



Phylogenomics of Globally Spread Clonal Groups 14 and 15 of *Klebsiella pneumoniae*

 Carla Rodrigues,^{a,*} Val F. Lanza,^{b,c}  Luísa Peixe,^{a,d} Teresa M. Coque,^{c,e}  Ângela Novais,^{a,d}
on behalf of the ESCMID Study Group for Epidemiological Markers (ESGEM)

^aUCIBIO, Applied Molecular Biosciences Unit, Department of Biological Sciences, Laboratory of Microbiology, Faculty of Pharmacy, University of Porto, Porto, Portugal

^bUnidad de Genómica Traslacional Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain

^cCIBER en Enfermedades Infecciosas (CIBERINFEC), Madrid, Spain

^dAssociate Laboratory i4HB—Institute for Health and Bioeconomy, Faculty of Pharmacy, University of Porto, Porto, Portugal

^eServicio de Microbiología, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain

ABSTRACT *Klebsiella pneumoniae* sequence type 14 (ST14) and ST15 caused outbreaks of CTX-M-15 and/or carbapenemase producers worldwide, but their phylogeny and global dynamics remain unclear. We clarified the evolution of *K. pneumoniae* clonal group 14 (CG14) and CG15 by analyzing the capsular locus (KL), resistome, virulome, and plasmidome of public genomes ($n = 481$) and *de novo* sequences ($n = 9$) representing main sublineages circulating in Portugal. CG14 and CG15 evolved independently within 6 main subclades defined according to the KL and the accessory genome. The CG14 ($n = 65$) clade was structured in two large monophyletic subclades, CG14-I (KL2, 86%) and CG14-II (KL16, 14%), whose emergences were dated to 1932 and 1911, respectively. Genes encoding extended-spectrum β -lactamase (ESBL), AmpC, and/or carbapenemases were mostly observed in CG14-I (71% versus 22%). CG15 clade ($n = 170$) was segregated into subclades CG15-IA (KL19/KL106, 9%), CG15-IB (variable KL types, 6%), CG15-IIA (KL24, 43%) and CG15-IIB (KL112, 37%). Most CG15 genomes carried specific GyrA and ParC mutations and emerged from a common ancestor in 1989. CTX-M-15 was especially prevalent in CG15 (68% CG15 versus 38% CG14) and in CG15-IIB (92%). Plasmidome analysis revealed 27 predominant plasmid groups (PG), including particularly pervasive and recombinant F-type ($n = 10$), Col ($n = 10$), and new plasmid types. While *bla*_{CTX-M-15} was acquired multiple times by a high diversity of F-type mosaic plasmids, other antibiotic resistance genes (ARGs) were dispersed by IncL (*bla*_{OXA-48}) or IncC (*bla*_{CMY/TEM-24}) plasmids. We first demonstrate an independent evolutionary trajectory for CG15 and CG14 and how the acquisition of specific KL, quinolone-resistance determining region (QRDR) mutations (CG15), and ARGs in highly recombinant plasmids could have shaped the expansion and diversification of particular subclades (CG14-I and CG15-IIA/IIB).

IMPORTANCE *Klebsiella pneumoniae* represents a major threat in the burden of antibiotic resistance (ABR). Available studies to explain the origin, the diversity, and the evolution of certain ABR *K. pneumoniae* populations have mainly been focused on a few clonal groups (CGs) using phylogenetic analysis of the core genome, the accessory genome being overlooked. Here, we provide unique insights into the phylogenetic evolution of CG14 and CG15, two poorly characterized CGs which have contributed to the global dissemination of genes responsible for resistance to first-line antibiotics such as β -lactams. Our results point out an independent evolution of these two CGs and highlight the existence of different subclades structured by the capsular type and the accessory genome. Moreover, the contribution of a turbulent flux of plasmids (especially multireplicon F type and Col) and adaptive traits (antibiotic resistance and metal tolerance genes) to the pangenome reflect the exposure and adaptation of *K. pneumoniae* under different selective pressures.

Editor Ruth M. Hall, University of Sydney

Copyright © 2023 Rodrigues et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Ângela Novais, aamorim@ff.up.pt, or Teresa M. Coque, mariateresa.coque@salud.madrid.org.

*Present address: Carla Rodrigues, Biodiversity & Epidemiology of Bacterial Pathogens, Institut Pasteur, Université Paris Cité, Paris, France.

The authors declare no conflict of interest.

Received 25 August 2022

Accepted 26 March 2023

Published 26 April 2023

KEYWORDS IncF plasmids, accessory genome, antimicrobial resistance, Col plasmids, genomic epidemiology, high-risk clones, mosaic plasmids, multireplicon plasmids, plasmidome, resistome

Klebsiella pneumoniae subsp. pneumoniae (here referred to as *K. pneumoniae*) is currently considered one of the urgent threats for the emergence and spread of antibiotic resistance (ABR) (1, 2) and one of the top six human pathogens causing infections with high mortality rates (3). The increase of nosocomial infections caused by multidrug-resistant (MDR) *K. pneumoniae* and community-acquired infections caused by hypervirulent strains represents a major public health problem (4, 5). Recent advances in the population structure of *K. pneumoniae* using core-genome multilocus sequence typing (cgMLST) or comparative analysis of whole-genome sequences (WGS) revealed the predominance of a few clonal groups (CGs) associated with ABR, namely, CG15, CG29, CG147, CG101, CG231, CG258, and CG307 (4, 6). However, comprehensive genomic analyses are available for only a few of these predominant CGs (CG258, CG307, CG101, and CG147) (7–10).

Strains belonging to *K. pneumoniae* sequence type 14 (ST14) and ST15 represent 2.1% and 5.2% of the publicly available genomes (accessed from the GenBank assembly repository on July 2020), respectively, and are frequently producers of extended-spectrum β -lactamases (ESBL) or carbapenemases that confer resistance to different antibiotics and are involved in hospital outbreaks worldwide (11, 12). ST15 and ST14 were initially considered to be highly related because all MLST alleles are identical except one point mutation in *infB* (<https://bigsd.b.pasteur.fr/klebsiella/>) (13, 14). Further analyses of disparate genome data sets (8 to 94 genomes; median = 23) led to variable conclusions, most considering CG15 and CG14 to be two different CGs (6, 15–17). Furthermore, the apparent frequent recombination events illustrated by the diversity of capsular types (KL) within ST15 and ST14 lineages circulating in Portugal since at least 2010 (18–21) and available recombination-free maximum likelihood phylogenetic inferences reinforce the need for a comprehensive phylogenetic analysis (15).

Available studies suggest that ST15 and ST14, as well as other MDR *K. pneumoniae* STs, are deep-branching lineages with a low nucleotide divergence (<0.5%) but highly variable accessory genomes (15). Such genomic variation is apparently due to large homologous recombination events involving the capsule locus or via exchange of plasmids, phages, or integrative conjugative elements (ICEs) (4, 15). However, few studies have provided detailed descriptions of the accessory genome, and especially of the plasmidome, in order to explain the adaptation and diversification of these *K. pneumoniae* CGs (16, 22). Furthermore, most studies have addressed plasmid diversity by replicon typing (23), which fails to establish whole plasmid entities. Technical limitations for assembling and sorting short-read sequences and the cost of long-read sequencing have limited the number of high-resolution analyses in large collections of isolates (24–27) until recently (28, 29).

This study analyzed the diversity and evolution of CG14 and CG15 *K. pneumoniae* genomes focusing on both its core and accessory genomes (antimicrobial resistance, capsule polysaccharide locus, and plasmidome) using high-resolution bioinformatic tools (30, 31). A detailed description of the plasmidome is also provided.

RESULTS

Phylogenomic analysis of global *K. pneumoniae* CG15 and CG14. We performed a detailed phylogenomic analysis of 235 nonduplicated CG14 ($n = 65$; 63 ST14 and 2 single-locus variants [SLV]) and CG15 ($n = 170$; 162 ST15 and 8 SLV) isolates. These included seven out of the nine *de novo*-sequenced genomes (5 ST15 and 2 ST14) representing the five major lineages circulating in Portugal since early 2000s (Table 1). They have been previously categorized by diverse typing methods, such as multilocus sequence typing (MLST), pulsed-field gel electrophoresis (PFGE), and Fourier transform infrared (FT-IR) spectroscopy, that together supported lineage definition (18–21). These were substantiated with publicly available genomes ($n = 228$; 63 CG14 and 165 CG15)

TABLE 1 *K. pneumoniae* isolates sequenced in this study and genome sequence information^a

Strain	Isolation yr	ST	KL type	No. of contigs	Sequencing coverage	Total length (bp)	GC (%)	N_{50}	Largest contig (bp)	Accession no.	Reference(s)
H49	2003	14	2	78	104	5,590,357	57.32	350,404	701,584	GCA_013504725.1	19, 21
H1122	2010	14	16	75	105	5,576,295	57.21	399,091	1,013,553	GCA_013504725.1	18, 19
44	2013	15	24	109	106	5,745,985	56.95	210,873	584,875	GCA_013620905.1	19
C1686*	2012	15	24	130	112	6,050,103	56.46	210,874	472,869	GCA_013521185.1	19
C1693	2012	15	24	132	105	5,649,365	57.12	184,651	371,151	GCA_013620925.1	19
H1119	2010	15	39	115	102	5,529,166	57.26	162,787	429,762	GCA_011090325.1	18, 19
C1699	2012	15	112	108	103	5,580,881	57.05	220,754	472,869	GCA_011065465.1	19
C1694*	2012	15	19	90	101	5,453,399	57.17	203,303	584,887	GCA_013620895.1	19
K47	2012	15	110	115	106	5,734,620	56.63	204,524	584,887	GCA_013620875.1	19, 79

^aAll strains were of phylogenetic group Kp1 and type O1. ST, sequence type; KL, capsular locus. Asterisks indicate strains that were closely related to strain 44 (shared <21 SNPs) and were excluded to avoid overrepresentation in the final data set (see Materials and Methods for details).

that met our inclusion criteria, as described in Materials and Methods. They correspond to strains isolated mainly from humans (97%) in Europe (55%), Western and Southeastern Asia (16%), and North America (7%) between 1980 and 2018 (see Fig. S1 in the supplemental material).

Figure 1 shows the maximum-likelihood phylogenetic tree based on the concatenation of 4,420 core genes (representing ~80% of a *K. pneumoniae* genome) which grouped the genomes in two large clades corresponding to the CG, and six main subclades (2 CG14 and 4 CG15) that were in accordance with the capsular type (KL). The CG14 clade is split into CG14-I (KL2, $n = 56$) and CG14-II (KL16, $n = 9$), whereas CG15 is structured into CG15-I and CG15-II, which are further subdivided into two subclades arbitrarily designed "A" and "B," as follows: CG15-IA (mainly KL19, $n = 16$), CG15-IB (variable KL types, $n = 10$), CG15-IIA (mainly KL24, $n = 73$), and CG15-IIB (KL112, $n = 63$). The single nucleotide polymorphism (SNP) median distance between CG14 and CG15 genomes was 2,452 SNPs/megabase (ranging between 7,500 and 12,143 SNPs) but much lower within CG14 (541 to 2,328 SNPs/megabase) or CG15 (252 to 1,084 SNPs/megabase) (Fig. S2).

In contrast with the correlation of particular KL types with the CG14 and CG15 subclades, virulence-factors (VF) appear to be widely distributed throughout the phylogeny (Table S1). Ninety-two percent of the genomes carried the O1 antigen type (92% of these are variant O1v1), while the occurrence of other O-antigen types was sporadic (3 O2 and 1 O4). All CG15 and CG14 genomes carried the *kfu* operon, encoding an iron ABC transporter, and a complete copy or remnants of the *kpi* chaperone-usher pilus system, the latter associated not only with CG15 as previously reported (32) but also with CG14. Eight different lineages of the yersiniabactin siderophore cluster (*ybt*) associated with eight different ICEKp structures and 48 *ybt* locus sequence types (YbSTs) were detected in 51% (120/235) of the genomes distributed throughout the phylogeny (Fig. 1). The highest diversity of *ybt* lineages was identified in subclades CG15-IIB (KL112) and CG14-I (KL2) (6 and 5 sublineages, respectively). The lineage *ybt*10/ICEKp4 was the most frequently detected (57% [68/120]), being overrepresented in the CG14-II (56%, 5/9) and CG15-IIA (51%, 37/73) subclades, respectively. The second most common (17% [20/120]) was the *ybt*9/ICEKp3 lineage, detected more frequently in CG15-IIB (14% [9/63]) and CG14-I (13% [7/56]). Minority lineages were *ybt*14/ICEKp5 and *ybt*13/ICEKp2, the former detected in 20% of the CG14-I genomes (Fig. 1 and Table S1). Only one strain (ST2174-KL2, SLV ST14) contained the putative virulence genes *rmpA2*, *iuc*, and *iro*, which encode a mucoid phenotype regulator and aerobactin and salmochelin siderophores, respectively. Three other ST15 genomes (2 KL112 and 1 KL24) carried the *clb* locus (encoding the genotoxin colibactin), *iuc*, and/or *rmpA2*.

Phylogenetic structure and diversity within *K. pneumoniae* CG14. The 65 CG14 genomes are grouped into two deeply branched subclades of variable size and diversity: CG14-I (86% KL2; 7 to 546 SNPs/megabase) and CG14-II (14% KL16; 29 to 108 SNPs/megabase) (Fig. 1 and Table S2). A time-scaled phylogenetic inference (root-to-tip regression analysis: $R^2 = 0.1474$ [Fig. S3]) suggested a deep-branching structure of

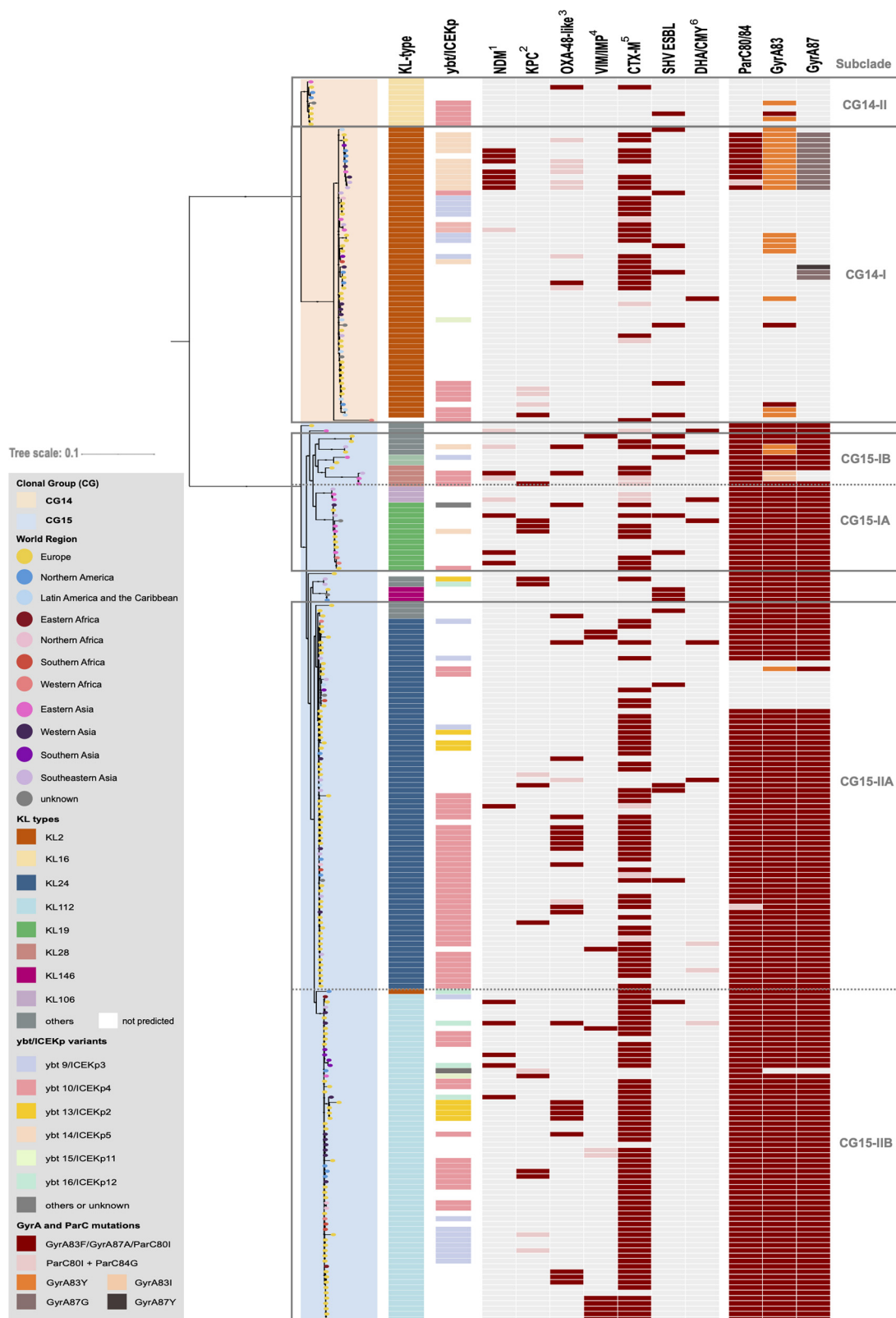


FIG 1 Phylogenetic structure of *K. pneumoniae* CG14 and CG15. Shown is a maximum likelihood tree (model GTR+F+ASC+R4) inferred from 35,783 SNPs extracted from the alignment of 4,420 core genes and rooted using two outgroups (*K. pneumoniae* ST540 Kpn0019 [SRA accession number [SRR2098710](#)] and *K. pneumoniae* ST101 Kp_Goe_33208 [GenBank assembly accession number [GCF_001902435.1](#)]). Branch lengths represent the number of nucleotide substitutions per site (scale, 0.01 substitution per site). The two main clades correspond to CG14 and CG15 and are shaded accordingly (see key), and main subclades thereof are delimited within (Continued on next page)

these two CG14 subclades and an evolutionary rate within CG14 of 4.76×10^{-7} substitutions/site/year (95% highest posterior density [HPD], 3.30×10^{-7} to 6.21×10^{-7}) (Fig. S4). According to this prediction, the emergence of the CG14 ancestor was estimated to be around 1904 (95% HPD, 1862 to 1941), and those of the CG14-I and CG14-II subclades were estimated to be around 1932 (95% HPD, 1904 to 1955) and 1911 (95% HPD, 1865 to 1952), respectively.

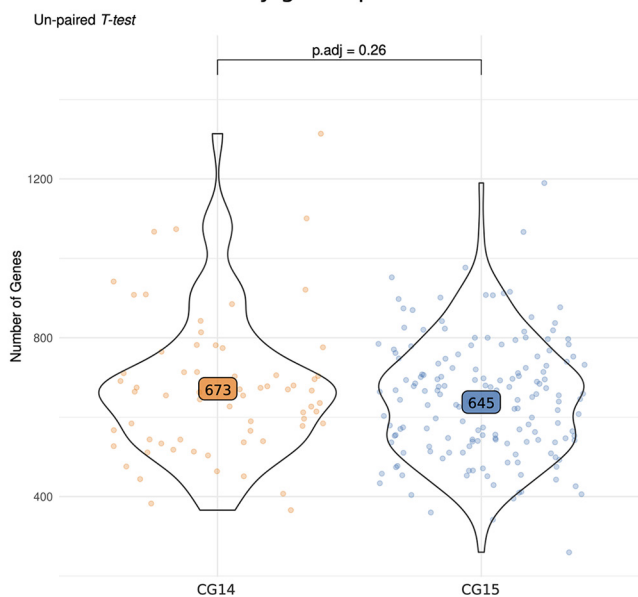
K. pneumoniae CG14 isolates were identified in 26 countries (45% in Europe, 27% in Asia, and 21% in America) between 1980 and 2018 (Fig. S1). One *K. pneumoniae* ST14-KL2 genome from Portugal dates back to 1980 and carries only *strAB* genes, while Portuguese genomes from the early 2000s ($n = 4$; 2003 to 2014) represent isolates causing infections in hospitalized patients from different studies that contributed to the dissemination of genes encoding SHV-106 (ST14-KL16), TEM-24, or KPC-3 (ST14-KL2) (27, 28) (Table 1). Genes encoding ESBL (*bla*_{CTX-M-15/-3/-9/-36}, *bla*_{TEM-24}, or *bla*_{SHV-12/-106}), AmpC (*bla*_{DHA-1}), and/or carbapenemases (*bla*_{OXA-232}, *bla*_{NDM-1}, or *bla*_{KPC-2/-3}) were observed in 65% of the CG14 genomes, and they were more frequently detected among CG14-I than among CG14-II isolates (71% versus 22%; $P = 0.0042$). Of note, 11 highly related CG14-I genomes (39 to 259 SNPs; mean, 124 SNPs) carried multiple antibiotic resistance genes [*bla*_{NDM-1}, *bla*_{CTX-M-15}, *bla*_{OXA-232}, *qnrB1*, *armA*, *aac(6')-Ib-cr*, *aac(3)-IId*, *aph(3')-VI*, *aadA2*, *dfrA1*, *dfrA12*, *catA1*, *catB4*, and/or *cmlA5*], specific mutations in quinolone resistance-determining regions (QRDR), and a specific mutation in *ompK36* (encoding an OmpK36GD variant) (33), showing dissemination of a pan-drug-resistant sublineage in 7 countries in Asia, the United States, and the United Kingdom between 2010 and 2016. Mutations in *GyrA* were uncommon and variable in CG14 genomes (6 to 19% 83Y, 83F, or 87G). *bla*_{SHV-28} was the most frequent narrow-spectrum chromosomal SHV β -lactamase gene identified (57%), but 7 other variants were also detected differing from *bla*_{SHV-28} by a 39-bp insertion (*bla*_{SHV-100}) or by 5 to 9 SNP (all other variants most of which were *bla*_{SHV-1}) (Table S1).

Phylogenetic structure and diversity within *K. pneumoniae* CG15. The 170 *K. pneumoniae* CG15 genomes were subdivided into two major subclades (CG15-I and CG15-II); each of these was further subdivided into two subclades, with all four identified putatively as CG15-IA (9% KL19 and KL106; 7 to 409 SNPs/megabase), CG15-IB (6% variable KL types; 10 to 777 SNPs/megabase), CG15-IIA (43% KL24; 6 to 299 SNPs/megabase), and CG15-IIB (37% KL112; 6 to 346 SNPs/megabase) (Fig. 1 and Table S2). Of note, strains corresponding to highly related genomes carrying KL106, KL110, or KL39 were isolated in neighboring countries and might correspond to local adaptive events (Table S1). The time-scaled phylogenetic inference (root-to-tip regression analysis: $R^2 = 0.3361$ [Fig. S3]) estimates the emergence of the ancestor of CG15 around 1952 (95% HPD, 1929 to 1969), from which derived initially a few diverse CG15-KL24 genomes with wild-type *gyrA* and *parC*, and around 1989 (95% HPD, 1982 to 1993) a major branch including all CG15-I and CG15-II genomes containing a specific set of mutations in the QRDR, namely, *GyrA83F*, *GyrA87A*, and *ParC80I* (Fig. S5). The CG15-IA subclade (KL19) emerged around 1995 (95% HPD, 1990 to 1999), whereas both the majority of CG15-IIA (KL24) and all CG15-IIB (KL112) emerged from a common ancestor (1990; 95% HPD, 1985 to 1994) in 1993 (95% HPD, 1988 to 1998) and 1992 (95% HPD, 1987 to 1996), respectively. The identification of a small set of genomes carrying KL24 ($n = 5$) in CG15-IIB suggests recombination and further selection of the KL112 capsular type. The evolutionary rate within CG15 is estimated at 7.78×10^{-7} substitutions/site/year (95% HPD, 7.28×10^{-7} to 9.33×10^{-7}).

FIG 1 Legend (Continued)

rectangles. The branch tips are colored by the world region of isolation (see key). Capsular locus (KL) type and yersiniabactin-carrying ICEKp are colored according to their variants (see key). Acquired β -lactamases and *GyrA* and *ParC* mutations are indicated by colored rectangles when present, the different variants represented by colors according to the key. 1, dark pink indicates NDM-1 and light pink indicates other NDM variants; 2, dark pink indicates KPC-2 and light pink indicates KPC-3; 3, dark pink indicates OXA-48 and light pink indicates other OXA-48-like variants; 4, dark pink indicates VIM and light pink stands for IMP; 5, dark pink indicates CTX-M-15 and light pink indicates other CTX-M variants; 6, dark pink indicates DHA-1 and light pink stands for CMY.

A Number of accessory genes per strain



B Intra-group shared accessory genes

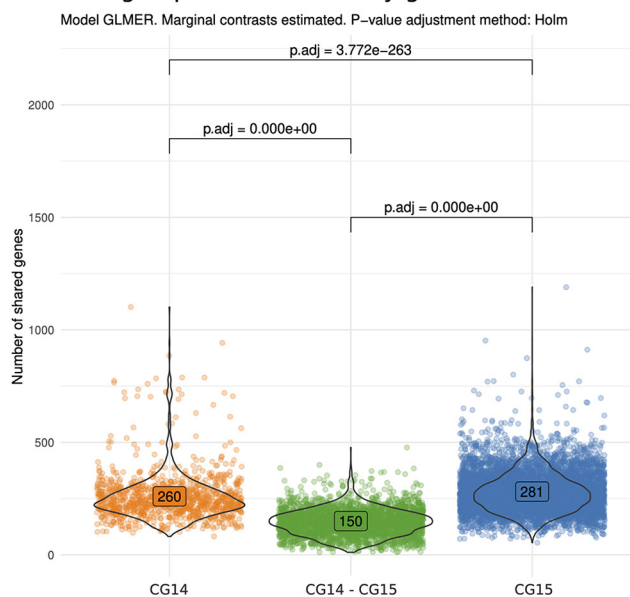


FIG 2 Statistical analysis of the accessory genome size and accessory gene sharedness between CG14 and CG15. (A) Average number of accessory genes in each CG; (B) accessory genes shared within CG14 or CG15 and between CG14 and CG15. Each set of shared genes between each genome was evaluated by correcting the model by the genomic distance between the genomes in order to mitigate the bias produced by sampling bias. The distribution was modeled with a GLME and a Poisson distribution.

The CG15 genomes corresponded to isolates identified between 1980 and 2017 in 32 countries, most of them collected in Europe (59% [100/170]) and in the framework of the EuSCAPE study (<https://pathogen.watch/collections/all?searchText=euscape>) (65% [65/100]). CTX-M-15 or CTX-M-15 and OXA-48 producers from Portugal and sequenced *de novo* in this study were closely related (from 61 to 187 SNPs) to public genomes from other countries in Europe and Western Africa (Table S2). Most CG15 genomes (87%) carried genes encoding ESBL, AmpC, and/or carbapenemases. *bla*_{CTX-M-15} was the predominant gene (68%) and significantly enriched in subclade CG15-II (64% in CG15-I versus 92% in CG15-II; $P = 0.0004$). Variants of *bla*_{CTX-M} (*bla*_{CTX-M-3}, *bla*_{CTX-M-14}, *bla*_{CTX-M-33}, *bla*_{CTX-M-55}, *bla*_{CTX-M-88}, and *bla*_{CTX-M-199}), *bla*_{SHV} (*bla*_{SHV-2} and *bla*_{SHV-12}), *bla*_{AmpC} (*bla*_{DHA-1}, *bla*_{CMY-6} and *bla*_{CMY-16}), or other *bla* genes (*bla*_{OXA-1}, *bla*_{OXA-10}, *bla*_{VEB-5} and *bla*_{GES-11}) were also detected. Among carbapenemase producers (38%), *bla*_{OXA-48-like} genes were predominant (41%), followed by *bla*_{KPC-2} (17%), *bla*_{NDM-1} (16%), and *bla*_{VIM-4} (9%). A higher diversity of β -lactamase-encoding variants was observed in genomes from the CG15-I subclade. Variants of *mcr* (*mcr-1*, *mcr-3*, and *mcr-9*) were identified in five genomes, four of them from Asia. Mutations/disruptions in *ompK35* (26%) and/or *ompK36* (12%) were present throughout the CG15 phylogeny. In contrast with the case of CG14, *bla*_{SHV-28} was identified in most (91%) CG15 isolates, whereas other variants (*bla*_{SHV-1} and *bla*_{SHV-11}) were only occasional (Table S1).

The accessory genome of *K. pneumoniae* CG15 and CG14. We used the Pangenome Analysis Toolkit (PATO) to define and represent the size and composition of the accessory genomes of CG14 and CG15 (see Materials and Methods for details) (30). A total of 15,563 genes were identified as representing the accessory genome of both clonal groups (genes present in <80% of the strains). Despite the average numbers of accessory genes per strain being similar between CG14 ($n = 673$) and CG15 ($n = 645$) (P value = 0.26), the number of genes shared by the genomes of each CG ($n = 260$ in CG14 and $n = 281$ in CG15) is higher than that shared between the genomes of the two CGs ($n = 150$) (P value < 0.001) (Fig. 2), suggesting distinct accessory genomes for CG14 and CG15. The accessory genome network reveals that genomes from subclades CG14-I, CG14-II, CG15-IIA, and CG15-IIB have also a characteristic set of accessory genes, some of them shared between the groups (Fig. S6A). Nevertheless, the size of the common accessory genome in

each subclade ranges from 274 and 329, in accordance with the core genome phylogenetic relatedness, and the differences between the subclades are significant ($P < 0.001$) (Fig. S6B). A number of proteins (104 in CG14 and 117 in CG15) were enriched in each CG (>80% of the genomes; P value < 0.001); 39% and 61% of these, respectively, were annotated as hypothetical proteins. We also found variable numbers of enriched proteins the main subclades as follows: CG14-I, $n = 97$; CG15-IA, $n = 81$; CG15-IIA, $n = 126$; and CG15-IIB, $n = 182$. A high frequency (54%) of subclade enriched proteins were annotated as hypothetical proteins; the remaining were involved in carbohydrate metabolism or were ABC transporters, surface and outer membrane proteins, prophage integrases, toxins, and proteins associated with plasmid conjugation and partitioning. Of note, most (75%) highly specific proteins ($n = 97$; present in >80% of the subclade and in <30% of the total genomes) were observed in the CG14-I subclade (Table S3). Our results show that each CG (and its respective subclades) has its own accessory genome repertoire but with a great plasticity, as it can be deduced from the number of genes shared between CG.

The plasmidome of *K. pneumoniae* CG15 and CG14. Plasmid sequences were identified using MOB-Suite and plasmids were reconstructed from the assembly FASTA files using MOB-recon (34). Plasmid typing on the whole data set (replicon family, relaxase type, mate-pair formation type, and predicted transferability) was performed using MOB-typer (34) and PATO (30). Plasmids sharing at least 50% coverage/similarity were considered a “plasmid group” (PG); data on replication origin, relaxase type, antibiotic resistance genes (ARG), and metal tolerance genes (MTG) were collated to each PG (see Materials and Methods for details). We detected a total of 1,050 plasmids, of which 533 clustered in 27 PGs composed of >5 plasmids (from 5 to 59 plasmids), 184 corresponded to small PGs (from 2 to 4 plasmids), and 333 were singletons (Fig. 3).

The content of the 27 main PGs in replicons and relaxases, their predicted mobility, and their distribution in CG15 and CG14 are shown in Fig. 4. Replicons identified belonged to 9 known types (FII, FIA, FIB, L, HI1B, C, N, R, and Q1), 10 were untypeable (not classified by incompatibility typing schemes available), and 6 were Col types, which were observed alone or in variable combinations. They were part of 10 F-type PGs, 10 Col PGs, and 7 variable non-F-type PGs, 16 of which were abundant (present in >10 genomes). Twelve PGs carried no relaxase, whereas MOBF, MOBPF, MOBPH, and MOBPC types were found, respectively, in 8, 4, 2, and 1 PG. Most plasmids were predicted as nonmobilizable (60%, based on the absence of conjugative machinery and/or relaxase) or mobilizable (32%, if conjugative machinery was absent), and a minority were predicted as conjugative (8%, when conjugative machinery was detected).

Ten PGs consisted of *K. pneumoniae* narrow-host-range F-type plasmids with multiple replicons, all except one highly abundant and distributed (from 16 to 45 genomes) (Fig. S7). They consisted of combinations of FII_{K1/K4} plus FIB_K with or without rep_2183 (50 to 300 kb; PG-2, -4, -6, and -42), FII_{K1}+rep_1418+rep_2183 (60 to 175 kb; PG-30), FIB_{pNDM-MAR}+HI1B_{pNDM-MAR}+rep_1254 (214 to 340 kb; PG-16), R+FIA_{HI1} (17 to 70 kb; PG-7), and different FIB plasmids (50 to 162 kb; PG-15, -52, and -173), most of which (90%) were associated with ABR (see below) (Fig. 4 and Table S4).

Ten PGs corresponded to Col plasmids, most of them (70%) of broad host range (Fig. S7). Four of them were particularly frequent, namely, ColpVC (2 kb; PG-9), ColRNA_rep1857 (6 to 12 kb; PG-32), ColRNAI_rep1987 (4 kb; PG-50), and ColYe4449 (5 to 6 kb; PG-49). Other less frequent Col plasmids included ColMG828 (1.6 kb; PG-82) and other combinations of known and new Col types ($n = 5$ PGs) (Fig. 4C). It is noteworthy that most Col plasmids have been associated with plasmids from non-*Klebsiella* species: *Salmonella enterica* (ColpVC), *Yersinia enterocolitica* (ColYe4449), and *Escherichia coli* (4 Col plasmids, including ColMG828) (Table S4). The remaining PGs included the broad-host-range IncL (46 to 128 kb; PG-1 in 37 genomes), IncN (40 to 50 kb; PG-132 in 8 genomes), or IncC (160 to 180 kb; PG-119 in 6 genomes), the narrow-host-range IncR (PG-232), and new replication types, including rep_cluster1+rep_cluster3 (50 to 60 kb; PG-10) or rep_1418 (10 kb; PG-21 or 20 to 30 kb; PG-41), some of which were highly abundant (PG-10 in 43 genomes and PG-21 in 11 genomes) (Fig. 4C).

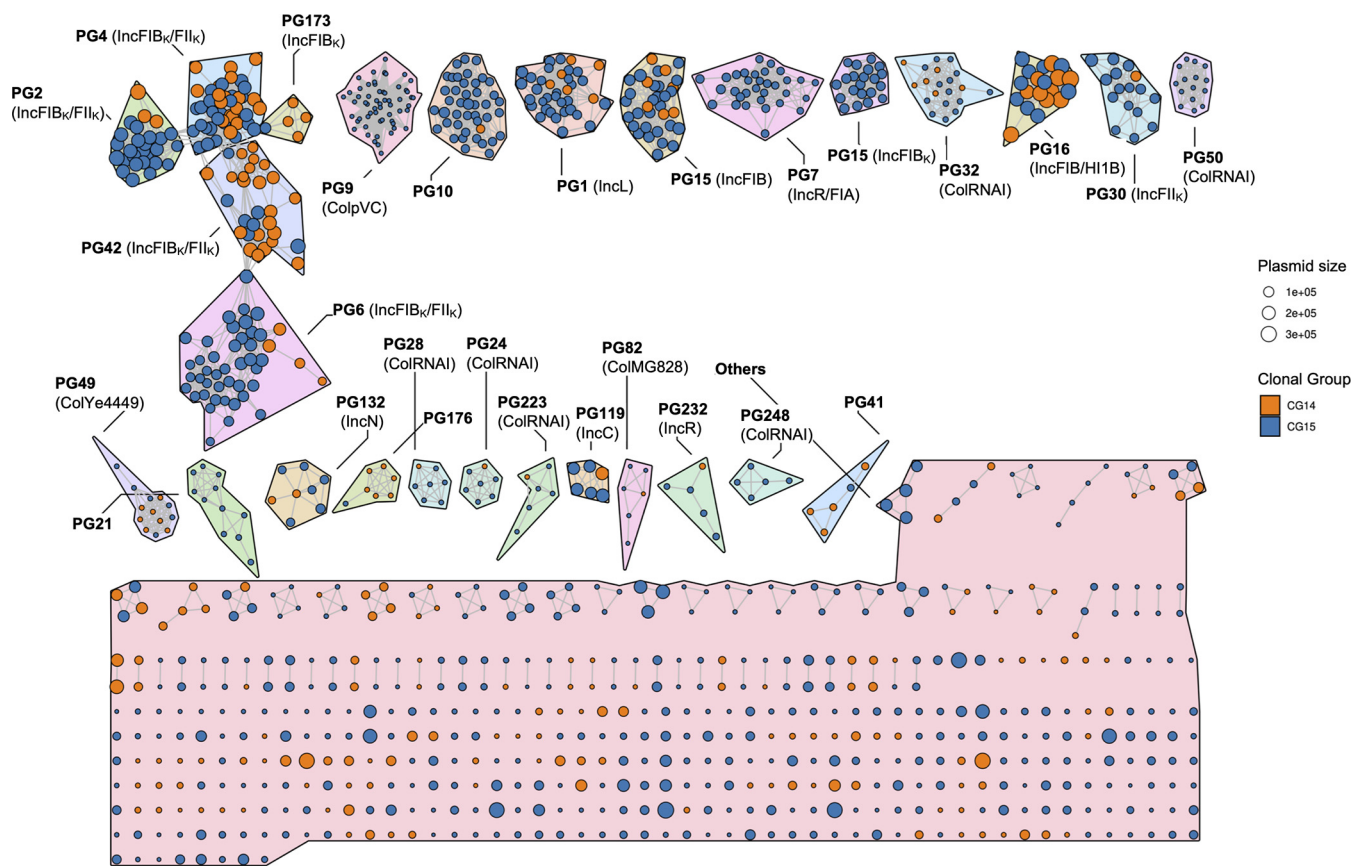


FIG 3 K-nearest neighbor network (K-NNN) of *K. pneumoniae* CG14 and CG15 plasmids. The network was built using the PATO k-NNN function. Each node represents one plasmid and is colored according to the CG. A group of similar plasmids (>5 plasmids) is represented by a plasmid group (PG), plasmid clusters being defined using the Louvain algorithm over the network structure. Each PG is arbitrarily designated with a number, and replicon content defined by MOB-typer is shown in parentheses (34). Differences in color shading between PGs is arbitrary. Each plasmid is connected with the 10 best hits if their Jaccard similarity is at least 0.5 and the difference in size is less than 50%.

We detected an average number of PGs per genome of 4.65 (between 3 and 6), the number increasing from CG14 (average = 3.8) to CG15 (average = 4.98), a statistically significant difference (P value = $1.58e-4$). Plasmids were unevenly distributed in the data set (243 in CG14 and 807 in CG15), some PGs being exclusively found in CG15 (PG-9, -10, -21, -50, and -248; 3 Col plasmids; 2 new plasmids). Thus, we used a hypergeometric test to identify PGs overrepresented in each CG (P value adjusted for multiple comparison < 0.05). We found that PG-4, PG-42, and PG-176 were overrepresented in CG14 and four PGs were overrepresented in subclades CG15-IIA and -B (PG-9 and PG-10; >80%) and CG15-IIB (PG-7 and PG-52; >80%). Only one of these was a multi-drug resistance plasmid (Fig. S8).

The acquired resistome and metalome of *K. pneumoniae* CG14 and CG15.

Fourteen out of the 27 PGs contained ARGs (Fig. S9) and nine of them also carried MTGs (Fig. S10). It is of note that they are plasmids of known IncF, IncC, IncL, IncN, and Col families, most ABR PGs ($n = 10/14$) carrying multiple resistance genes being multireplicon F plasmids. Four PGs carried only *bla*_{TEM-1_r}, *bla*_{OXA-232_r}, *tetA*, or *aph*(3'')-Ib+*aph*(6)-I_d.

The association of main ESBL and carbapenemase genes with particular PGs across the CG14 and CG15 phylogeny reflects the complex trajectories and acquisition routes of these genes among these CGs (Fig. 5). The *bla*_{CTX-M-15} gene was frequently linked to *bla*_{TEM-1A_r}, *bla*_{OXA-1_r}, *df*rA14, *tetA*, *aac*(6')-Ib-cr, *aph*(6)-I_d, and *aph*(3'')Ib on a diversity of plasmid backgrounds ($n = 9$, consisting of variable combinations of FII_{K1/K4_r}, FIB [FIB_{K_r}, FIB_{pNDM-MAR} or FIB_{pKPHS1_r}], HI1B, FIA_{HI1_r}, R, or new replicons L and C) (Fig. S9), suggesting their location on a widespread transposable unit (35). Of note, this putative transposable unit was repeatedly acquired by variable mosaic F-type plasmids several times

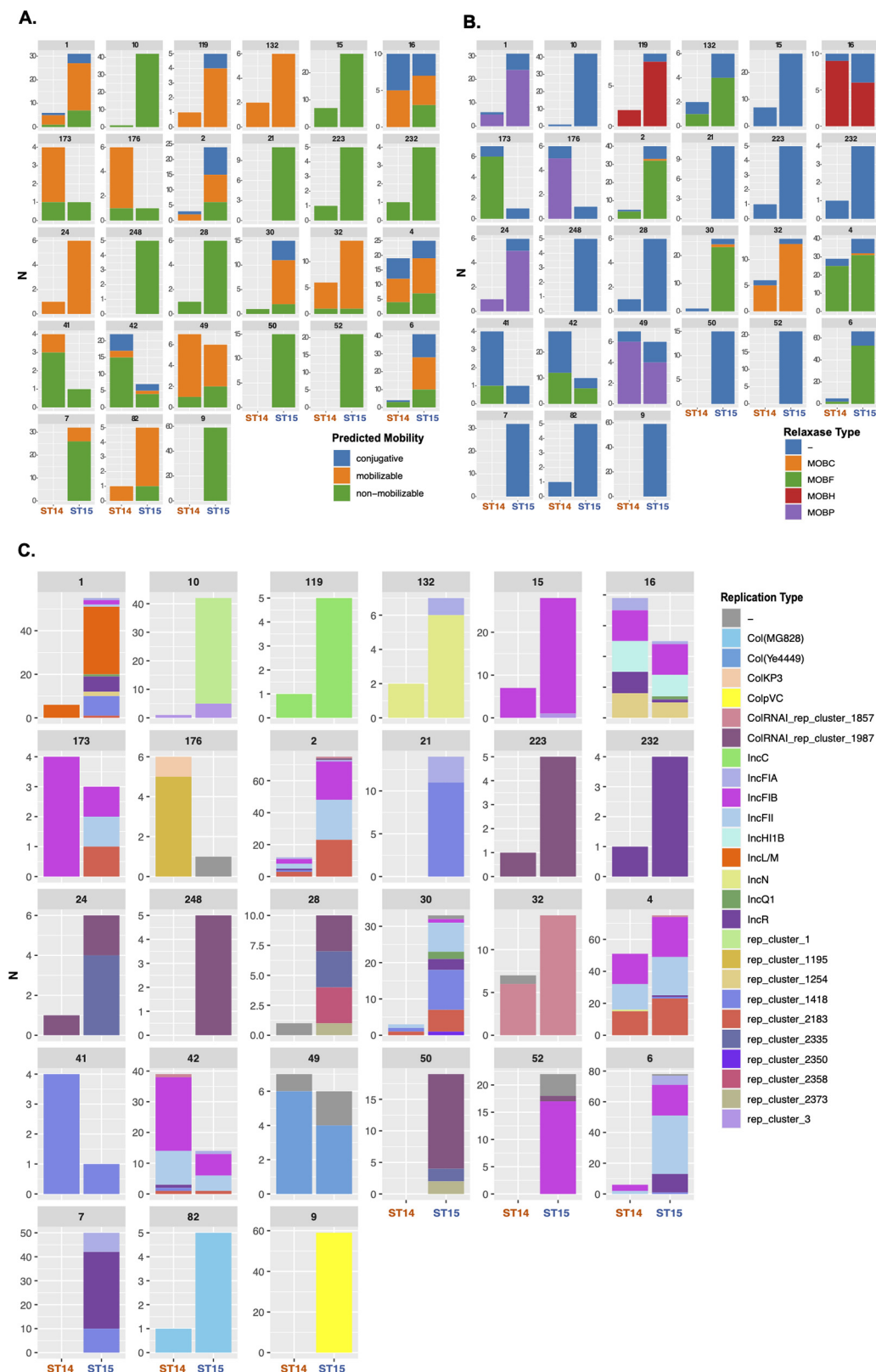


FIG 4 Predicted mobility and replicase and relaxase contents of each plasmid group (PG). (A) Distribution of plasmid predicted mobility within each PG in CG15 and CG14; (B) distribution of replicases (incompatibility group) within each PG in CG15 and CG14; (C) distribution of relaxases (MOB type) within each PG in CG15 and CG14.

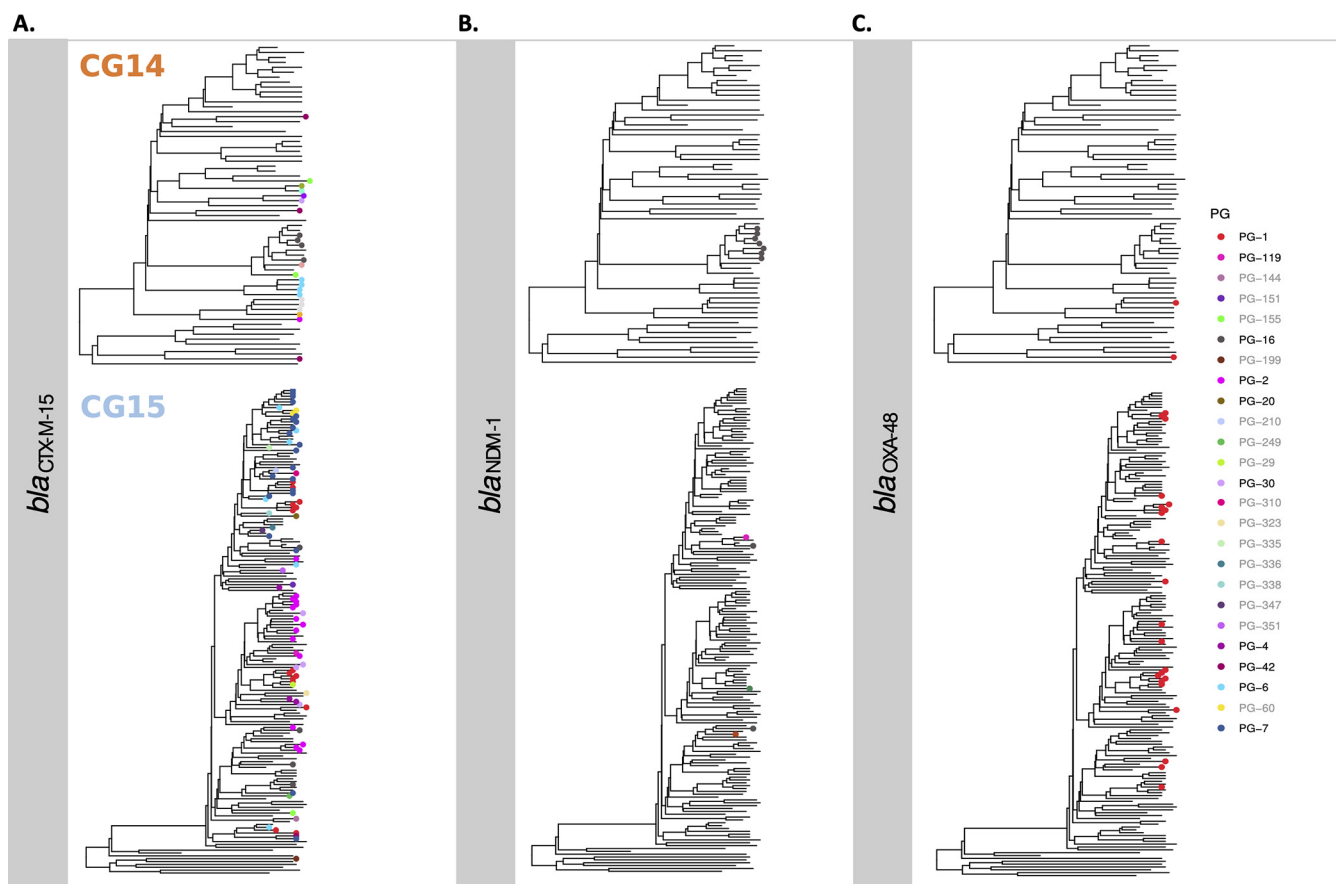


FIG 5 Distribution of plasmid groups carrying *bla*_{CTX-M-15}, *bla*_{NDM-1}, and *bla*_{OXA-48} across the phylogeny. Phylogenetic trees of each CG were obtained with BEAST (represented in Fig. S4 and S5). PGs carrying each β -lactamase gene were associated with each genome in the phylogenetic tree. In the key, black labels represent common PGs (>5 plasmids), and gray labels represent uncommon PGs (<5 plasmids).

across the CG14 and CG15 phylogeny (Fig. 5A). Several of these are highly homologous to those of known hybrid plasmids of *K. pneumoniae*, such as pKPN307-D (GenBank accession number [KY271407](#)), pKP6402 (GenBank accession number [AP018752](#)), and pUUH239 (GenBank accession number [CP002474](#)), all derived from pKPN3 plasmid (GenBank accession number [NC_009649](#) [Table S4]). *bla*_{NDM-1} was acquired once in CG14 by PG-16 (IncHI1B+FIB) and at least four times in CG15 by four different PGs (IncHI1B+FIB, IncC, or other), on some occasions together with *bla*_{CTX-M-15} (Fig. 5B). In contrast, *bla*_{OXA-48} was located only on IncL plasmids related to pOXA-48 (GenBank accession number [JN626286](#)), which was acquired multiple times throughout the CG15 phylogeny and less frequently in CG14 (Fig. 5C).

*bla*_{OXA-232} was exclusively associated with a hybrid Col-type PG-176 in closely related CG14-I (KL2) (Table S4). The *bla*_{TEM-24} or *bla*_{CMY} genes were located on IncC and the *bla*_{SHV-12} or *bla*_{KPC-2} genes on IncN, whereas *bla*_{VIM} and *bla*_{KPC-3} variants were occasionally detected in different PGs of 10 and 7 genomes, respectively. It is of interest that certain predominant PGs carry diverse ESBL or carbapenemase genes, namely, IncC (*bla*_{CTX-M-15}, *bla*_{NDM-1}, *bla*_{SHV-5}, *bla*_{CMY}, or *bla*_{TEM-24}), IncN (*bla*_{KPC-2}, *bla*_{NDM-1}, or *bla*_{SHV-12}) or the FIB-HI1B multireplicon plasmid (*bla*_{CTX-M-15}, *bla*_{NDM-1}, or *bla*_{DHA-1}). PG-16 (FIB+HI1B) carried the highest number and diversity of ARG and showed regions identical to reference NDM-1-encoding plasmids such as pPMK1-NDM and pNDM-MAR (GenBank accession numbers [CP008933](#) and [JN420336](#), respectively).

MTG were mainly located on AMR-PGs ($n = 8/9$; 7 multireplicon F plasmids, 1 IncL plasmid, and 1 IncC plasmid). All multireplicon FII_{K1/K4}+FIB_K plasmids (PG-2, -4, -6, and -42) were enriched in mercury (*mer*), copper (*pco*), silver (*sil*), and arsenic (*ars*) operon genes. Other F-type plasmids carried *mer*, *pco*, and *sil* (PG-16; FIB_{pNDM-MAR}+HI1B), *pco*

and *sil* operons (PG-52, FIB_k) or only *merA* or *merE* genes (PG-7, FIA+R). IncN and IncC plasmids carried different *mer* operons (Fig. S10).

DISCUSSION

This study revealed that the globally distributed *K. pneumoniae* CG14 and CG15 have independent evolutionary trajectories. Moreover, it provided the first comprehensive analysis of the plasmidome of global *K. pneumoniae* CGs showing the contribution of specific plasmids to the expansion of particular CGs and subclades.

The discrimination between CG14 and CG15 described here is also supported by another recent work that consider them as sublineage 15 (SL15) and SL14 according to a nomenclature based on cgMLST and life identification number (LIN) (17). According to this proposal, only the major CG14 subclades are differentiated in two distinct clonal groups (CG14-I as CG14 and CG14-II as CG10201), whereas CG15 subclades belong to the same clonal group (CG15). The deep-branching structure of their core genome shows a diversification that involves capsular switches and further expansion of the dominant subclades carrying specific KL types (KL2 in CG14 and KL24/KL112 in CG15). The most recent common ancestor of the major CG15 subclade (CG15-II) seems to have appeared in 1990s, contemporaneously to other globally distributed sequence types such as ST307 (1994; 95% HPD, 1974 to 2006), ST147 (KL64; 1994; 95% HPD, 1990 to 1998), ST101 (1989; 95% HPD, 1964 to 2007), and ST258 (1995; 95% HPD, 1988 to 1999) (7–10). Conversely, the common recent ancestor of major CG14 subclades seems to have emerged earlier (around 1911 or 1932), which is consistent with the differentiation and nomenclature proposed by Hennart et al. (17). Despite the differences in the predicted emergence dates of different CGs, the evolutionary rate of both CG14 and CG15 is lower (4.76×10^{-7} to 7.78×10^{-7} substitutions/site/year) than that observed for most global CGs (1.03×10^{-6} to 2.85×10^{-6} substitutions/site/year) but similar to that of hypervirulent ST23 (3.4×10^{-7} substitutions/site/year) (7–10). Nonetheless, these differences should be interpreted with caution since they might reflect the heterogeneity and variable robustness of the different data sets, especially those comprising relatively short timescales (36).

We found that CG and subclade differentiation can be explained by specific events involving the capsule and acquisition of mutations associated with fluoroquinolone resistance. Recombination events involving the capsule locus could have allowed the adaptation to human immunity and consequent selection and amplification of certain *K. pneumoniae* subpopulations (37–39), similarly to what has been described for *Streptococcus pneumoniae* or *Neisseria meningitidis* (40, 41). In fact, the recognition of these KL types as suitable evolutionary markers led to their use in diagnostics for categorizing *K. pneumoniae* isolates (including CG15 and CG14) in hospital surveillance and infection control programs using FT-IR spectroscopy (19, 42, 43). Besides the capsule, the acquisition of resistance to fluoroquinolones by most CG15 isolates through specific mutations in *GyrA* and *ParC* also mimics what has been reported for other contemporary high-risk clones of *K. pneumoniae*, *Escherichia coli*, *Staphylococcus aureus*, *Clostridioides difficile*, or *Enterococcus faecium* (7, 8, 44–47). In fact, fluoroquinolone resistance seems to be an advantage for the fitness of high-risk clones of different species (48, 49). Fluoroquinolone-resistant clones predominate in the elderly and people who often have a long history of exposure to health care centers, a risk factor for the acquisition of other antibiotic resistance genes (50, 51).

Though the sample analyzed is overrepresented by genomes from European countries, the detection of all CG14 and CG15 subclades in different continents and wide geographical areas reflects their global distribution (11, 12), similar to what is reported for CG147, CG258, CG307, or CG101. Most genomes carried genes encoding extended-spectrum β -lactamases, acquired AmpCs, and/or carbapenemases (81%), reflecting the sampling bias toward MDR strains from the nosocomial setting. The role of plasmids in the expansion of MDR *K. pneumoniae* CGs remains largely unexplored, probably due to the difficulty in analyzing plasmidomes and mobilomes in large genome data sets.

Although a few recent studies have analyzed plasmids of a high number of genomes from *Enterobacteriales* using large-scale network analysis (28, 29), this work is the first study using such an approach to characterize the plasmidome of particular *K. pneumoniae* CGs.

Our results revealed an overrepresentation of F and Col plasmids, whose distribution in other CGs is largely unknown. The high degree of mosaicism observed, especially on narrow-host-range multireplicon F-type plasmids, is in agreement with previous observations and is suggestive of the impact of recombination in the dissemination of ARG and/or *K. pneumoniae* adaptation and of a degree of specialization resulting from long-term plasmid-host coevolution (52–54). The high heterogeneity of small plasmids (essentially ColE1 type) long evolving in *Enterobacteriales*, some of them particularly abundant and linked to CG15 (ColpVC and ColRNAI_rep1987), is also of note. Col plasmids have been implicated in extraintestinal virulence and/or gastrointestinal colonization in *E. coli* (55, 56). In *K. pneumoniae*, their role is still largely unexplored, although the association with adaptive traits such as ARG (e.g., *bla*_{OXA-232}), cell metabolism, virulence, defense from phages, and heavy metal resistance, together with their variability and transposase content (IS26, IS4321, Tn3-like), suggests an important role in adaptation (26, 55, 57–59). The diversity of plasmids harboring replicons untypeable by classical plasmid typing schemes (23, 60) reinforces the need to improve current *K. pneumoniae* plasmid categorization approaches (34, 61). Cryptic plasmids (50 to 60 kb; PG-10 or 10 kb; PG-21; noncarriers of ABR genes) were common in a major CG15 subclade, suggesting a role for unknown features in adaptation. The high occurrence of nonconjugative/nonmobilizable plasmids is not surprising and reflects the high recombination of the mobile genetic elements within and/or between high-risk clones with frequent deletion and acquisition events (61). However, their mobilization is not compromised because of the association with helper entities (e.g., conjugative plasmids and ICEs) (52, 62, 63).

One of the most important observations is that only a few known PGs are responsible for the dissemination of ARG and MTG, specifically, multireplicon F plasmids and those of classical L, C, and N incompatibility groups. Of particular relevance is the high frequency of mosaic pKPN-3-like plasmids in CG15, previously observed in other CGs, which confers on them a significant role in the acquisition, spread, and persistence of contemporaneous ESBL and carbapenemase genes (*bla*_{CTX-M-15}, *bla*_{KPC}, and *bla*_{NDM-1}) in *K. pneumoniae* (63–68). The acquisition of ARG and/or MTG by these plasmid backgrounds seems to have contributed to the diversification and expansion of CG15, particularly the GC15-IIB subclade (KL112). In fact, *bla*_{CTX-M-15} is the most frequently acquired *bla*_{ESBL} gene (38% in CG14 versus 68% in CG15; $P = 0.00006$), increasing in prevalence from CG14-I (45%), CG15-IIA (62%), and CG15-IIB (92%). The repeated acquisition of the putative transposable unit containing *bla*_{CTX-M-15} by multiple F-type plasmids might have been key for this expansion, as reported for *E. coli* ST131 (69). It is also worth highlighting the high proportion of genes encoding acquired β -lactamases (71%; especially *bla*_{CTX-M-15}, *bla*_{OXA-232}, and/or *bla*_{NDM-1}) identified in the CG14-I subclade, considering that KL2 is frequently associated with hypervirulent and antibiotic-susceptible *K. pneumoniae* (22). Our data reveal multiple trajectories for the dissemination of ARG across different plasmid and genetic backgrounds: (i) a single plasmid across multiple lineages (e.g., *bla*_{OXA-48}-IncL), (ii) multiple plasmids across multiple lineages (e.g., *bla*_{CTX-M-15} or *bla*_{NDM-1} in multiple plasmids across the phylogeny), and (iii) multiple plasmids in one lineage (e.g., *bla*_{CTX-M-15} in plasmids from CG15-IIB). By providing opportunities for plasmid shuffling and evolution, these trajectories contribute to the expansion and maintenance of critical ARG in the *K. pneumoniae* population.

Conclusions. This study demonstrates an independent evolutionary trajectory for CG15 and CG14, marked by expansion of subclades with particular KL types (CG14-KL2 and CG15-KL24/KL112), GyrA and ParC mutations (within CG15), and ARG/MTG in a high diversity of mosaic plasmids. A comprehensive analysis of the plasmidome revealed the contribution of a turbulent flux of plasmids, particularly pervasive F-type and Col plasmids, and adaptive traits (ARGs, MTG, and VFs) to the pangenome, which

reflect the exposure to disparate selective pressures of MDR *Enterobacteriales*. Finally, this study further highlights the relevance of plasmid data for the understanding of the evolutionary trajectories of ARG and *K. pneumoniae* subpopulations.

MATERIALS AND METHODS

Bacterial isolates and sequencing. Seven MDR ST15 (3 K24, 1 KL39, 1 KL112, 1 KL19, and 1 KL110) and 2 MDR ST14 (1 KL2 and 1 KL16) isolates were selected for whole-genome sequencing (WGS). These isolates are representative of five CG15 and CG14 lineages previously defined by MLST, PFGE, and FT-IR spectroscopy, circulating as ESBL and/or carbapenemase producers in Portuguese hospitals between 2003 and 2013 (18–21) (Table 1). Bacterial genomic DNA was extracted using the QIAamp DNA minikit (Qiagen), and DNA concentration and purity were measured using a Qubit fluorometer (Life Technologies) and NanoDrop 2000 instrument (Thermo Scientific), respectively. DNA libraries were prepared using the Nextera XT kit (Illumina, San Diego, CA, USA), and 2×300 -bp paired-end sequence reads with mean coverage of $100\times$ were generated on the MiSeq platform (Illumina). *De novo* assembly was performed with SPAdes v3.9.0 using *k*-mers of 101, 111, 121, and 127 (70), and the quality of the assemblies was evaluated using QUAST (71). The whole-genome shotgun project was deposited in DDBJ/EMBL/GenBank under BioProject accession number [PRJNA408270](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA408270), and assembly statistics concerning the nine isolates sequenced are available in Table 1.

Publicly available genomes and global data set. In order to capture CG15 and CG14 diversity and provide robustness to phylogenetic analysis, a total data set of 481 publicly available CG14 ($n = 110$; 1980 to 2018) and CG15 ($n = 371$; 1980 to 2017) *K. pneumoniae* genomes from NCBI RefSeq (accessed on November 2019) were downloaded, provided they complied with the control quality criteria defined (genome size [4.9 to 6.2 Mb] and GC content [56 to 58%] matching with *K. pneumoniae* and fewer than 1,000 contigs or N_{50} of $>20,000$) (Table S5 and Fig. S11). Including our 9 genomes, we had an initial data set of 490 *K. pneumoniae* genomes. Of these, all the genomes without available information concerning the isolation year ($n = 20$) were excluded. Furthermore, 226 genomes (30 CG14 and 196 CG15) showing <21 single nucleotide polymorphisms (SNPs) in the core genome alignment obtained with Roary were discarded since they were epidemiologically related, considering the cutoff proposed by David et al. (11). Furthermore, to evaluate the strength of the temporal signal of our data set and to depict problematic or erroneous sequences (“outliers”), we conducted a linear regression analysis of the root-to-tip genetic distances as a function of the sample collection year, using TempEst v1.5.3 (<http://tree.bio.ed.ac.uk/software/tempest/>), after which we discarded 9 additional genomes, resulting in 235 deduplicated high-quality genomes.

Our global final data set ($n = 235$) included 65 CG14 (63 ST14 and 2 SLV) and 170 CG15 (162 ST15 and 8 SLV) nonduplicated genomes originated mainly from Europe (55%), Western and South-Eastern Asia (16%), and North America (7%) (Fig. S1). The vast majority were recovered from humans (97%, 75% in the context of infection) and carried genes encoding extended-spectrum β -lactamases, acquired AmpCs, and/or carbapenemases (81%). All the epidemiological and genomic data concerning the final data set of genomes are shown in Table S1.

Core genome phylogenetic analysis. To conduct the phylogenetic analyses, the 235 genomes from our final data set were previously annotated using Prokka 1.12 (72) and a core genome alignment was constructed with Roary v3.12 (73) using a blastP identity cutoff of 90% and core genes defined as those being present in more than 90% of the isolates, resulting in a total of 4,420 core genes. Afterwards, 35,783 single-nucleotide variants (SNVs) were extracted from the core genome alignment with SNP-sites (74) and used to construct a maximum likelihood phylogenetic tree with IQ-TREE v1.6.11 (model GTR+F+ASC+R4) rooted with two closely related outgroup genomes based on the study by Wyres et al. (15): *K. pneumoniae* ST540 isolate Kpn0019 (SRA accession number [SRR2098710](https://www.ncbi.nlm.nih.gov/sra/SRR2098710)) and *K. pneumoniae* ST101 isolate Kp_Goe_33208 (GenBank assembly accession number [GCF_001902435.1](https://www.ncbi.nlm.nih.gov/genbank/GCF_001902435.1)).

Since CG15 and CG14 evolved separately (Fig. 1) and to improve the reliability of the temporal signal and the inferred predictions, we decided to perform timescaled phylogenetic inference for each CG independently. For this purpose, we generated a recombination-free alignment using Snippy v4.6.0 (<https://github.com/tseemann/snippy>) and Gubbins v1.12 (75), and we evaluated the strength of the temporal signal for each CG (R^2 for CG14 = 0.1474; R^2 for CG15 = 0.3361) (Fig. S3A and B). The timescaled phylogenetic inference was obtained with BEAST v2.6.2 (run with a Markov chain Monte Carlo length of 1×10^9 , sampling every 5×10^3 steps) (76). We used model parameters that had the best fit: GTR substitution model, lognormal relaxed clock, and constant population size. Parameter estimates were computed using Tracer v1.7.1, and a maximum clade credibility tree was obtained with TreeAnnotator v2.6.0 and visualized in FigTree v1.4.4. Additionally, to confirm the significance of our temporal signatures, we conducted a series of date randomization tests for our CG14 data set, running BEAST with true dates in duplicate and in 5 randomized tip date data sets (Fig. S3C).

MLST/cgMLST and genomic analyses of surface polysaccharides, antimicrobial resistance, virulence, and heavy metal tolerance genes. MLST and cgMLST were performed using the Institut Pasteur *Klebsiella* MLST database (<https://bigsdbs.pasteur.fr/klebsiella/>) (6, 13, 17). Genes associated with antimicrobial resistance and virulence (siderophores [yersiniabactin, aerobactin, and salmochelin], colibactin, and regulators of mucoid phenotype [RmpA and RmpA2]) were searched using Kleborate v1.0 (<https://github.com/katholt/Kleborate/wiki>) (12). Characterization of the capsular locus (KL) and lipopolysaccharide (LPS) O antigen was performed through Kaptive, integrated in Kleborate (12, 77). Additional virulence factors, such as iron uptake systems (*kfu* operon), as well as heavy metal tolerance

genes (arsenic [*ars* operon], copper [*pco* cluster], silver/copper [*sil* cluster], tellurite [*ter* operon], and mercury [*mer* operon]) were searched using BIGSdb (<https://bigsdbs.pasteur.fr/klebsiella/>), whereas the *kpi* chaperone-usher pilus system was searched using BLAST (32). Statistical analyses to check the association of the different categorical variables within the clades defined were performed using the χ^2 test (P values of <0.05 were considered statistically significant).

Accessory genome analysis. The accessory genome was defined as all genes present in less than 80% of the genomes and was obtained using PATO with default parameters (80% identity and 80% coverage) (30), followed by a detailed study of the protein enrichment in each of the CGs. Accessory gene enrichment analysis according to CG and the respective subclades (KL types) was performed using the function `accnet_enrichment_analysis` from PATO. This function performs a multihyperparametric test to find overrepresented genes in a cluster in comparison with the general population (i.e., the genome data set). Table S3 lists the genes enriched in the CGs with a P value of <0.001 and present in more than 75% of the genomes of the CG, together with their respective annotation. To illustrate the accessory network, we use Gephi for the rearrangement and then we loaded the layout in R to plot the network (Fig. S6A).

The statistical test of accessory genes shared within each CG or within each subclade was performed modeling the distribution of pair sharedness of genes between genomes of each clade/subclade and further correction according to the similarity between genomes. A generalized linear mixed-effects model (lmer R package) was built using the sharedness distribution under a Poisson distribution and corrected with the genomic distance calculated by MASH. The statistical test of the number of accessory genes per genome was performed by nonpaired Student t test (stats R package).

Plasmid identification and analysis. Plasmids were reconstructed using MOB-recon from MOB-suite software (34). MOB-recon automatically reconstructs sequence plasmids from assembly FASTA files. All 1,050 plasmids predicted and reconstructed by MOB-recon were typed with MOB-typer and then annotated with Prokka (72). A plasmidome was set with PATO by clustering the proteins of each plasmid in protein families of $>70\%$ identity. Using the data of presence/absence of each protein family produced by PATO, we built a similarity matrix of Jaccard distances. Then we created a k -nearest neighbor network (K-NNN) with PATO with a parameter of 10 neighbors, allowing reciprocal connections and with a similarity threshold of 0.5 (Jaccard distance) (Table S6). This means that each plasmid is linked to its 10 most similar plasmids as long as they share more than 0.5 Jaccard similarity. In order to remove spurious links, we removed those links in which there was a difference in sequence length between the plasmids of $>50\%$ (i.e., the smallest plasmid had to represent at least 50% of the largest plasmid). The plasmid network was arranged by Cytoscape software (78) and imported to R again using the `tidygraph` R package. Once the network was built, we used the Louvain cluster algorithm (`igraph` R package). Finally, we incorporated the plasmid typing information (predicted mobility, replicases, and relaxases) to the network. Plasmid distribution over phylogenetic trees was built using `ggtree` R software. All the data manipulation and visualization were performed using the `tidyverse` R metapackage.

Data availability. The whole-genome shotgun project was deposited in DDBJ/EMBL/GenBank under BioProject accession number [PRJNA408270](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA408270).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 4.4 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.9 MB.

ACKNOWLEDGMENTS

This work is financed by national funds from FCT—Fundação para a Ciência e a Tecnologia, I.P., in the scope of project UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences—UCIBIO and project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy—i4HB. Carla Rodrigues was supported by a fellowship from FCT through POCH (grant number SFRH/BD/84341/2012), and work developed at T.M.C.'s lab was supported by a FEMS Research Grant (FEMS-RG-2014-0089). Ângela Novais is supported by national funds through FCT, I.P., through the Scientific Employment Stimulus Program [2021.02252.CEECIND/CP1662/CT0009]. T.M.C.'s lab is supported by the European Commission (grant MISTAR AC21/2 00041) and the Instituto de Salud Carlos III (ISCIII) of Spain, cofinanced by the European Development Regional Fund (A Way to Achieve Europe program; Spanish Network for Research in Infectious Diseases grants PI18/1942, PI21/02027, and CIBER CB21/13/00084) and the Fundación Ramón Areces (BioMetasep). Francisco Soria Melguizo (MicroCarbaflux). V.F.L. acknowledges financial support from a Miguel Servet contract from Instituto de Salud Carlos III (ISCIII) (grant no. CP22/00164), cofunded by the European Social Fund, 'Investing in your future' and Centro de Investigación Biológica en Red de Enfermedades Infecciosas (CIBERINFEC), cofunded with European Regional Development Fund.

We declare that no competing interests exist.

REFERENCES

- Navon-Venezia S, Kondratyeva K, Carattoli A. 2017. *Klebsiella pneumoniae*: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol Rev* 41:252–275. <https://doi.org/10.1093/femsre/fux013>.
- Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outtersson K, Patel J, Cavalieri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* 18: 318–327. [https://doi.org/10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3).
- Antimicrobial Resistance Collaborators. 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399:629–655. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0).
- Wyres KL, Lam MMC, Holt KE. 2020. Population genomics of *Klebsiella pneumoniae*. *Nat Rev Microbiol* 18:344–359. <https://doi.org/10.1038/s41579-019-0315-1>.
- Russo TA, Marr CM. 2019. Hypervirulent *Klebsiella pneumoniae*. *Clin Microbiol Rev* 32:e00001-19. <https://doi.org/10.1128/CMR.00001-19>.
- Bialek-Davenet S, Criscuolo A, Ailloud F, Passet V, Jones L, Delannoy-Vieillard AS, Garin B, Le Hello S, Arlet G, Nicolas-Chanoine MH, Decré D, Brisse S. 2014. Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis* 20:1812–1820. <https://doi.org/10.3201/eid2011.140206>.
- Wyres KL, Hawkey J, Hetland MAK, Fostervold A, Wick RR, Judd LM, Hamidian M, Howden BP, Löhr IH, Holt KE. 2019. Emergence and rapid global dissemination of CTX-M-15-associated *Klebsiella pneumoniae* strain ST307. *J Antimicrob Chemother* 74:577–581. <https://doi.org/10.1093/jac/dky492>.
- Rodrigues C, Desai S, Passet V, Gajjar D, Brisse S. 2022. Genomic evolution of the globally disseminated multidrug-resistant *Klebsiella pneumoniae* clonal group 147. *Microb Genom* 8:000737. <https://doi.org/10.1099/mgen.0.000737>.
- Roe CC, Vazquez AJ, Esposito EP, Zarrilli R, Sahl JW. 2019. Diversity, virulence, and antimicrobial resistance in isolates from the newly emerging *Klebsiella pneumoniae* ST101 lineage. *Front Microbiol* 10:542. <https://doi.org/10.3389/fmicb.2019.00542>.
- Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D, de Man T, Rasheed JK, Engelthaler DM, Keim P, Limbago BM. 2015. Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS One* 10:e0133727. <https://doi.org/10.1371/journal.pone.0133727>.
- David S, ESGEM Study Group, Reuter S, Harris SR, Glasner C, Feltwell T, Argimon S, Abudahab K, Goater R, Giani T, Errico G, Aspbury M, Sjunnebo S, Feil EJ, Rossolini GM, Aanensen DM, Grundmann H. 2019. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat Microbiol* 4:1919–1929. <https://doi.org/10.1038/s41564-019-0492-8>.
- Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. 2021. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat Commun* 12:4188. <https://doi.org/10.1038/s41467-021-24448-3>.
- Diancourt L, Passet V, Verhoef J, Grimont PAD, Brisse S. 2005. Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *J Clin Microbiol* 43:4178–4182. <https://doi.org/10.1128/JCM.43.8.4178-4182.2005>.
- Breurec S, Guessennd N, Timinouni M, Le TAH, Cao V, Ngandjio A, Randrianirina F, Thiberge JM, Kinana A, Dufougeray A, Perrier-Gros-Claude JD, Boisier P, Garin B, Brisse S. 2013. *Klebsiella pneumoniae* resistant to third-generation cephalosporins in five African and two Vietnamese major towns: multiclonal population structure with two major international clonal groups, CG15 and CG258. *Clin Microbiol Infect* 19:349–355. <https://doi.org/10.1111/j.1469-0691.2012.03805.x>.
- Wyres KL, Wick RR, Judd LM, Froumine R, Tokolyi A, Gorrie CL, Lam MMC, Duchêne S, Jenney A, Holt KE. 2019. Distinct evolutionary dynamics of horizontal gene transfer in drug resistant and virulent clones of *Klebsiella pneumoniae*. *PLoS Genet* 15:e1008114. <https://doi.org/10.1371/journal.pgen.1008114>.
- Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, Brisse S, Cao H, Wilksch J, Gorrie C, Schultz MB, Edwards DJ, van Nguyen K, Nguyen TV, Dao TT, Mensink M, Minh V, Nhu NTK, Schultsz C, Kuntaman K, Newton PN, Moore CE, Strugnell RA, Thomson NR, van Nguyen K, Nguyen TV, Dao TT, Mensink M, Le Minh V, Nhu NTK, Schultsz C, Kuntaman K, Newton PN, Moore CE, Strugnell RA, Thomson NR. 2015. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci U S A* 112:E3574–E3581. <https://doi.org/10.1073/pnas.1501049112>.
- Hennart M, Guglielmini J, Bridel S, Maiden MCJ, Jolley KA, Criscuolo A, Brisse S. 2022. A dual barcoding approach to bacterial strain nomenclature: genomic taxonomy of *Klebsiella pneumoniae* strains. *Mol Biol Evol* 39:msac135. <https://doi.org/10.1093/molbev/msac135>.
- Rodrigues C, Machado E, Ramos H, Peixe L, Novais A. 2014. Expansion of ESBL-producing *Klebsiella pneumoniae* in hospitalized patients: a successful story of international clones (ST15, ST147, ST336) and epidemic plasmids (IncR, IncFIK). *Int J Med Microbiol* 304:1100–1108. <https://doi.org/10.1016/j.ijmm.2014.08.003>.
- Rodrigues C, Sousa C, Lopes JA, Novais A, Peixe L. 2020. A front line on *Klebsiella pneumoniae* capsular polysaccharide knowledge: Fourier transform infrared spectroscopy as an accurate and fast typing tool. *mSystems* 5:e00386-19. <https://doi.org/10.1128/mSystems.00386-19>.
- Rodrigues C, Bavlovič J, Machado E, Amorim J, Peixe L, Novais A. 2016. KPC-3-producing *Klebsiella pneumoniae* in Portugal linked to previously circulating non-CG258 lineages and uncommon genetic platforms (Tn4401d-IncFIA and Tn4401d-IncN). *Front Microbiol* 7:1000. <https://doi.org/10.3389/fmicb.2016.01000>.
- Machado E, Coque TM, Cantón R, Novais A, Sousa JC, Baquero F, Peixe L, Cantón R, Novais A, Sousa JC, Baquero F, Peixe L, Portuguese Resistance Study Group, Cantón R, Novais A, Sousa JC, Baquero F, Peixe L. 2007. High diversity of extended-spectrum beta-lactamases among clinical isolates of Enterobacteriaceae from Portugal. *J Antimicrob Chemother* 60: 1370–1374. <https://doi.org/10.1093/jac/dkm381>.
- Wyres KL, Holt KE. 2016. *Klebsiella pneumoniae* population genomics and antimicrobial-resistant clones. *Trends Microbiol* 24:944–956. <https://doi.org/10.1016/j.tim.2016.09.007>.
- Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, Aarestrup FM, 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>.
- Achtman M, Zhou Z. 2014. Distinct genealogies for plasmids and chromosome. *PLoS Genet* 10:e1004874. <https://doi.org/10.1371/journal.pgen.1004874>.
- Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics* 32:3380–3387. <https://doi.org/10.1093/bioinformatics/btw493>.
- Lanza VF, de Toro M, Pilar Garcillá N-, Barcia M, Mora A, Blanco J, Coque TM, de La Cruz F. 2014. Plasmid flux in *Escherichia coli* ST131 sublineages, analyzed by Plasmid Constellation Network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. *PLoS Genet* 10:e10047. <https://doi.org/10.1371/journal.pgen.1004766>.
- Orlek A, Stoesser N, Anjum MF, Doumith M, Ellington MJ, Peto T, Crook D, Woodford N, Walker AS, Phan H, Sheppard AE. 2017. Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front Microbiol* 8:182. <https://doi.org/10.3389/fmicb.2017.00182>.
- Redondo-Salvo S, Bartomeus-Peñalver R, Vielva L, Tagg KA, Webb HE, Fernández-López R, de la Cruz F. 2021. COPLA, a taxonomic classifier of plasmids. *BMC Bioinformatics* 22:390. <https://doi.org/10.1186/s12859-021-04299-x>.
- Lipworth S, Matlock W, Shaw L, Vihta K-D, Rodger G, Chau K, Barker L, George S, Kavanagh J, Davies T, Vaughan A, Andersson M, Jeffery K, Oakley S, Morgan M, Hopkins S, Peto T, Crook D, Walker AS, Stoesser N. 4 April 2022. The mobilome associated with Gram-negative bloodstream infections: a large-scale observational hybrid sequencing based study. *medRxiv*. <https://doi.org/10.1101/2022.04.03.22273290>.
- Fernández-De-Bobadilla MD, Talavera-Rodríguez A, Chacon L, Baquero F, Coque TM, Lanza VF. 2021. PATO: Pangenome Analysis Toolkit. *Bioinformatics* 37:4564–4566. <https://doi.org/10.1093/bioinformatics/btab697>.
- Fernández-de-Bobadilla MD, Coque TM, Lanza VF. 2020. Statistical analysis of accessory genome. *Methods Mol Biol* 2075:341–353. https://doi.org/10.1007/978-1-4939-9877-7_24.
- Gato E, Vázquez-Ucha JC, Rumbo-Feal S, Álvarez-Fraga L, Vallejo JA, Martínez-Gutián M, Beceiro A, Vivas JR, Sola Campoy PJ, Pérez-Vázquez M, Iglesias JO, Rodiño-Janeiro BK, Romero A, Poza M, Bou G, Pérez A. 2020. Kpi, a chaperone-usher pili system associated with the worldwide-disseminated high-risk clone *Klebsiella pneumoniae* ST-15. *Proc Natl Acad Sci U S A* 117: 17249–17259. <https://doi.org/10.1073/pnas.1921393117>.
- David S, Wong JLC, Sanchez-Garrido J, Kwong H-S, Low WW, Morecchiato F, Giani T, Rossolini GM, Brett SJ, Clements A, Beis K, Aanensen DM, Frankel G. 2022. Widespread emergence of *OmpK36* loop 3 insertions

- among multidrug-resistant clones of *Klebsiella pneumoniae*. *PLoS Pathog* 18:e1010334. <https://doi.org/10.1371/journal.ppat.1010334>.
34. Robertson J, Nash JHE. 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom* 4:e000206. <https://doi.org/10.1099/mgen.0.000206>.
 35. Abdallah R, Kuete Yimagou E, Hadjadj L, Mediannikov O, Ibrahim A, Davoust B, Barciela A, Hernandez-Aguilar RA, Diatta G, Sokhna C, Raoult D, Rolain JM, Baron SA. 2022. Population diversity of antibiotic resistant Enterobacteriales in samples from wildlife origin in Senegal: identification of a multidrug resistance transposon carrying blaCTX-M-15 in *Escherichia coli*. *Front Microbiol* 13:838392. <https://doi.org/10.3389/fmicb.2022.838392>.
 36. Duchêne S, Holt KE, Weill F-X, Le Hello S, Hawkey J, Edwards DJ, Fourment M, Holmes EC. 2016. Genome-scale rates of evolutionary change in bacteria. *Microb Genom* 2:e000094. <https://doi.org/10.1099/mgen.0.000094>.
 37. Patro LPP, Rathinavelan T. 2019. Targeting the sugary armor of *Klebsiella* species. *Front Cell Infect Microbiol* 9:367. <https://doi.org/10.3389/fcimb.2019.00367>.
 38. Chen L, Mathema B, Pitout JDD, DeLeo FR, Kreiswirth BN. 2014. Epidemic *Klebsiella pneumoniae* ST258 is a hybrid strain. *mBio* 5:e01355-14. <https://doi.org/10.1128/mBio.01355-14>.
 39. Mostowy RJ, Holt KE. 2018. Diversity-generating machines: genetics of bacterial sugar-coating. *Trends Microbiol* 26:1008–1021. <https://doi.org/10.1016/j.tim.2018.06.006>.
 40. Croucher NJ, Finkelstein JA, Pelton SJ, Mitchell PK, Lee GM, Parkhill J, Bentley SD, Hanage WP, Lipsitch M. 2013. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat Genet* 45:656–663. <https://doi.org/10.1038/ng.2625>.
 41. Snape MD, Pollard AJ. 2005. Meningococcal polysaccharide-protein conjugate vaccines. *Lancet Infect Dis* 5:21–30. [https://doi.org/10.1016/S1473-3099\(04\)01251-4](https://doi.org/10.1016/S1473-3099(04)01251-4).
 42. Guerra AM, Lira A, Lameirão A, Selaru A, Abreu G, Lopes P, Mota M, Novais Â, Peixe L. 2020. Multiplicity of carbenemase-producers three years after a KPC-3-producing *K. pneumoniae* ST147-K64 hospital outbreak. *Antibiotics (Basel)* 9:806. <https://doi.org/10.3390/antibiotics9110806>.
 43. Novais Â, Ferraz RV, Viana M, da Costa PM, Peixe L. 2022. NDM-1 introduction in Portugal through a ST11 KL105 *Klebsiella pneumoniae* widespread in Europe. *Antibiotics (Basel)* 11:92. <https://doi.org/10.3390/antibiotics11010092>.
 44. Stoesser N, Sheppard AE, Pankhurst L, de Maio N, Moore CE, Sebra R, Turner P, Anson LW, Kasarskis A, Batty EM, Kos V, Wilson DJ, Phetsouvanh R, Wyllie D, Sokurenko E, Manges AR, Johnson TJ, Price LB, Peto TEA, Johnson JR, Didelot X, Walker AS, Crook DW. Modernizing Medical Microbiology Informatics Group (MMMIG). 2016. Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *mBio* 7:e02162-15. <https://doi.org/10.1128/mBio.02162-15>.
 45. Willems RJJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F, Grundmann H, Bonten MJM. 2005. Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerg Infect Dis* 11:821–828. <https://doi.org/10.3201/1106.041204>.
 46. Wasels F, Kuehne SA, Cartman ST, Spigaglia P, Barbanti F, Minton NP, Mastrantonio P. 2015. Fluoroquinolone resistance does not impose a cost on the fitness of *Clostridium difficile* in vitro. *Antimicrob Agents Chemother* 59:1794–1796. <https://doi.org/10.1128/AAC.04503-14>.
 47. Holden MTG, Hsu LY, Kurt K, Weinert LA, Mather AE, Harris SR, Strommenger B, Layer F, Witte W, de Lencastre H, Skov R, Westh H, Žemličková H, Coombs G, Kearns AM, Hill RLR, Edgeworth J, Gould I, Gant V, Cooke J, Edwards GF, McAdam PR, Templeton KE, McCann A, Zhou Z, Castillo-Ramírez S, Feil EJ, Hudson LO, Enright MC, Balloux F, Aanensen DM, Spratt BG, Fitzgerald JR, Parkhill J, Achtman M, Bentley SD, Nübel U. 2013. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res* 23:653–664. <https://doi.org/10.1101/gr.147710.112>.
 48. Fuzi M, Szabo D, Cserssik R. 2017. Double-serine fluoroquinolone resistance mutations advance major international clones and lineages of various multi-drug resistant bacteria. *Front Microbiol* 8:2261. <https://doi.org/10.3389/fmicb.2017.02261>.
 49. Redgrave LS, Sutton SB, Webber MA, Piddock LJV. 2014. Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success. *Trends Microbiol* 22:438–445. <https://doi.org/10.1016/j.tim.2014.04.007>.
 50. Baquero F, Martínez JL, Novais Â, Rodríguez-Beltrán J, Martínez-García L, Coque TM, Galán JC. 2021. Allogeneous selection of mutational collateral resistance: old drugs select for new resistance within antibiotic families. *Front Microbiol* 12:757833. <https://doi.org/10.3389/fmicb.2021.757833>.
 51. Fuzi M, Rodríguez Baño J, Toth A. 2020. Global evolution of pathogenic bacteria with extensive use of fluoroquinolone agents. *Front Microbiol* 11:271. <https://doi.org/10.3389/fmicb.2020.00271>.
 52. Sandegren L, Linkevicius M, Lytsy B, Melhus Å, Andersson DI. 2012. Transfer of an *Escherichia coli* ST131 multiresistance cassette has created a *Klebsiella pneumoniae*-specific plasmid associated with a major nosocomial outbreak. *J Antimicrob Chemother* 67:74–83. <https://doi.org/10.1093/jac/dkr405>.
 53. Pesesky MW, Tilley R, Beck DAC. 2019. Mosaic plasmids are abundant and unevenly distributed across prokaryotic taxa. *Plasmid* 102:10–18. <https://doi.org/10.1016/j.plasmid.2019.02.003>.
 54. Wang X, Zhao J, Ji F, Chang H, Qin J, Zhang C, Hu G, Zhu J, Yang J, Jia Z, Li G, Qin J, Wu B, Wang C. 2021. Multiple-replicon resistance plasmids of *Klebsiella* mediate extensive dissemination of antimicrobial genes. *Front Microbiol* 12:754931. <https://doi.org/10.3389/fmicb.2021.754931>.
 55. Ares-Arroyo M, Rocha EPC, Gonzalez-Zorn B. 2021. Evolution of ColE1-like plasmids across γ -Proteobacteria: from bacteriocin production to antimicrobial resistance. *PLoS Genet* 17:e1009919. <https://doi.org/10.1371/journal.pgen.1009919>.
 56. Cummins ML, Reid CJ, Djordjevic SP. 2022. F plasmid lineages in *Escherichia coli* ST95: implications for host range, antibiotic resistance, and zoonosis. *mSystems* 7:e01212-21. <https://doi.org/10.1128/mSystems.01212-21>.
 57. Ramirez MS, Tran T, Tolmasky ME, Lin DL, Traglia GM. 2014. Plasmid-mediated antibiotic resistance and virulence in Gram-negatives: the *Klebsiella pneumoniae* paradigm. *Microbiol Spectr* 2:2.5.36. <https://doi.org/10.1128/microbiolspec.PLAS-0016-2013>.
 58. Rodríguez-Beltrán J, Hernández-Beltrán JCR, Delafuente J, Escudero JA, Fuentes-Hernández A, MacLean RC, Peña-Miller R, San Millán A. 2018. Multicopy plasmids allow bacteria to escape from fitness trade-offs during evolutionary innovation. *Nat Ecol Evol* 2:873–881. <https://doi.org/10.1038/s41559-018-0529-z>.
 59. Hendrickx APA, The Dutch CPE surveillance Study Group, Landman F, de Haan A, Borst D, Witteveen S, van Santen-Verheul MG, van der Heide HGJ, Schouls LM, Halaby T, Steingrover R, Cohen Stuart JWT, Melles DC, van Dijk K, Spijkerman IJB, Notermans DW, Oudbier JH, van Ogtrop ML, van Dam A, den Reijer M, Kluytmans JAJW, van der Linden MPM, Mattsson EE, van der Vusse M, de Jong E, Majjer-Reuwer A, van Trijp M, van Griethuysen AJ, Ott A, Bathoorn E, Sinnige JC, Heikens E, de Brauer EIGB, Stals FS, Silvius W, Dorigo-Zetsma JW, Waar K, van Mens SP, Roescher N, Voss A, Wertheim H, Slingerland BCGC, Frenay HME, Schulien T, Diederens BMW, Bode L, van Rijn M, Dinant S, Damen M, de Man P, Leversteijn-van Hall MA, van Elzakker EPM, Muller AE, Schneeberger P, van Dam DW, Buiting AGM, Vlek ALM, Stam A, Troelstra A, Overdeest ITMA, Bosboom RW, Trienekens TAM, Wolfhagen MJHM, Paltansing S. 2020. Plasmid diversity among genetically related *Klebsiella pneumoniae* blaKPC-2 and blaKPC-3 isolates collected in the Dutch national surveillance. *Sci Rep* 10:16778. <https://doi.org/10.1038/s41598-020-73440-2>.
 60. Garcillán-Barcia MP, Redondo-Salvo S, Vielva L, de la Cruz F. 2020. MOBscan: automated annotation of MOB relaxases. *Methods Mol Biol* 2075:295–308. https://doi.org/10.1007/978-1-4939-9877-7_21.
 61. Coluzzi C, Garcillán-Barcia MP, de la Cruz F, Rocha EPC. 2022. Evolution of plasmid mobility: origin and fate of conjugative and nonconjugative plasmids. *Mol Biol Evol* 39:msac115. <https://doi.org/10.1093/molbev/msac115>.
 62. Ares-Arroyo M, Coluzzi C, Rocha EPC. 2022. Origins of transfer establish networks of functional dependencies for plasmid transfer by conjugation. *Nucleic Acids Res* 29:gkac1079. <https://doi.org/10.1093/nar/gkac1079>.
 63. Löhr IH, Hülter N, Bernhoff E, Johnsen PJ, Sundsfjord A, Naseer U. 2015. Persistence of a pKPN3-like CTX-M-15-encoding IncFIIK plasmid in a *Klebsiella pneumoniae* ST17 host during two years of intestinal colonization. *PLoS One* 10:e0116516. <https://doi.org/10.1371/journal.pone.0116516>.
 64. García-Fernández A, Villa L, Carta C, Venditti C, Giordano A, Venditti M, Mancini C, Carattoli A. 2012. *Klebsiella pneumoniae* ST258 producing KPC-3 identified in Italy carries novel plasmids and OmpK36/OmpK35 porin variants. *Antimicrob Agents Chemother* 56:2143–2145. <https://doi.org/10.1128/AAC.05308-11>.
 65. Shin J, Soo Ko K. 2014. Single origin of three plasmids bearing blaCTX-M-15 from different *Klebsiella pneumoniae* clones. *J Antimicrob Chemother* 69:969–972. <https://doi.org/10.1093/jac/dkt464>.
 66. Stoesser N, Giess A, Batty EM, Sheppard AE, Walker AS, Wilson DJ, Didelot X, Bashir A, Sebra R, Kasarskis A, Sthapit B, Shakya M, Kelly D, Pollard AJ, Peto TEA, Crook DW, Donnelly P, Thorson S, Amatya P, Joshi S. 2014. Genome sequencing of an extended series of NDM-producing *Klebsiella pneumoniae* isolates from neonatal infections in a Nepali hospital characterizes the extent of community- versus hospital-associated

- transmission in an endemic setting. *Antimicrob Agents Chemother* 58:7347–7357. <https://doi.org/10.1128/AAC.03900-14>.
67. Villa L, Feudi C, Fortini D, Brisse S, Passet V, Bonura C, Endimiani A, Mammina C, Ocampo AM, Jimenez JN, Doumith M, Woodford N, Hopkins K, Carattoli A. 2017. Diversity, virulence, and antimicrobial resistance of the KPC-producing *Klebsiella pneumoniae* ST307 clone. *Microb Genom* 3:e000110. <https://doi.org/10.1099/mgen.0.000110>.
 68. David S, the European Survey of Carbapenemase-Producing Enterobacteriaceae (EuSCAPE) Working Group, Cohen V, Reuter S, Sheppard AE, Giani T, Parkhill J, Rossolini GM, Feil EJ, Grundmann H, Aanensen DM, Koraqi A, Lacey D, Apfalter P, Hartl R, Glupczynski Y, Huang T-D, Strateva T, Marteva-Proevska Y, Andrasevic AT, Butic I, Pieridou-Bagatzouni D, Maikanti-Charalampous P, Hrabak J, Zemlickova H, Hammerum A, Jakobsen L, Ivanova M, Pavelkovich A, Jalava J, Österblad M, Dortet L, Vaux S, Kaase M, Gatermann SG, Vatopoulos A, Tryfinopoulou K, Tóth Á, Jánvári L, Boo TW, McGrath E, Carmeli Y, Adler A, Pantosti A, Monaco M, Raka L, Kurti A, Balode A, Saule M, Miciuleviciene J, Mierauskaite A, Perrin-Weniger M. 2020. Integrated chromosomal and plasmid sequence analyses reveal diverse modes of carbapenemase gene spread among *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A* 117:25043–25054. <https://doi.org/10.1073/pnas.2003407117>.
 69. Johnson TJ, Danzeisen JL, Youmans B, Case K, Llop K, Munoz-Aguayo J, Flores-Figueroa C, Aziz M, Stoesser N, Sokurenko E, Price LB, Johnson JR. 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>.
 70. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 71. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
 72. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 73. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
 74. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2:e000056. <https://doi.org/10.1099/mgen.0.000056>.
 75. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 43:e15. <https://doi.org/10.1093/nar/gku1196>.
 76. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, de Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, Du Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 15:e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>.
 77. Wick RR, Heinz E, Holt KE, Wyres KL. 2018. Kaptive Web: user-friendly capsule and lipopolysaccharide serotype prediction for *Klebsiella* genomes. *J Clin Microbiol* 56:e00197-18. <https://doi.org/10.1128/JCM.00197-18>.
 78. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>.
 79. Rodrigues C, Novais Â, Machado E, Peixe L. 2014. Detection of VIM-34, a novel VIM-1 variant identified in the intercontinental ST15 *Klebsiella pneumoniae* clone. *J Antimicrob Chemother* 69:274–275. <https://doi.org/10.1093/jac/dkt314>.