichia coli* producing extended-spectrum β-lactamases. Int. J. Antimicrob. Agents 42, 347–351. https://doi.org/10.1016/j.ijantimicag.2013.06.017.

Dahbi, G., Mora, A., Mamani, R., López, C., Alonso, M.P., Marzoa, J., Blanco, M., Herrera, A., Viso, S., García-Garrote, F., Tchesnokova, V., Billig, M., de la Cruz, F., de Toro, M., González-López, J.J., Prats, G., Chaves, F., Martínez-Martínez, L., López-Cerezo, L., Denamur, E., Blanco, J., 2014. Molecular epidemiology and virulence of *Escherichia coli* O16:H5-ST131: comparison with *H*30 and *H*30-Rx subclones of O25b:H4-ST131. Int. J. Med. Microbiol. 304, 1247–1257. https://doi.org/10.1016/j.ijmm.2014.10.002.

Decano, A.G., Ludden, C., Feltwell, T., Judge, K., Parkhill, J., Dowing, T., 2019. Complete assembly of Escherichia coli sequence type 131 genomes using long reads demonstrates antibiotic resistance gene variation within diverse plasmid and chromosomal contexts. Abstr. Gen. Meet. Am. Soc. Microbiol. 1–12.

Dolejska, M., Brhelova, E., Dobiasova, H., Krivdova, J., Jurankova, J., Sevcikova, A., Dubska, L., Literak, I., Cizek, A., Vavrina, M., Kutnikova, L., Sterba, J., 2012. Dissemination of IncFIIK-type plasmids in multiresistant CTX-M-15-producing Enterobacteriaceae isolates from children in hospital paediatric oncology wards. Int. J. Antimicrob. Agents 40, 510–515. https://doi.org/10.1016/j.ijantimicag.2012.07.016.

Downing, T., 2015. Tackling drug resistant infection outbreaks of global pandemic *Escherichia coli* ST131 using evolutionary and epidemiological genomics. Microorganisms 3, 236–267. https://doi.org/10.3390/microorganisms3020236.

Forde, B.M., Ben Zakour, N.L., Stanton-Cook, M., Phan, M.D., Totsika, M., Peters, K.M., Chan, K.G., Schembri, M.A., Upton, M., Beatson, S.A., 2014. 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 9. https://doi.org/10.1371/journal.pone.0104400.

Forde, B.M., Phan, M., Gawthorne, J.A., Ashcroft, M.M., Stanton-cook, M., Sarkar, S., 2015. Lineage-specific methyltransferases define the methylome of the globally disseminated *Escherichia coli* ST131 clone. mBio 6, 1–12. https://doi.org/10.1128/mBio.01602-15.

Gazal, L.E.S., Puno-Sarmiento, J.J., Medeiros, L.P., Cyoia, P.S., Da Silveira, W.D., Kobayashi, R.K.T., Nakazato, G., 2015. Presence of pathogenicity islands and virulence genes of extraintestinal pathogenic *Escherichia coli* (ExPEC) in isolates from avian organic fertilizer. Poult. Sci. 94, 3025–3033. https://doi.org/10.3382/ps/pev278.

Guidoni, E.B.M., Dalpra, V.A., Figueiredo, P.M., da Silva Leite, D., Mímica, L.M.J., Yano, T., Blanco, J.E., Toporovski, J., 2006. *E. coli* virulence factors in children with neurogenic bladder associated with bacteriuria. Pediatr. Nephrol. 21, 376–381. https://doi.org/10.1007/s00467-005-2133-0.

Hall, J.P.J., Brockhurst, M.A., Harrison, E., 2017. Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria. Philosophical Transactions of Royal Society A. B 372, 20160424. https://doi.org/10.1098/rstb.2016.0424.Harris.

Harris, P.N.A., Zakour, N.L.Ben, Roberts, L.W., Alexander, M., Zowawi, H.M., Tambyah, P.A., Lye, D.C., Lee, T.H., Yin, M., Izharuddin, E., Looke, D., Rogers, B., Bhally, H., Crowe, A., Mark, A., Beatson, S., Paterson, D.L., 2017. Dominance of ST131 *Escherichia coli* carrying blaCTX-M in patients with bloodstream infections caused by cephalosporin-resistant strains in Australia, New Zealand and Singapore : whole genome analysis of isolates from a randomised trial. bioRixiv.

Joensen, K.G., Tetzschner, A.M.M., Iguchi, A., Aarestrup, F.M., Scheutz, F., 2015. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J. Clin. Microbiol. 53, 2410–2426. https://doi.org/10.1128/JCM.00008-15.

Johnson, J.R., Stell, A.L., 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. J. Infect. Dis. 181 (1), 261–272. https://doi.org/10.1086/315217.

Johnson, R.J., Urban, C., Weissman, J.S., Jorgensen, H.J., Lewis, S.J., Hansen, G., Edelstein, H.P., Robicsek, A., Cleary, T., Adachi, J., Paterson, D., Quinn, J., Hanson, D.N., Johnston, D.B., Clabots, C., Kuskowski, A.M., AMERECUS investigators, 2012. Molecular Epidemiological Analysis of *Escherichia coli* sequence type ST131 (O25:H4) and blaCTX-M-15 among Extended-Spectrum-β-Lactamase-Producing *E. coli* from the Unidade States, 2000 to 2009. Antmicrobial Agents and Chemotherapy 56. https://doi.org/10.1128/AAC.05824-11.

Johnson, T.J., Danzeisen, J.L., Youmans, B., Case, K., Llop, K., Munoz-Aguayo, J., Flores-Figueroa, C., Aziz, M., Stoesser, N., Sokurenko, E., Price, L.B., Johnson, J.R., 2016. Separate F-type plasmids have shaped the evolution of the H30 subclone of *Escherichia coli* sequence type 131. mSphere 1, e00121–16. https://doi.org/10.1128/mSphere.00121-16.

Kanamori, H., Parobek, C.M., Juliano, J.J., Johnson, J.R., Johnston, B.D., Johnson, T.J., Weber, D.J., Rutala, W.A., Anderson, D.J., 2017. Genomic analysis of multidrug-resistant *Escherichia coli* from North Carolina community hospitals: ongoing circulation of CTX-M-Producing ST131- *H*30Rx and ST131- *H*30R1 strains. Antimicrob. Agents Chemother. 61 https://doi.org/10.1128/aac.00912-17.

Koga, V.L., Tomazetto, G., Cyoia, P.S., Neves, M.S., Vidotto, M.C., Nakazato, G., Kobayashi, R.K.T., 2014. Molecular screening of virulence genes in extraintestinal pathogenic *Escherichia coli* isolated from human blood culture in Brazil. Biomed Res. Int. 2014 https://doi.org/10.1155/2014/465054.

Lyimo, B., Buza, J., Subbiah, M., Temba, S., Kipasika, H., Smith, W., Call, D.R., 2016. IncF plasmids are commonly carried by antibiotic resistant *Escherichia coli* isolated from drinking water sources in northern Tanzania. Int. J. Microbiol. 2016 https://doi.org/10.1155/2016/3103672, 3103672.

Mathers, A.J., Peirano, G., Pitout, J.D.D., 2015. *Escherichia Coli* ST131: the Quintessential Example of anInternational Multiresistant High-risk Clone, Advances in Applied Microbiology. Elsevier Ltd. https://doi.org/10.1016/bs.aambs.2014.09.002

Matsumura, Y., Yamamoto, M., Nagao, M., Ito, Y., Takakura, S., Ichiyama, S., 2013. Association of fluoroquinolone resistance, virulence genes, and IncF plasmids with extended-spectrum-β-lactamase-producing *Escherichia co*li sequence type 131 (ST131) and ST405 clonal groups. Antimicrob. Agents Chemother. 57, 4736–4742. https://doi.org/10.1128/AAC.00641-13.

Miyoshi-akiyama, T., Sherchan, J.B., Doi, Y., Nagamatsu, M., Sherchand, J.B., Tandukar, S., Ohmagari, N., Kirikae, T., Ohara, H., Hayakawa, K., 2016. Comparative genome analysis of Extended-Spectrum-β -Lactamase-Producing *Escherichia coli* sequence type 131 strains from Nepal and Japan. mSphere 1, 1–14. https://doi.org/10.1128/mSphere.00289-16.Editor.

Östblom, A., Adlerberth, I., Wold, A.E., Nowrouzian, F.L., 2011. Pathogenicity island markers, virulence determinants malX and usp, and the capacity of *Escherichia coli* to persist in infants' commensal microbiotas. Appl. Environ. Microbiol. 77, 2303–2308. https://doi.org/10.1128/AEM.02405-10.

Partridge, S.R., Zong, Z., Iredell, J.R., 2011. Recombination in IS26 and Tn2 in the evolution of multiresistance regions carrying bla CTX-M-15 on conjugative IncF plasmids from *Escherichia coli*. Antimicrob. Agents Chemother. 55, 4971–4978. https://doi.org/10.1128/aac.00025-11.

Paulshus, E., Thorell, K., Guzman-Otazo, J., Joffre, E., Colque, P., Kühn, I., Möllby, R., Sørum, H., Sjöling, Å., 2019. Repeated isolation of Extended-Spectrum-β-Lactamase-Positive *Escherichia coli* sequence types 648 and 131 from community wastewater

indicates that sewage systems are important sources of emerging clones of antibiotic-resistant Bacteria. Antimicrob. Agents Chemother. 63, 1–12. https://doi.org/10.1128/aac.00823-19.

Peirano, G., Pitout, J.D.D., 2014. Fluoroquinolone-resistant *Escherichia coli* sequence type 131 isolates causing bloodstream infections in a canadian region with a centralized laboratory system: rapid emergence of the *H*30-RX sublineage. Antimicrob. Agents Chemother. 58, 2699–2703. https://doi.org/10.1128/AAC.00119-14.

Price, L.B., Johnson, J.R., Aziz, M., 2013. The epidemic of extended-spectrum- β -Lactamase-Producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, *H*30-Rx. mBio 4, e00377–13. https://doi.org/10.1128/mBio.00377-13.

Ranjan, A., Shaik, S., Nandanwar, N., Hussain, A., Tiwari, S.K., Semmler, T., Jadhav, S., Wieler, L.H., Alam, M., Colwell, R.R., Ahmed, N., 2017. Comparative genomics of *Escherichia coli* isolated from skin and Soft tissue and other extraintestinal infections. mBio 8. https://doi.org/10.1128/mBio.01070-17 e01070-17.

Robertson, J., Nash, J.H.E., 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb. Genom. 1–7. https://doi.org/10.1099/mgen.0.000206.

Sarowska, J., Koloch, B.F., Kmiecik, A.J., Madrzak, M.F., Ksiazczyk, M., Ploskonska, G.B., Krol, I.C., 2019. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources : recent reports. Gut Pathog. 1–16. https://doi.org/10.1186/s13099-019-0290-0.

Song, L., Wang, X., Zhang, W., Ye, L., Feng, X., 2018. Low-intensity ultrasound promotes the horizontal transfer of resistance genes mediated by plasmids in *E. coli*. 3 Biotech 8 (5), 224. https://doi.org/10.1007/s13205-018-1247-6.

Subashchandrabose, S., Mobley, H.L.T., 2015. Virulence and fitness determinant of uropethogenic *Escherichia coli*. Microbiol. Spectr. 3, 1–32. https://doi.org/10.1128/microbiolspec.UTI-0015-2012.Virulence.

Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27, 1009–1010. https://doi.org/10.1093/bioinformatics/btr039.

Weissman, S.J., Adler, A., Qin, X., Zerr, D.M., 2013. Emergence of extended-spectrum β-lactam resistance among *Escherichia coli* at a US academic children's hospital is clonal at the sequence type level for CTX-M-15, but not for CMY-2. Int. J. Antimicrob. Agents 41, 414–420. https://doi.org/10.1016/j.ijantimicag.2013.01.006.

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. BIOINFORMATICS APPLICATIONS NOTE Bandage : interactive visualisation of de novo genome assemblies. Bioinformatics 1–2.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13, 1–20. https://doi.org/10.1371/journal.pcbi.1005595.

Yoon, S.H., Hur, C.-G., Kang, H.-Y., Kim, Y.H., Oh, T.K., Kim, J.F., 2005. A computational approach for identifying pathogenicity islands in prokaryotic genomes. BMC Bioinformatics 6, 184. https://doi.org/10.1186/1471-2105-6-184.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67, 2640–2644. https://doi.org/10.1093/jac/dks261.

Zaw, M.T., Yamasaki, E., Yamamoto, S., Nair, G.B., Kawamoto, K., Kurazono, H., 2013. Uropathogenic specific protein gene, highly distributed in extraintestinal uropathogenic *Escherichia coli*, encodes a new member of H-N-H nuclease superfamily. Gut Pathog. 5, 1. https://doi.org/10.1186/1757-4749-5-13.

Zowawi, H.M., Harris, P.N.A., Roberts, M.J., Tambyah, P.A., Schembri, M.A., Pezzani, D. M., Williamson, D.A., Paterson, D.L., 2015. The emerging threat of multidrug-resistant Gram-negative bacteria in urology. Nat. Rev. Urol. 12 (10), 570–584. https://doi.org/10.1038/nrurol.2015.199.