


RESEARCH

Open Access



Genomic analysis of the international high-risk clonal lineage *Klebsiella pneumoniae* sequence type 395

Elvira R. Shaidullina^{1†}, Michael Schwabe^{2†}, Thomas Rohde², Valeria V. Shapovalova³, Marina S. Dyachkova³, Alina D. Matsvay³, Yuliya A. Savochkina³, Andrey A. Shelenkov⁴, Yulia V. Mikhaylova⁴, Katharina Sydow², François Lebreton⁵, Evgeny A. Idelevich^{6,7}, Stefan E. Heiden², Karsten Becker⁶, Roman S. Kozlov¹, German A. Shipulin³, Vasily G. Akimkin⁴, Michael Lalk⁸, Sebastian Guenther⁹, Andreas E. Zautner¹⁰, Jürgen A. Bohnert⁶, Ayslu M. Mardanova¹¹, Ruth Bouganim¹², Dror Marchaim¹³, Katharina J. Hoff¹⁴, Katharina Schaufler^{2,15*†}  and Mikhail V. Edelstein^{1†}

Abstract

Background *Klebsiella pneumoniae*, which is frequently associated with hospital- and community-acquired infections, contains multidrug-resistant (MDR), hypervirulent (hv), non-MDR/non-hv as well as convergent representatives. It is known that mostly international high-risk clonal lineages including sequence types (ST) 11, 147, 258, and 307 drive their global spread. ST395, which was first reported in the context of a carbapenemase-associated outbreak in France in 2010, is a less well-characterized, yet emerging clonal lineage.

Methods We computationally analyzed a large collection of *K. pneumoniae* ST395 genomes ($n = 297$) both sequenced in this study and reported previously. By applying multiple bioinformatics tools, we investigated the core-genome phylogeny and evolution of ST395 as well as distribution of accessory genome elements associated with antibiotic resistance and virulence features.

Results Clustering of the core-SNP alignment revealed four major clades with eight smaller subclades. The subclades likely evolved through large chromosomal recombination, which involved different *K. pneumoniae* donors and affected, *inter alia*, capsule and lipopolysaccharide antigen biosynthesis regions. Most genomes contained acquired resistance genes to extended-spectrum cephalosporins, carbapenems, and other antibiotic classes carried by multiple plasmid types, and many were positive for hypervirulence markers, including the siderophore aerobactin. The detection of “hybrid” resistance and virulence plasmids suggests the occurrence of the convergent ST395 pathotype.

Conclusions To the best of our knowledge, this is the first study that investigated a large international collection of *K. pneumoniae* ST395 genomes and elucidated phylogenetics and detailed genomic characteristics of this emerging high-risk clonal lineage.

[†]Elvira R. Shaidullina and Michael Schwabe are shared first authors.

[†]Katharina Schaufler and Mikhail V. Edelstein are shared last authors.

*Correspondence:

Katharina Schaufler

schaufler@infmed.uni-kiel.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords *Klebsiella pneumoniae*, Recombination, Plasmid, Hypervirulence, Multidrug-resistance, Whole genome sequencing, Carbapenem resistance, ESBL

Background

Klebsiella pneumoniae is an opportunistic pathogen and one of the most common causes of hospital- and community-acquired infections [1]. It has been traditionally differentiated into classic (frequently but not exclusively multidrug-resistant [MDR]) and hypervirulent *K. pneumoniae* (hvKp) with the latter appearing clinically as mostly antibiotic-susceptible and invasive pathogen able to induce multiple site infection [2–4]. However, a recent study revealed that definition boundaries are blurred and that a strict assignment to particular contexts (hospital- or community-associated) not always applies [5]. Various virulence-associated features enable some *Klebsiella* strains' pathogenicity including significant siderophore secretion and hypermucoviscosity [6]. The presence of certain biomarkers, for example, *peg-344* (metabolite transporter), *iucA* (aerobactin), *iroB* (salmochelin), and *rmpA* (regulator hypermucoviscosity), has been shown to predict hypervirulence phenotypes and distinguish hvKp from non-hvKp strains [6, 7]. However, the above-mentioned study [5] also suggests that for some settings, these markers are rather unreliable.

Several sequence types (ST) and clonal lineages have been associated with either antimicrobial resistance (AMR) or hypervirulence characteristics; however, the convergence of both has also been increasingly reported during the last few years [8]. The emergence of isolates resistant to most or even all clinically available antibiotics yet exacerbates the overall tense situation and is mostly driven by the acquisition of mobile resistance genes [9]. This includes, for example, the production of the plasmid-encoded extended-spectrum beta-lactamase (ESBL) and carbapenemase (KPC, NDM, and OXA-48) enzymes as well as mobilizable colistin resistance (*mcr*) genes [10]. MDR *K. pneumoniae* isolates often belong to clonal complex (CC)258 (including ST11, ST258, ST340, and ST512) and other established high-risk clonal lineages: ST15, ST101, ST147, and ST307, among others, as defined by multilocus sequence typing (MLST) [11–14], that disseminate globally and are often involved in nosocomial outbreaks [15]. A so far less well-characterized genetic lineage, ST395, is an emerging international high-risk clonal lineage that has been associated with MDR and found to harbor ESBL and carbapenemases, as well as many other resistance determinants including 16S rRNA methyltransferases conferring pan-aminoglycoside resistance [16]. ST395 differs from ST11, the closest member within CC258, in three of seven MLST loci and,

therefore, does not belong to CC258 by classical MLST definition. However, genome-wide analysis of ST395 revealed that it has likely evolved from ST11 through recombination and acquisition of large genomic regions from unrelated STs, and thus, in a broader sense, it may be considered a sub-lineage within CC258 [17]. ST395 was first reported from an outbreak of OXA-48 producing *K. pneumoniae* in a French hospital in 2010 [18]. Shortly after, in 2011, OXA-48 producing isolates from The Netherlands and Morocco were linked to the same sequence type/clone based on MLST and repetitive element sequence-based PCR [19]. ST395 isolates carrying OXA-48 plasmids have spread widely and were recovered either sporadically or from outbreaks in many countries in Europe (the Czech Republic, Denmark, France, Germany, Hungary, Ireland, Italy, Romania, Russia, Sweden, the UK), North Africa (Egypt, Algeria), Southeast Asia (Malaysia), and the Middle East (Israel, Kuwait) [18–37]. Besides OXA-48, isolates of this genotype have been reported to carry different carbapenemases, such as KPC-2 in China [38], KPC-3 in Italy [39], NDM-1 in Russia [40, 41], and NDM-1 and NDM-5 in Germany [8]. Most of the above reports noted simultaneous production of carbapenemases and ESBL, predominantly of the CTX-M group.

This study aimed to genomically characterize the *K. pneumoniae* ST395 clonal lineage in-depth by using different bioinformatics approaches including phylogenetics, genome evolution, and distribution of acquired AMR and virulence genes.

Methods

Data origin

A sample library of 297 *K. pneumoniae* ST395 genomes (assembled genomes and paired-end reads) was acquired, including both previously unpublished and publicly available data. A list of all genomic data used, and their origin, can be found in Table 1 and Additional file 1: Table S1. Whole genome sequence data were either provided by collaborating partners ($n = 138$) or downloaded from public databases ($n = 159$). For the former, a total of 121 new whole genome sequences of ST395 isolates with the corresponding metadata were obtained, which included 108 isolates collected from 20 cities across Russia between 2013 and 2018, ten isolates collected from three cities of Belarus in 2014, and three isolates collected from one city in Kazakhstan in 2017. All these isolates were referred to the central laboratory of the

Table 1 Summary of the whole genome sequences used in this study

Country	Genome source ^a	Isolation source ^b	Carbapenemase gene-positive	ESBL and/or AmpC gene-positive	16S rRNA methyltransferase gene-positive	Hypervirulence gene-positive
Bangladesh	- / 1	1 / - / - / -	1	1	1	-
Belarus	10 / 10	20 / - / - / -	15	18	15	6
Brazil	- / 1	1 / - / - / -	1	-	-	-
China	- / 8	4 / 3 / - / 1	5	8	6	3
Estonia	- / 10	10 / - / - / -	-	10	-	-
Finland	- / 1	1 / - / - / -	1	1	-	1
France	- / 1	1 / - / - / -	-	-	-	-
Georgia	- / 1	1 / - / - / -	1	1	-	-
Germany	1 / 8	9 / - / - / -	6	8	2	3
India	- / 7	7 / - / - / -	5	3	3	1
Israel	- / 7	7 / - / - / -	1	7	1	-
Italy	- / 13	13 / - / - / -	10	10	-	10
Kazakhstan	3 / -	3 / - / - / -	3	3	3	3
Lebanon	- / 1	- / - / 1 / -	-	-	-	-
Lithuania	- / 1	1 / - / - / -	-	1	1	-
Luxembourg	- / 2	1 / - / - / 1	-	2	-	-
Netherlands	- / 1	- / - / - / 1	-	1	-	-
Nigeria	- / 7	7 / - / - / -	6	7	-	-
Portugal	- / 1	1 / - / - / -	-	1	-	-
Romania	- / 5	1 / - / 4 / -	4	5	-	-
Russia	108 / 59	163 / - / - / 4	153	147	66	102
South Korea	- / 4	4 / - / - / -	-	4	-	-
Switzerland	- / 2	2 / - / - / -	2	2	2	-
Turkey	- / 8	8 / - / - / -	8	6	8	-
Ukraine	- / 3	3 / - / - / -	3	3	-	-
USA	- / 9	9 / - / - / -	2	8	1	1
Vietnam	- / 4	4 / - / - / -	3	4	4	-
Total:	122 / 175	282 / 3 / 5 / 7	230	261	113	130

^a Genome Source: This study / Public databases

^b Isolations source: Human / Animal / Environment / Not available

Institute of Antimicrobial Chemotherapy (IAC) as part of the national sentinel AMR surveillance program [42]. In this surveillance, each participating center (hospital) contributed annually up to 150 consecutive non-duplicate (one per patient/case of infection) clinical isolates of different species recovered from representative specimens (blood, cerebrospinal fluid, tissue biopsies, bronchoalveolar lavage, endotracheal aspirate, sputum, urine, etc.) of hospitalized patients with clinical symptoms of infection. Isolates recovered from patients without clinical symptoms of infection or from hospital environment for epidemiological screening purposes, as well as repeated isolates of the same species recovered from the same patient, were spared. In the IAC laboratory, the species identity of isolates was confirmed by MALDI-TOF mass spectrometry with the Microflex LT-MALDI Biotyper

System (Bruker Daltonics, Bremen, Germany), susceptibility to various antibiotics, including meropenem, was determined by broth microdilution method according to ISO 20776–2:2021 [43] and EUCAST [44] methodology, and the presence of carbapenemase genes was assessed by means of real-time PCR in all *Enterobacterales* isolates with decreased susceptibility to meropenem (MIC > 0.125 mg/L) as recommended by EUCAST [45]. A maximum of 20 carbapenemase gene-positive *K. pneumoniae* isolates consecutively collected from each hospital participating in the surveillance were then selected for genome sequencing. Short-read paired-end sequencing of these isolates was performed on Illumina platforms and then used to construct draft genome assemblies. For a subset of isolates, long-read sequencing on MinION (Oxford Nanopore Technologies) was

additionally performed to obtain hybrid genome assemblies. The genome sequences of ST395 isolates from Russia, Kazakhstan, and Belarus obtained for this study were submitted to the European Nucleotide Archive under the project accession number PRJEB49683 [46].

In addition, collaboration partners provided previously unpublished sequence read data of 17 isolates from Israel, the United States (US) of America and Germany. The sequence read data from Israel comprised six isolates (020-1_S1, 020-12_S12, 020-3_S1, 020-4_S4, 020-7_S7, 020-8_S8), which have been previously described by Bouganim *et al.* [47]. The nine genomes from the USA corresponded to isolates from six American patients (MRSN31307, MRSN31314, MRSN31330, MRSN31332, MRSN31333, MRSN8909), one German patient (MRSN363431), and two Ukrainian patients (MRSN567231, MRSN740795). The two German isolates were MRGN_20, which has been previously described by Zautner *et al.* [48] and PBIO2028, a clinical isolate from Greifswald. The associated metadata is listed in Additional file 1: Table S1. Raw reads for the isolates from Israel and Germany were submitted to the European Nucleotide Archive under the project accession number PRJEB53512 [49]. Assemblies of the US samples have been deposited with links to BioProject accession number PRJNA849970 [50] in the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>).

Additionally, to obtain already published *K. pneumoniae* ST395 genomic data, first, literature research was performed using the keywords: “*Klebsiella pneumoniae*” or “*K. pneumoniae*” and “sequence type 395” or “ST395.” Then, paired-end reads were downloaded, upon availability. Second, the NCBI (National Center for Biotechnology Information) assembly database was searched with the taxonomic ID of *Klebsiella pneumoniae* “taxid573.” The sequence types (ST) of genomes were determined using MLST (v. 2.19.0) [51]. For ST395 genomes, the NCBI’s Short Read Archive (SRA) was searched for paired-end read data. If no paired-end reads were available, the available assemblies were used for further analysis.

Short-read whole genome sequencing

For the abovementioned whole genome-sequenced isolates, extracted genomic DNA was used for library preparation after preliminary ultrasonic fragmentation using M220 Focused-ultrasonicator (Covaris, USA). Repair and dA-tailing of DNA fragments was performed using the NEBNext Ultra II End Repair / dA-Tailing Module (New England Biolabs, USA) according to the manufacturer’s protocol. Ligation of 0.25 μ M KAPA DI Adapters (Roche Diagnostics GmbH, Germany) was performed overnight using T4 DNA ligase (New England Biolabs, USA). Libraries were amplified using the DreamTaq PCR Master

Mix (2x) (Thermo Fisher Scientific, USA) according to the manufacturer’s protocol, PCR was stopped at saturation. All clean-up and size selection procedures were carried out using AMPure XP beads (Beckman Coulter Life Sciences, USA) according to the manufacturer’s protocol. Quality of libraries was assessed by capillary electrophoresis using the Agilent 2100 Bioanalyzer system (Agilent, USA). Library concentration was measured as described above. Sequencing was performed on the Illumina NextSeq 550 System with the NextSeq 500/550 Mid Output Kit v2.5 (300 Cycles) (Illumina, USA).

Long-read whole genome sequencing

Extracted genomic DNA was used for library preparation with NEBNext Companion Module (New England Biolabs, USA) and Ligation Sequencing Kit (Oxford Nanopore Technologies, UK) in accordance with the manufacturer’s protocol. Samples were multiplexed using Native Barcoding Expansion 1–12 (Oxford Nanopore Technologies, UK). Library concentration was measured with the Qubit 4 Fluorometer (Thermo Fisher Scientific, USA) using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, USA). Sequencing was performed on the MinION platform (Oxford Nanopore Technologies, UK) using the R9.4.1 Flow Cell.

Data quality and processing of raw data

Raw sequencing reads were quality checked with FastQC [52]. Adapter sequences and low-quality bases were trimmed with Trimmomatic v.0.39 run in paired-end mode [53]. After trimming, eukaryotic contaminations were identified with KRAKEN2 (v.2.1.2) [54] using the standard database (v.2021.05.17). The KRAKENTOOLS toolkit (v.1.2) was used to remove the eukaryotic reads (taxid: 2759) from the paired-end FASTQ files. De novo genome assemblies of contamination-filtered reads were performed with SPAdes [55]. The resulting contig files were used for further analyses.

Base-calling and demultiplexing of Oxford Nanopore data were conducted using Guppy v.5.0.16 with the “sup” model. The quality control was performed with MinIONQC (v.1.4.2) [56]. Hybrid assemblies were prepared by using the Trycycler pipeline v.0.5.1 [57]. The consensus contigs were polished with Medaka v.1.4.4 [58]. Short-read polishing was performed by Polypolish tool [59] and POLCA script [60].

Processing of assembled data

The contamination removal method was based on determining coverage cutoffs for each assembly after visual inspection of coverage-versus-length (CVL) plots as described by Douglass *et al.* [61]. The k-mer coverage value was provided by SPAdes assembler.

Phylogenetic analyses

We performed core-genome SNP analysis of all the genomes against the previously published, fully assembled genome of PBIO1951 using Parsnp v.1.5.6 [62]. PBIO1951 has been selected due to its complete genome and the described epidemiologic link between St. Petersburg (Russia) and Greifswald (Germany) [8]. The core-genome alignment was clustered using BAPS with the fastBAPS R package 1.0.6 [63]. Visualization and annotation of phylogenetic trees were performed using iTOL v.6 [64].

Identification of large recombination events and putative donors

The complete genome of the PBIO1951 isolate was used as the reference genome. All other ST395 genomes from the dataset were mapped to the reference genome using Snippy [65]. Putative regions of recombination were predicted by Gubbins [66]. The Gubbins output files were used to calculate the mean recombination count per base (Additional file 2).

The method described by Comandatore *et al.* in [67] was then used to identify the donors of the large recombination events detected within ST395. Briefly, a total of 1303 complete genome assemblies of *K. pneumoniae* isolates representing different STs were retrieved from the Patric database [68]. They were merged to the ST395 dataset and aligned to the reference genome of PBIO1951. The core SNPs were called again using Snippy, and those corresponding to each recombination region of >100 kbp predicted by Gubbins were extracted and subjected to separate phylogenetic analysis using the FastTree tool [69, 70] with 100 resamples.

Each resulting tree was manually analyzed as follows: (i) the ST395 recipients of the recombination were identified on the tree and (ii) when a monophylum including all the recipients and one or more non-ST395 strains was detected, the latter were considered as putative donors of the recombination.

Identification of virulence, antimicrobial resistance genes, and essential genetic elements

Identification of open reading frames (ORFs) and gene contents in the assembled genomes was performed using Prokka v.1.14.6 [71]. Antimicrobial resistance genes were annotated using AMRFinderPlus tool v.3.10.5 [72] with BLASTP, BLASTX, and HMMER algorithms. Detection of virulence factors was performed by using the ABRicate program v.1.0.1 with BIGSdb-Kp virulence typing databases [73, 74]. Kleborate v.2.1.0 [75] and Kaptive [76] databases were used for MLST, capsule polysaccharide (K), and

lipopolysaccharide (O) antigen molecular typing. The K-locus was classified as “ambiguous” when Kaptive was unable to confidently identify a K-locus due to fragmented or incomplete genome assemblies, or the presence of atypical *wzi* allelic variants or of atypical gene composition of a K-locus. BIGSdb-Kp [77] online tools and databases were used for core-genome MLST (cgMLST) and virulence typing (YbST) and for identification of ICEKp elements [12, 74]. Insertion sequences were determined by ISfinder online tool [78]. Prophages were searched using the online PHASTER tool [79]. Partially matching sequences, point mutations, and indels were further manually curated by using QIAGEN CLC Genomics Workbench 22.

Plasmid analysis

MOB-suite with MOB-recon and MOB-typer software modules [80] was used to partially reconstruct and type the plasmids carrying the carbapenemase and aerobactin cluster genes from short-read genome assemblies. Representative plasmid sequences shown in Fig. 4 were extracted from hybrid short- and long-read assemblies obtained using Trycycler as described above and visualized with the BLAST Ring Image Generator (BRIG) tool v.0.95 [81] and NCBI BLAST + v.2.12.0 [82].

A detailed overview of the used programs and settings can be found in Additional file 3: Detailed Methods.

Results

ST395 genomes were obtained from different countries with the majority originating from Russia

Overall, we included 297 *K. pneumoniae* ST395 genomes from 27 different countries and four continents, which were either collected and sequenced in this study or had sequence data available in public repositories. Most of the Russian isolates (108/167; 64.7%) were collected as part of the national sentinel AMR surveillance program [42] and subjected to whole genome sequencing to identify carbapenemase-producing *K. pneumoniae*. Therefore, the sample set was also biased towards carbapenemase- and MDR-positive isolates. Human clinical isolates represented the vast majority (282/297; 94.9%) compared to environmental (5/297; 1.7%) and animal (3/297; 1%) isolates (for the remaining 7/297 [2.4%] genomes retrieved from public databases, source information was missing [Table 1]).

The documented diseases from which the human clinical isolates were recovered included respiratory (52/199; 26.1%), urinary (46/199; 23.1%), bloodstream (42/199; 21.1%), skin and soft tissue (14/199; 7%), and other (14/199; 7%) infections (Additional file 1: Table S1).

ST395 subclades evolved through recombination and insertion of large chromosomal regions

Alignment of all ST395 genomes against the fully assembled PBIO1951 reference revealed a common core genome of about 3.87 Mb with a total of 12,379 SNP positions identified by Parsnp analysis that were used to construct the phylogenetic tree (Fig. 1). BAPS clustering of the core-SNP alignment assigned all genomes to four well-separated clades designated A, B, C, and D, and eight smaller subclades named A1, A2, B1, B2, C, D1, D2, and D3, respectively. The minimum inter-clade distances (expressed as the number of SNPs between the closest representatives of two clades) ranged from 765 for clades A and B (median distance: 810.0) to 3034 for clades C and D (median distance: 3499.0) (Fig. 1B).

We observed a significant correlation between phylogenetic clustering and the landscape of certain chromosomal markers (Fig. 2). Most notably, we found strict concordance between the type of capsule synthesis locus (KL), which determines the capsule serotype (K-type), and the subclades within ST395. Thus, all but four isolates of subclade A1 carried the KL108 (K47 type) (the remaining four isolates showed recombinations in the K-locus and, therefore, were not assigned to a particular K-type). The only isolate of subclade A2 carried KL102. All isolates of subclade B1 carried KL2KL30 (K2 type), while the isolates of subclade B2 showed the KL39 type. While the clade C-genomes showed the KL64 (K14.K64 type) type, the D1-, D2-, and D3-subclade isolates demonstrated the KL3 (K3 type), KL36, and KL24 (K24 type) types, respectively (Fig. 1A).

Based on these results, we explored the hypothesis that, similar to other *K. pneumoniae* MDR clonal lineages [17, 83], ST395 diverged into distinct clades and subclades through chromosomal recombination, which eventually led to the observed capsule locus diversity. Analysis of the entire genome collection by Gubbins revealed a total of 678 putative recombination events, twelve (1.8%) of which involved large (> 100 kbp) regions. Most recombination events occurred in chromosomal regions but did not affect any of the MLST loci; however, the largest recombination events spanned a wide (1.1 Mbp) area of the chromosome which included capsule (K) and LPS antigen (O) biosynthesis loci (Fig. 2). In this area, the core-SNP profiles as well as core-genome MLST (cgMLST) profiles showed considerable diversity among different subclades (Additional file 1: Table S2). We therefore used two different approaches to identify putative recombination donors: (i) a core-SNP-based phylogenetic approach, which involved analysis of *K. pneumoniae* collections of genomes retrieved from the Patric database, and (ii) a query by cgMLST locus combination for the putative recombination regions in the

BIGSdb database. While we did not identify donors for all ST395 clades, we found that the ST147 *Klebsiella* lineage was involved in one or more recombination events, which resulted in the acquisition of a large fragment (> 661 kbp) by ST395 clade C within the genomic area mentioned above. We also identified fragments of 101 to 388 kbp spanning the same area that were likely acquired from donors of other *Klebsiella* sequence types: ST377 in subclade A2, ST896 in subclade D1, ST437 and ST11 in subclade D2, and ST15 and ST11 in subclade D3 (Fig. 2, Additional file 4). As expected, the regions corresponding to K and O loci (highlighted in red vertical lines in Fig. 2) showed peaks in recombination counts. Another major peak defining a recombination hot-spot within this area corresponded to the integrative conjugative element ICEKp (highlighted in orange vertical lines in Fig. 2), which encodes the siderophore yersiniabactin and its receptor (*ybt* locus). We detected different ICEKp elements that represent distinct *ybt* lineages in 223 of 297 (75.1%) ST395 isolates. ICEKp12 (*ybt16*) was the predominant type, detected in 182 of 189 (96.3%) genomes of subclade B1 as well as all 14 isolates of subclade B2 and eleven isolates of clade C. Other ICEKp variants were less common: ICEKp11 (*ybt15*) was found in three of 76 (3.9%) isolates of subclade A1 and both genomes of subclade D1; ICEKp5 (*ybt14*) in six of 76 (7.9%) isolates of subclade A1; ICEKp2 (*ybt13*) in both isolates of subclade D3; ICEKp4 (*ybt10*), ICEKp3 (*ybt9*), and the unclassified ICEKp variant in one isolate each of subclades A1, B1, and B1, respectively. Sixty-six of 76 (86.8%) isolates of subclade A1, the only isolate of subclade A2, 5/189 (2.6%) isolates of subclade B1, and both isolates of subclade D2 lacked ICEKp. Finally, we identified eight regions in the chromosome that showed peaks in recombination frequency associated with prophage-related sequences. Five of these regions (highlighted in green vertical lines in Fig. 2) were called by PHASTER as a questionable complete phage (Φ 1: similar to *Salmonella* phage SEN34 [NC_028699]) or intact phages (phages Φ 1a, Φ 2, Φ 3, and Φ 4: similar, respectively, to SEN34, *Klebsiella* phage ST512-KPC3phi13.2 [NC_049452], *Edwardsiella* phage GF-2 [NC_026611], and *Klebsiella* phage 3LV2017 [NC_047817]).

ST395 isolates often carried carbapenemase and ESBL genes simultaneously

Characterization of acquired resistance genes, especially carbapenemase, extended-spectrum β -lactamase (ESBL), and molecular class C cephalosporinase (AmpC) genes, was one of the other interests of this study (Table 1, Fig. 3A, B). The genes of the OXA-48 family carbapenemases (*bla*_{OXA-48-like}) were the most common, found in 163 of 297 (54.9%) of all isolates. The dominant variant,

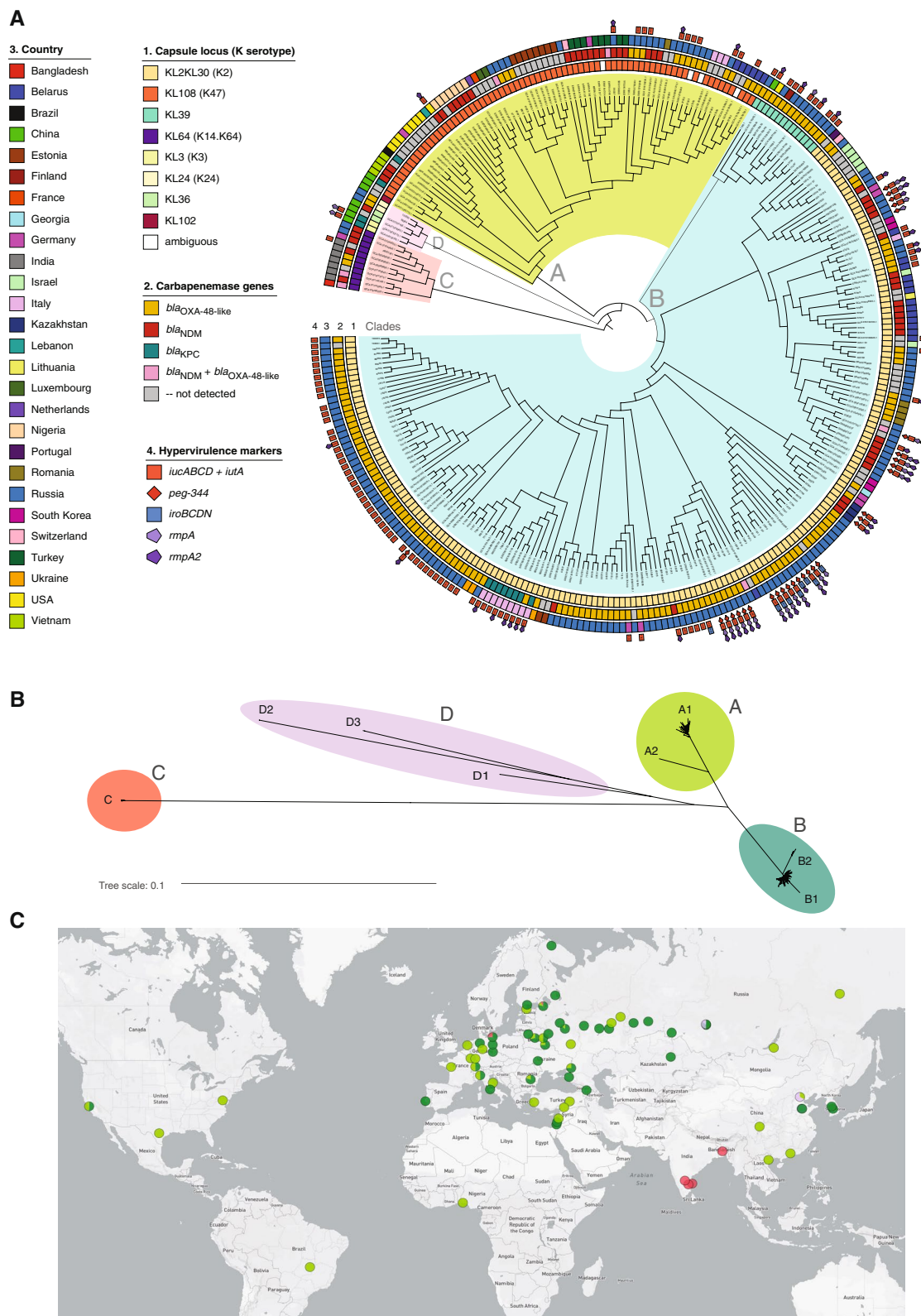


Fig. 1 Core-genome SNP phylogeny and geographic distribution of ST395. Phylogenetic tree of international ST395 isolates inferred from core-genome SNP analysis by Parsnp and visualized with iTOL. **A** Midpoint-rooted tree with unscaled branch lengths shows the classification of isolates into four clades and their characteristics according to the color legend. **B** Unrooted tree with scaled branch lengths shows the evolutionary distance between clades and subclades. **C** The map shows the geographic distribution of clades

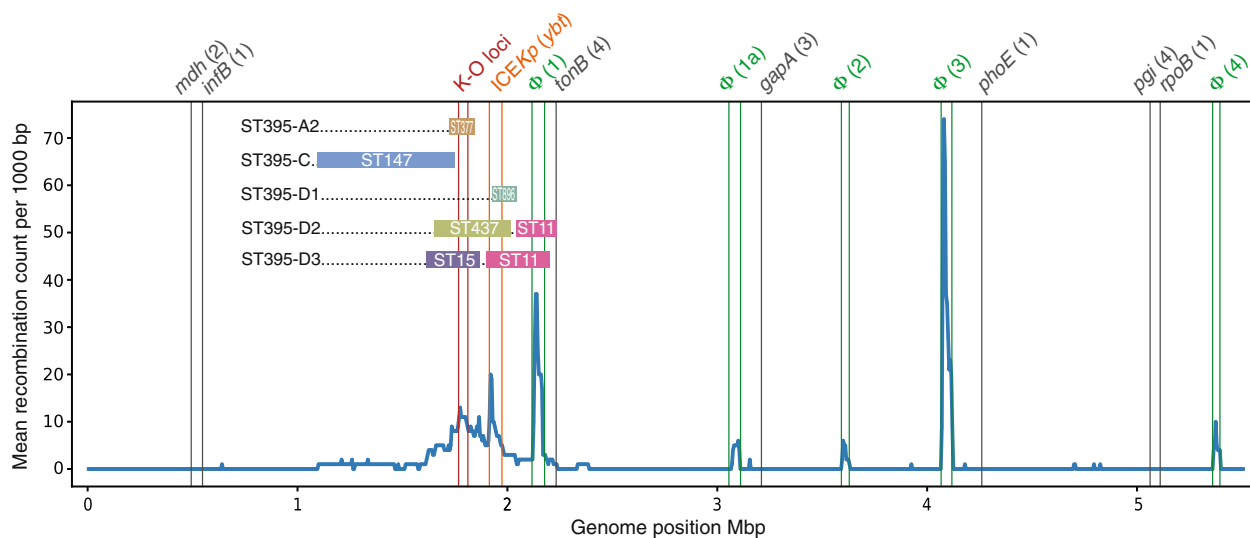


Fig. 2 Mean recombination counts across ST395 chromosomes and putative large recombination events. The positions of seven MLST loci are indicated by gray vertical lines with corresponding allele names and numbers in brackets shown on top. K and O loci boundaries are indicated by red vertical lines. The ICEKp element and prophage integration regions are bounded, respectively, by orange and green vertical lines. Prophages are denoted by the “Φ” symbol followed by a number in brackets: 1 and 1a, similar to *Salmonella* phage SEN34 (NC_028699); 2, similar to *Klebsiella* phage ST512-KPC3phi13.2 (NC_049452); 3, similar to *Edwardsiella* phage GF-2 (NC_026611); 4, similar to *Klebsiella* phage 3LV2017 (NC_047817). Large subclade-specific recombination regions are shown by colored horizontal bars indicating the putative donor lineages

*bla*_{OXA-48} was identified in the majority (133/167; 79.6%) of the Russian isolates but also in isolates from other European countries. The less common variants, *bla*_{OXA-232} and *bla*_{OXA-244} encoding enzymes with weaker carbapenem-hydrolyzing activity were identified, respectively, in the genomes of four isolates from India and Bangladesh, and two isolates from Russia. The genes of the NDM family carbapenemases (*bla*_{NDM}) were the second most prevalent (61/297; 20.5%) in isolates of all ST395 clades from various world regions. These included the most typical *bla*_{NDM-1} gene (52/297; 17.5%) detected in Belarus, Germany, Israel, Kazakhstan, Nigeria, Russia, Switzerland, Turkey, USA, and Vietnam, *bla*_{NDM-5} (8/297; 2.7%) in Bangladesh, China, Germany, India, and Russia, and *bla*_{NDM-16} in one (0.3%) Russian isolate. The genes for KPC type carbapenemases (*bla*_{KPC}) were the least common: *bla*_{KPC-2} (4/297; 1.3%) was found in the isolates of different ST395 clusters from Brazil and China, and *bla*_{KPC-3} (8/297, 2.7%) in closely related isolates from Italy. Six isolates (2%) from Bangladesh, India, Russia, and Turkey were found to carry a combination of *bla*_{NDM-1} and one of the *bla*_{OXA-48-like} genes. Carbapenemase-negative ST395 isolates occurred in many countries (Fig. 3A).

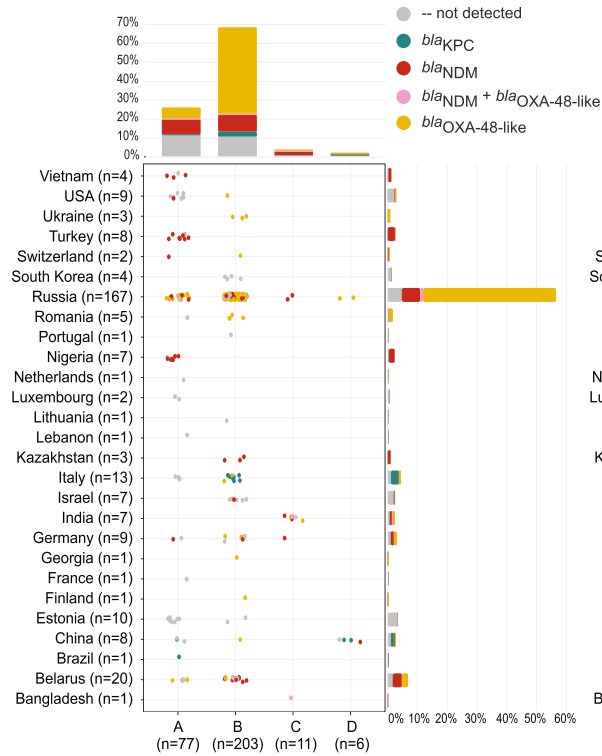
ESBL genes were almost universally present among global ST395 isolates (Fig. 3B). The majority (259/297; 87.2%) possessed the CTX-M-gene family, of which *bla*_{CTX-M-15} was the most common type (247/297; 83.2%). Other ESBL genes were detected sporadically: *bla*_{CTX-M-27} (8/297; 2.7%), *bla*_{CTX-M-3} (4/297; 1.3%), *bla*_{SHV-31}

(2/297; 0.7%), *bla*_{SHV-2}, *bla*_{VEB-1}, and *bla*_{CTX-M-55} (1/297; 0.3% each). In total 68.0% (202/297) of the isolates carried carbapenemase and ESBL resistance genes simultaneously. Six (2%) isolates representing different countries and ST395 clades carried the genes for acquired AmpC cephalosporinases: *bla*_{DHA-1} (4/297; 1.3%), *bla*_{CMY-2} and *bla*_{CMY-6} (1/297; 0.3% each), either alone or in combination with *bla*_{CTX-M} (Fig. 3B).

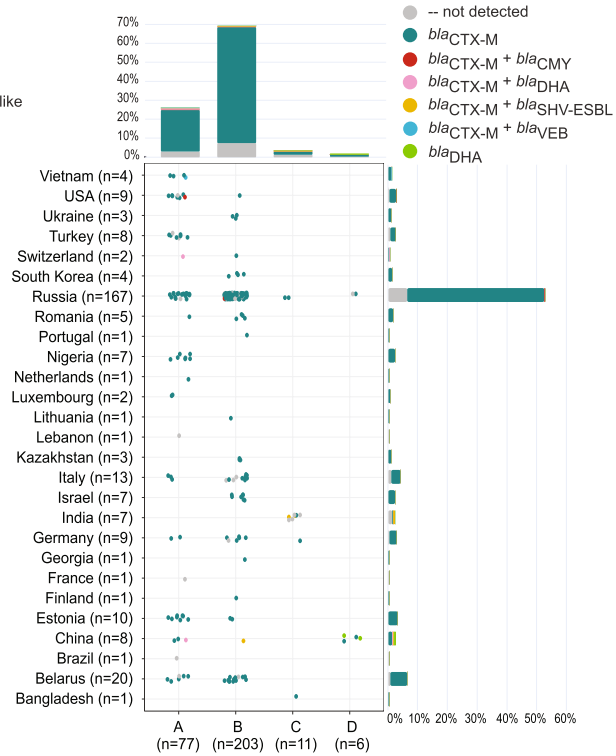
ST395 carried an unusually high abundance and diversity of 16S rRNA methyltransferase genes

Besides β-lactamases, we identified a great variety of resistance genes of plasmid and/or chromosomal origin to all antibiotic classes in the ST395 genomes (Table 1, Additional file 1: Table S1). Some of these genes, such as the dual-targeting aminoglycoside- and fluoroquinolone-modifying acetyltransferase genes (*aac*(6′)-Ib-cr), were highly common (232/297; 78.1%), while others, such as the mobile colistin resistance gene (*mcr*-8.2), were found in single isolates only. Besides *aac*(6′)-Ib-cr, ST395 isolates carried many other quinolone resistance determinants, including genes of the quinolone family protective proteins (*qnrB*, *qnrD1*, and *qnrS1*), resistance-nodulation-cell division (RND) and major facilitator superfamily (MFS) efflux pumps (*oqxA*, *oqxB*, and *qepA*). All, except for ten isolates, also harbored a variety of genes of aminoglycoside-modifying enzymes in various combinations. Most notably, however, we found an unusually high abundance (113/297; 38.1%) and diversity of 16S rRNA

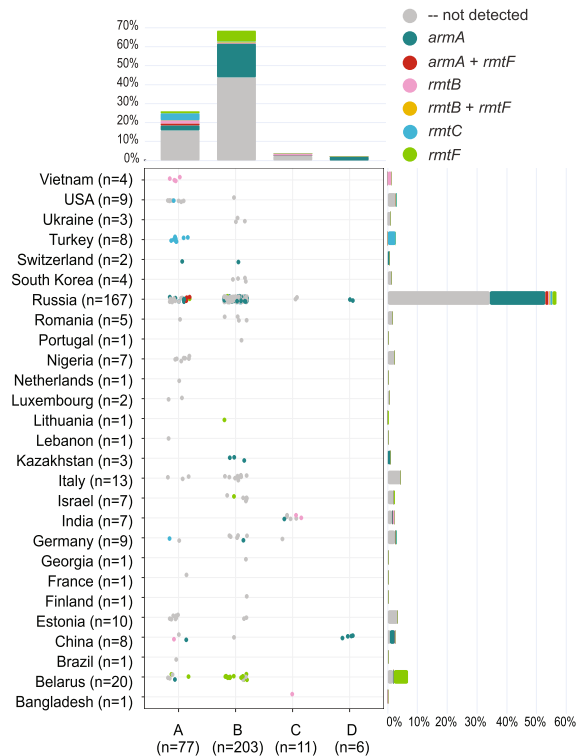
A. Carbapenemase genes



B. ESBL and AmpC genes



C. 16S rRNA methyltransferase genes



D. Hypervirulence marker genes

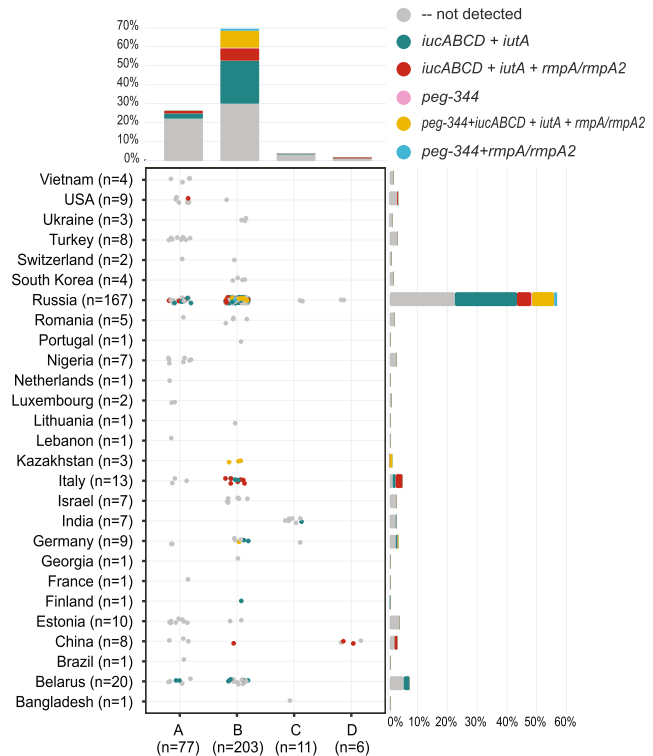


Fig. 3 Distribution of major antibiotic resistance and hypervirulence markers by ST395 clades and countries. Data shown summarizes distribution of carbapenemase genes (A), ESBL and AmpC genes (B), 16S rRNA methyltransferase genes (C), and hypervirulence marker genes (D) by country and clade. Each circle represents a genome colored based on the marker gene (combination). Barplots summarize the percentage of genomes from each clade (top) and country (right) and are colored based on the marker gene (combination)

methyltransferase genes, which confer resistance to all aminoglycosides (Fig. 3C). *ArmA* was the most prevalent (71/297; 23.9%) type of methyltransferase detected. Isolates carrying *armA* genes represented all ST395 subclades and seven countries of origin. Other types of methyltransferase genes, in order of decreasing prevalence, *rmtF* (24/297; 8.1%), *rmtC* (12/297; 4%), and *rmtB* (10/297; 3.4%), were more specifically associated with certain countries and geographic regions. For instance, *rmtF* and *rmtC* were most common among the isolates from Belarus and Turkey, respectively, but occurred only sporadically in the isolates from other countries. The genomes of four (1.3%) Russian isolates revealed the simultaneous presence of two methyltransferase genes (*rmtF* and either of *armA* or *rmtB*).

Hypervirulence markers were mostly found in ST395 clade B genomes

Next, we explored the distribution of hypervirulence (hv) genes in ST395. While the studied genomes contained many different biomarkers that may contribute to virulence in *K. pneumoniae* (Table 1, Additional file 1: Table S1), we assessed more specifically the plasmid virulence markers that have been identified as most important for differentiating the hvKp from non-hvKp representatives, namely, the aerobactin (*iucABCD-iutA*) and salmochelin (*iroBCDN*) iron uptake systems, regulators of mucoid phenotypes (*rmpA* and *rmpA2*), and the metabolite transporter gene (*peg-344*) [6]. These biomarkers were detected in 130 of 297 (43.8%) of the isolates but were not in complete linkage disequilibrium with each other (Fig. 3D). Aerobactin synthesis and receptor genes were detected in 126 of 297 (42.4%) isolates, either alone (74/297; 24.9%) or in combination with other markers. *rmpA* and *rmpA2* in 51 of 297 (17.2%), *peg-344* in 26/297 (8.8%), and the salmochelin gene cluster in six of 297 (2%) only. The distribution of these biomarkers among different ST395 clades and countries was rather uneven. Most isolates carrying hv genes belonged to clade B (115/130, 88.5%) and originated from Russia (102/130; 78.5%), Italy (10/130; 7.7%), and Belarus (6/130; 4.6%) (Fig. 3D).

Carbapenemase genes were associated with different plasmid replicon types and occasionally located on “hybrid” plasmids

Using MOB-suite, we were able to partially reconstruct and characterize the plasmids carrying carbapenemase genes

for 202 draft genome assemblies of overall 230 carbapenemase-positive isolates (87.8%) (Additional file 1: Table S3). Carbapenemase genes were tentatively associated with plasmids of 14 different replicon types with the highest diversity observed for *bla*_{NDM}-carrying plasmids (nine types) and lower diversity for *bla*_{OXA-48-like}- and *bla*_{KPC}-harboring plasmids—five and three types, respectively. *bla*_{OXA-48-like} genes were found predominantly on IncL (80/145; 55.2%) and IncM2 plasmids (56/145; 38.6%), *bla*_{NDM} genes were mostly located on multi-replicon plasmids of IncFIB/IncFII (22/53; 41.5%) and IncFIB/IncHI1B types (16/53; 30.2%), and *bla*_{KPC-3} genes were always associated with pKPC_Kp02-like plasmids. While we did not attempt to systematically assemble the plasmids carrying ESBL genes, we found different replicon types associated with ESBL and AmpCs (data not shown).

Similarly, we determined the genetic backgrounds of aerobactin cluster genes for 114/126 isolates (90.5%). In all but two cases, we identified the presence of *iucABCD-iutA* on IncFIB/IncHI1B dual-replicon plasmids. Notably, 96 of 112 (85.7%) of these typical hv plasmids were found to co-harbor various antibiotic resistance determinants, including carbapenemase genes (as exemplified in Fig. 4B) and, therefore, were classified as “hybrid” or “mosaic” plasmids. A triple-replicon IncFIB/IncHI1B/IncR plasmid, which carried *iucABCD-iutA* and *bla*_{NDM-1} genes, was detected in one isolate, and an IncFIA/IncFII plasmid, which carried *iucABCD-iutA*, *rmpA*, *rmpA2*, and *peg-344*, was detected in the remaining one.

Using a combination of Illumina short-read and Oxford Nanopore long-read sequencing, we obtained the complete genomes of two Russian isolates that belonged to the largest subclade B1. One of them (87569) carried *bla*_{OXA-48} and the other (105623) had a combination of *bla*_{OXA-48} and *bla*_{NDM-1}. Both isolates also possessed additional acquired resistance genes including *bla*_{CTX-M-15} and *armA*, as well as plasmid-encoded virulence factors including the aerobactin cluster (*iucABCD-iutA*). The *bla*_{OXA-48} gene was found as part of a Tn1999 insertion in the chromosome of isolate 87569, and, in isolate 105623, it was located on a typical IncL pOXA48-like plasmid (Fig. 4A). BLAST analysis showed that the 63,577-bp OXA-48-encoding plasmid of 105623 (pOXA48_105623) was almost identical (99% query coverage, 99.85% identity) to the recently reported 63,589-bp plasmid of a Russian ST23 *K. pneumoniae* isolate (GenBank accession no. CP063280; [84]) and was very similar to other related plasmids reported from different

(See figure on next page.)

Fig. 4 Structure of carbapenemase gene-carrying plasmids. BRIG images show similarities between plasmid sequences of selected ST395 isolates (central rings) and other published plasmid sequences identified by BLAST (concentric rings colored according to legends). Key plasmid genes are annotated by colored arrows and description: maintenance and transfer genes (gray), AMR genes (red), IS elements (orange), virulence genes (blue), and hypervirulence marker genes (deep blue). **A** IncL pOXA48-like plasmid of isolate 15623 and 87569, **B** IncFIB/IncHI1B hybrid virulence and resistance plasmids of isolates 105623 (central ring) and 87569 (outer ring)

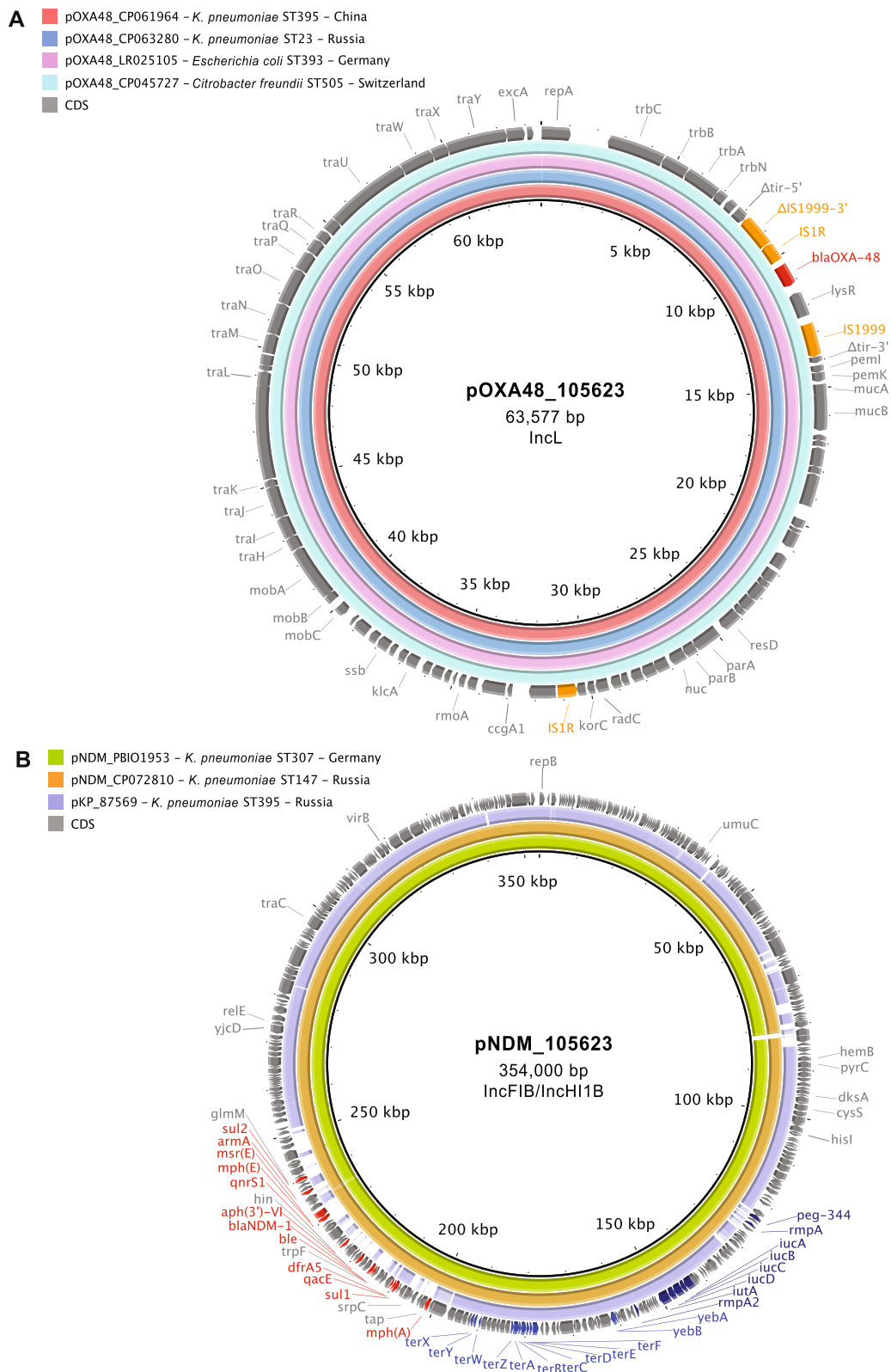


Fig. 4 (See legend on previous page.)

countries (Fig. 4). On these plasmids, the *bla*_{OXA-48} gene was carried by inverted Tn1999.2, which contains an insertion of IS1R and provides a stronger promoter for the gene expression compared to Tn1999.1 [85, 86]. Both isolates, 87569 and 105623, also harbored large hybrid (resistance and virulence) plasmids of the IncFIB/IncHI1B type, designated pKP_87569 (307,622 bp) and pNDM_105623 (354,000 bp), respectively (Fig. 4B). These plasmids shared a common backbone/core region and important hyper-virulence biomarker genes including the aerobactin cluster (*iucABCD-iutA*), activator of capsular polysaccharide synthesis (*rmpA2*), and tellurite resistance (*terABCDEFZ*) genes (pNDM_105623 additionally carried *rmpA* and the *peg-344* transporter genes) but differed in the content and arrangement of antibiotic resistance genes (Fig. 4B). pKP_87569 contained only four genes for chloramphenicol, aminoglycoside, and sulfonamide resistance (*catA1*, *ant(3'')-Ia*, *ant(2'')-Ia*, and *sul1*), while pNDM_105623 contained a large array of resistance genes to many antibiotic classes (*mph(A)*, *msr(E)*, *mph(E)*, *sul1*, *sul2*, *qnrS1*, *aph(3'')-VI*, *armA*, *dfra5*) and, most notably, the *bla*_{NDM-1} carbapenemase gene. BLAST analysis revealed the highest similarity of pNDM_105623 to the previously reported plasmids of Russian ST147 *K. pneumoniae* isolate (GenBank accession no. NZ_CP072810.1; [87]) and a German ST307 *K. pneumoniae* isolate (PBIO1953; [8]) (Fig. 4B). Finally, hybrid genome assemblies of the isolates 87569 and 105623 revealed the presence of IncR plasmids that carried an ESBL gene (*bla*_{CTX-M-15}) and additional resistance genes to penicillins and penicillin-inhibitor combinations (*bla*_{TEM-1} and *bla*_{OXA-1}), aminoglycosides, and fluoroquinolones (*aac(6'')-Ib-cr* and *qnrS1*), tetracycline (*tet(A)*), chloramphenicol (*catA1* and *catB3*), sulfonamides (*sul1*), and trimethoprim (*dfra1*).

Discussion

Here we describe the characteristics of *K. pneumoniae* ST395, which is an emerging high-risk clonal lineage. The investigated collection comprised 297 genomes and was geographically widely dispersed over 27 countries and four continents. More than 40% of the ST395 genomes were newly obtained as part of this study and represent Belarus, Kazakhstan, Germany, the USA, and, predominantly, Russia with the latter potentially leading to a bias in terms of interpreting geographical distributions and ST395 characteristics. Also, the fact that most studies and surveillance programs frequently use whole genome sequencing for outbreak investigations caused by MDR isolates obtained from humans might lead to underrepresentation of antibiotic-susceptible strains, which could have an additional impact on the diversity of our sample set. This limitation also applies to the interpretation of the emergence of convergent pathotypes that

combine resistance with virulence. Despite these biases, our data clearly point towards high-level endemicity of *K. pneumoniae* ST395 in hospital settings across Russia and extend upon previous epidemiological reports from other groups in Russia [40, 41, 88]. In addition, our results suggest the occurrence of local outbreaks and multiple sporadic cases of ST395 infections in other countries across the world.

Expanding on the study by Wyres *et al.* [17], which revealed the descent of ST395 lineage from ST11 (CC258) through large recombination events, we explored the role of chromosomal recombination in subsequent evolutionary processes and the intra-clade diversification of ST395. We found that the early ST395 isolates addressed by Wyres *et al.* belonged to the “original” phylogenetic clade (designated clade A in this study), which continuously spreads in multiple countries but is now being superseded by the potentially more successful clade B, at least in Europe and Asia. Overall, we assigned four distinct phylogenetic clades and eight subclades within ST395, which are well supported by hierarchical Bayesian clustering analysis of the core genome. We showed that the ST395 subclades evolved through recombination and insertion of large DNA fragments within a 1.1-Mbp chromosomal area that showed evidence of genetic exchange with unrelated clonal lineages, including *K. pneumoniae* ST377, ST147, ST437, ST15, ST896, and other yet unidentified donors. Consistent with previous reports [17, 89], our data therefore indicate that intra-clonal diversification by large-scale chromosome recombination is a common evolutionary strategy employed by typical clones of CC258 as well as ST395. Notably, the chromosomal area which showed extensive recombination in ST395 includes capsule and LPS biosynthesis loci, which likely explains the observed K and O antigen diversity between the subclades of ST395. Similar findings have been recently reported for various *Klebsiella* clonal lineages [83]. We also found that additional pan-genome diversity of ST395 is driven by the integrative conjugative element ICEKp, which encodes the major virulence factor yersiniabactin [74]. The non-universal presence of ICEKp and identification of different ICEKp types in isolates of different ST395 subclades indicates their likely independent and frequent chromosomal acquisition and loss. Finally, we identified several prophage-related regions variably present in genomes of ST395 isolates, a trait apparently common in CC258 and other MDR clonal lineages [90].

Our results revealed endemic tendencies regarding the occurrence of certain carbapenemases. In general, the distribution of different carbapenemase genes among ST395 isolates from different countries was rather uneven and reflects the geographical endemicity patterns of

certain carbapenemases. The most frequent carbapenemase genes were of the OXA-48-type and mostly distributed in Russia and Europe but also found in single isolates from the USA and China. In the retrospective surveillance of OXA-48 producing *Enterobacterales* collected between 2001 and 2011 in Europe and North Africa, *K. pneumoniae* ST395 was identified as the second most prevalent clone [91] whereas, in the very recent report from the SENTRY global surveillance, it was identified as the most prevalent genotype of OXA-48-producing *K. pneumoniae*, with the majority of isolates originating from Russia, Belarus, and Turkey [92].

Most genome sequences analyzed in this study were obtained using Illumina short-read sequencing. However, determining the plasmid content and identifying the plasmids carrying AMR and virulence genes from short-read WGS data is known to be challenging due to incomplete assembly of plasmids. The genes of interest cannot be directly assigned to a particular plasmid if found in different contigs than the plasmid marker genes. Assembly of AMR gene-carrying plasmids is especially problematic because they are usually large in size, low in copy number and frequently built as mosaic-like structures, containing mobile components that can be found in different plasmid backbones [93]. Thus, several software tools have been developed that use graph-based or reference-based approaches or their combination. Here, we used MOB-suite [80, 94] for reconstructing and characterizing plasmids carrying carbapenemase genes. MOB-suite allows “aggregation” of input sequence contigs into putative plasmids according to the reference hits obtained from the reference database of clustered complete plasmid sequences and determination of plasmid types using marker sequence databases of known replicon and relaxase genes. This software has been successfully used to characterize AMR gene-carrying plasmids in many large-scale bacterial population genetics studies, and in a recent study comparing the performance of different available plasmid prediction tools, has been shown to be the only tool to correctly reconstruct the majority of *Escherichia coli* plasmids, including large ESBL-encoding plasmids [93]. All available bioinformatics approaches, however, have important limitations and none can correctly reconstruct all plasmids from draft genome assemblies. Therefore, the data on association of carbapenemase and virulence genes with certain plasmid types presented in our study must be interpreted with caution. Also, we did not establish plasmid support for carbapenemase and aerobactin cluster genes in about 12% and 10% of *K. pneumoniae* ST395 genome sequences.

Consistent with many previous reports, in our study, *bla*_{OXA-48-like} genes were predominantly found on IncL and IncM2 plasmids, which are known as primary

vehicles for *bla*_{OXA-48} dissemination worldwide [95]. *bla*_{NDM} genes were often harbored by multi-replicon IncFIB/IncFII plasmids and hybrid (virulence and resistance) IncFIB/IncHI1B plasmids [87, 96], and *bla*_{KPC-3} genes were harbored by pKPC_Kp02-like plasmids, which has also been previously described [97].

To further characterize the typical carbapenemase-encoding plasmids, we obtained the complete genome sequences of two Russian isolates from the largest subclade B1. The OXA-48-encoding plasmid of one of those isolates was almost identical to the IncL plasmids reported previously in different strains of *Enterobacterales*: *K. pneumoniae* ST23 from Russia, *K. pneumoniae* ST395 from China, *E. coli* ST393 from Germany and *Citrobacter freundii* ST505 from Switzerland (Fig. 4A). Literature suggests that the Chinese isolate originated from a human and was collected in 2015, whereas the Swiss one was obtained from the grass surroundings of a veterinary clinic in 2019. For the German isolate, neither the collection date nor the origin location was available. However, the German sequence has been previously described in several publications [84, 98, 99]. In a recent report on a *K. pneumoniae* ST23 outbreak in Göttingen, Germany, in 2013, it was shown, that at least one outbreak isolate seemingly carried an identical plasmid [98]. Note that the occurrence of many carbapenemase-negative genomes in multiple countries likely indicates that production of these enzymes is not the only driver of successful dissemination of ST395, however. ESBL genes, mostly of CTX-M-type, were frequently present (87%) and other types occurred less often (<3%). The simultaneous occurrence of ESBL and carbapenemase genes is not unusual and consistent with previous findings. In addition to these, partly last-resort drug resistances, the presence of 16S rRNA methyltransferases intensifies the overall tense situation, where life-threatening infections with MDR *Klebsiella* cannot be reliably treated [16]. Therefore, the presence of 16S rRNA methyltransferase genes in >38% of the investigated isolates is alarming, especially in Russia, where ~40% of the isolates were positive for 16S rRNA methyltransferase genes. These rates are high compared to revealed rates from Spain (5%) and the UK (~16–18%), and similar to those in Asian countries (50%) [16]. We only found one genome with mobile colistin resistance (*mcr*) gene, which seemingly mirrors the current global situation of mostly colistin-susceptible ST395 isolates. However, we cannot exclude phenotypic colistin resistance conveyed through chromosomal adaptations. Given that colistin is still frequently used in veterinary medicine, it is not surprising that the *mcr*-carrying isolate in our study was obtained from a chicken.

ST395 was initially reported as non-hypervirulent type associated with multiple antibiotic resistances in nosocomial isolates. However, considering recent findings [41, 88, 100], we explored the acquired hv markers and their plasmid backbones. In total, we found 130 (43.8%) isolates carrying at least one hv marker gene. We found 96 isolates with plasmids classified as “hybrid” or “mosaic” plasmids, co-harboring hv and various AMR genes. As our sample set is biased towards MDR *K. pneumoniae*, however, these convergent genotypes might be overrepresented in this study, possibly not accurately reflecting the prevalence in larger datasets that contain a higher number of susceptible isolates.

We also confirmed that, in ST395, aerobactin was more common than other hv biomarkers, as it has been shown for other *K. pneumoniae* clonal lineages [101]. Interestingly, most of the isolates co-harboring virulence and resistance genes belonged to clade B, mostly of capsular type K2. This adds to the results of Lazareva from 2020 [41]. The convergence of ST395 hypervirulence and MDR has already been highlighted by some reports, especially in recent years [41, 88, 100]. Finding a similar hybrid (resistance and virulence) plasmid in three Russian isolates (two ST395 and one ST147) and one German (ST307) isolate indicates transmission of AMR-encoding plasmids among ST395 and other *Klebsiella* sequence types.

Conclusions

This study revealed the ability of ST395 to evolve and diversify into a highly successful, antibiotic-resistant, and virulent clonal lineage through both chromosomal recombination and acquisition of multiple extrachromosomal elements and provided evidence for its dissemination across countries and continents. Our results also suggest that controlling the further spread of this and other pathogenic bacterial lineages will remain challenging.

Abbreviations

AMR	Antimicrobial resistance
CC	Clonal complex
ESBL	Extended-spectrum beta-lactamase
hvKp	Hypervirulent <i>Klebsiella pneumoniae</i>
hv	Hypervirulent
IAC	Institute of Antimicrobial Chemotherapy
MDR	Multidrug-resistant/multidrug resistance
MFS	Major facilitator superfamily
MLST	Multilocus sequence typing
ORF	Open reading frame
RND	Resistance-nodulation-cell division
ST	Sequence type

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-023-01159-6>.

Additional file 1: Table S1. Characteristics of the isolates from the ST395 dataset. **Table S2.** Comparison of cgMLST profiles of isolates representing ST395 subclades. **Table S3.** Location of carbapenemase genes.

Additional file 2. Calculation of mean recombination counts. Formula for calculation of mean recombination counts.

Additional file 3. Detailed Methods. Overview of used bioinformatics tools and parameters.

Additional file 4: Identification of large recombination events and putative donors. **Table S4.** Large (>100 kbp) chromosome recombination events predicted by Gubbins and the putative donors identified by core-SNP phylogenetic analysis from the expanded genome dataset. **Fig. S1.** An example tree (A) and subtree (B) used to identify the ST437 as a putative donor for recombination event 13 in ST395 subclade D2.

Acknowledgements

We thank the entire MRSN staff for the collection and sequencing of a subset of US ST395 isolates. We also thank Dr. Thomas Riedel, Prof. Jacob Moran-Gilad, and Dr. Motoro for providing sequencing data and meta-information.

Authors' contributions

Providing genome sequence data: MSD, ADM, YAS, AAS, YVM, FL, RSK, GAS, VGA, AEZ, JAB, AMM, RB, DM. Study design: EAI, SG, KB, ML, KJH, KSc, MVE. Study analysis: ERS, MS, TR, VVS, KSy, SEH, MVE. Manuscript preparation: ERS, MS, TR, VVS, SEH, KSc, MVE. All authors read and approved the final version of the manuscript.

Authors' information

Elvira R. Shaidullina and Michael Schwabe (shared first authorship) as well as Katharina Schaufler and Mikhail V. Edelstein (shared senior authorship) contributed equally to this work.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work and the position of MS was supported by a grant from the Federal Ministry of Education and Research (BMBF) to KSc entitled “Disarming pathogens as a different strategy to fight antimicrobial-resistant Gram-negatives” (01KI2015). The funding sources had no influence on the design of the study and collection, analysis, and interpretation of data, and writing of the manuscript.

Availability of data and materials

All raw and assembled Illumina and Oxford Nanopore sequence data are available from the European Nucleotide Archive under the study accession nos. PRJEB49683 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB49683>) [46] and PRJEB53512 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB53512>) [49]. Assemblies of the US samples have been deposited with links to BioProject accession number PRJNA849970 [50] in the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA849970>). Individual accession numbers for raw reads and/or de novo assemblies are also available in the Additional file 1: Table S1. The entire ST395 genome dataset can be accessed as a Microreact project at <https://microreact.org/project/uyaJw eHqXXdVudHU3QDwW>.

Declarations

Ethics approval and consent to participate

This research conformed to the principles of the Helsinki Declaration. No ethics approval or consent to participate was required for the already published data.

The studied bacterial isolates from the National Antimicrobial Surveillance Program (PRJEB49683) were recovered at local clinical hospitals in Russia as part of routine diagnostic sampling. Informed consent was not sought for this study because it did not involve any additional interventions; no patient personal data were collected and the data about sources of isolates were collected anonymously. The National Antimicrobial Surveillance Program was approved by the Review Board of the Smolensk State Medical University. The MRSA isolates from the USA (PRJNA849970) were collected as part of public health surveillance activities, and informed consent was waived, as determined by the WRAIR branch director and Human Subjects Protection Branch (HSPB) which granted ethical approval.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Antimicrobial Chemotherapy, Smolensk State Medical University, Smolensk, Russia. ²Pharmaceutical Microbiology, Institute of Pharmacy, University of Greifswald, Greifswald, Germany. ³Federal State Budgetary Institution "Centre for Strategic Planning and Management of Biomedical Health Risks" of the Federal Medical Biological Agency, Moscow, Russia. ⁴Central Research Institute of Epidemiology, Moscow, Russia. ⁵Multidrug-Resistant Organism Repository and Surveillance Network, Walter Reed Army Institute of Research, Silver Spring, USA. ⁶Friedrich Loeffler-Institute of Medical Microbiology, University Medicine Greifswald, Greifswald, Germany. ⁷Institute of Medical Microbiology, University Hospital Münster, Münster, Germany. ⁸Institute of Biochemistry, University of Greifswald, Greifswald, Germany. ⁹Pharmaceutical Biology, Institute of Pharmacy, University of Greifswald, Greifswald, Germany. ¹⁰Institute of Medical Microbiology and Hospital Hygiene, Medical Faculty, Otto-Von-Guericke University Magdeburg, Magdeburg, Germany. ¹¹Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan, Russia. ¹²Department of Internal Medicine A, Shamir (Assaf Harofeh) Medical Center, Zerifin, Israel. ¹³Infection Control Unit, Shamir (Assaf Harofeh) Medical Center and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. ¹⁴Institute of Mathematics and Computer Science, University of Greifswald, Greifswald, Germany. ¹⁵Institute of Infection Medicine, Christian-Albrecht University Kiel and University Medical Center Schleswig-Holstein, Kiel, Germany.

Received: 23 June 2022 Accepted: 20 January 2023

Published online: 13 February 2023

References

- Wu K-M, Li L-H, Yan J-J, Tsao N, Liao T-L, Tsai H-C, et al. Genome sequencing and comparative analysis of Klebsiella pneumoniae NTUH-K2044, a strain causing liver abscess and meningitis. *J Bacteriol.* 2009;191:4492–501 (American Society for Microbiology).
- Pomakova DK, Hsiao C-B, Beanan JM, Olson R, MacDonald U, Keynan Y, et al. Clinical and phenotypic differences between classic and hypervirulent Klebsiella pneumoniae: an emerging and under-recognized pathogenic variant. *Eur J Clin Microbiol Infect Dis.* 2012;31:981–9.
- Lai Y-C, Lu M-C, Hsueh P-R. Hypervirulence and carbapenem resistance: two distinct evolutionary directions that led high-risk Klebsiella pneumoniae clones to epidemic success. *Expert Rev Mol Diagn.* 2019;19:825–37 (Taylor & Francis).
- Kochan TJ, Nozick SH, Medernach RL, Cheung BH, Gatesy SWM, Lebrun-Corbin M, et al. Genomic surveillance for multidrug-resistant or hypervirulent Klebsiella pneumoniae among United States bloodstream isolates. *BMC Infect Dis.* 2022;22:603.
- Lipworth S, Vihta K-D, Chau K, Barker L, George S, Kavanagh J, et al. Ten-year longitudinal molecular epidemiology study of *Escherichia coli* and *Klebsiella* species bloodstream infections in Oxfordshire. *UK Genome Medicine.* 2021;13:144.
- Russo TA, Olson R, Fang C-T, Stoesser N, Miller M, MacDonald U, et al. Identification of biomarkers for differentiation of hypervirulent *Klebsiella pneumoniae* from classical *K. pneumoniae*. *J Clin Microbiol.* 2018;56:e00776-18.
- Russo TA, Olson R, MacDonald U, Beanan J, Davidson BA. Aerobactin, but not Yersiniabactin, Salmochelin, or Enterobactin, enables the growth/survival of hypervirulent (hypermucoviscous) *Klebsiella pneumoniae* ex vivo and in vivo. *Infect Immun.* 2015;83:3325–33 (American Society for Microbiology).
- Heiden SE, Hübner N-O, Bohnert JA, Heidecke C-D, Kramer A, Balou V, et al. A *Klebsiella pneumoniae* ST307 outbreak clone from Germany demonstrates features of extensive drug resistance, hypermucoviscosity, and enhanced iron acquisition. *Genome Medicine.* 2020;12:113.
- Chen L, Todd R, Kiehlbauch J, Walters M, Kallen A. Notes from the field: pan-resistant New Delhi metallo-beta-lactamase-producing *Klebsiella pneumoniae* - Washoe County, Nevada, 2016. *MMWR Morb Mortal Wkly Rep.* 2017;66:33.
- Du F-L, Huang Q-S, Wei D-D, Mei Y-F, Long D, Liao W-J, et al. Prevalence of carbapenem-resistant *Klebsiella pneumoniae* co-harboring blaKPC-carrying plasmid and pLVPK-like virulence plasmid in bloodstream infections. *Front Cell Infect Microbiol.* 2020;10:556654.
- Lee MY, Ko KS, Kang C-I, Chung DR, Peck KR, Song J-H. High prevalence of CTX-M-15-producing *Klebsiella pneumoniae* isolates in Asian countries: diverse clones and clonal dissemination. *Int J Antimicrob Agents.* 2011;38:160–3.
- Bialek-Davenet S, Criscuolo A, Ailloud F, Passet V, Jones L, Delannoy-Vieillard A-S, et al. Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis.* 2014;20:1812–20.
- Peirano G, Chen L, Kreiswirth BN, Pitout JDD. Emerging antimicrobial-resistant high-risk *Klebsiella pneumoniae* clones ST307 and ST147. *Antimicrob Agents Chemother.* 2020;64:e01148-e1220.
- Rodrigues C, Desai S, Passet V, Gajjar D, Brisse S. Genomic evolution of the globally disseminated multidrug-resistant *Klebsiella pneumoniae* clonal group 147. *Microb Genom.* 2022;8:000737 (Microbiology Society).
- Karki D, Dhungel B, Bhandari S, Kunwar A, Joshi PR, Shrestha B, et al. Antibiotic resistance and detection of plasmid mediated colistin resistance mcr-1 gene among *Escherichia coli* and *Klebsiella pneumoniae* isolated from clinical samples. *Gut Pathogens.* 2021;13:45.
- Arca-Suárez J, Rodiño-Janeiro BK, Pérez A, Gujjarro-Sánchez P, Vázquez-Ucha JC, Cruz F, et al. Emergence of 16S rRNA methyltransferases among carbapenemase-producing Enterobacterales in Spain studied by whole-genome sequencing. *Int J Antimicrob Agents.* 2022;59:106456.
- Wyres KL, Gorrie C, Edwards DJ, Wertheim HFL, Hsu LY, Van Kinh N, et al. Extensive capsule locus variation and large-scale genomic recombination within the *Klebsiella pneumoniae* clonal group 258. *Genome Biol Evol.* 2015;7:1267–79.
- Loucif L, Kassah-Laouar A, Saidi M, Messala A, Chelaghma W, Rolain J-M. Outbreak of OXA-48-producing *Klebsiella pneumoniae* involving a sequence type 101 clone in Batna University Hospital. *Algeria Antimicrob Agents Chemother.* 2016;60:7494–7.
- Potron A, Kalpoe J, Poirel L, Nordmann P. European dissemination of a single OXA-48-producing *Klebsiella pneumoniae* clone. *Clin Microbiol Infect.* 2011;17:E24–6.
- Kovács K, Nyul A, Mestyán G, Melegh S, Fenyvesi H, Jakab G, et al. Emergence and interhospital spread of OXA-48-producing *Klebsiella pneumoniae* ST395 clone in Western Hungary. *Infect Dis.* 2017;49:231–3 (Taylor & Francis).
- Giani T, Conte V, Di Pilato V, Aschbacher R, Weber C, Larcher C, et al. *Escherichia coli* from Italy producing OXA-48 carbapenemase encoded by a novel Tn1999 transposon derivative. *Antimicrob Agents Chemother.* 2012;56:2211–3 (American Society for Microbiology).
- Dimou V, Dhanji H, Pike R, Livermore DM, Woodford N. Characterization of Enterobacteriaceae producing OXA-48-like carbapenemases in the UK. *J Antimicrob Chemother.* 2012;67:1660–5.
- Oteo J, Ortega A, Bartolomé R, Bou G, Conejo C, Fernández-Martínez M, et al. Prospective multicenter study of carbapenemase-producing enterobacteriaceae from 83 hospitals in Spain reveals high in vitro susceptibility to colistin and meropenem. *Antimicrob Agents Chemother.* 2015;59:3406–12 (American Society for Microbiology).
- Low Y-M, Yap PS-X, Abdul Jabbar K, Ponnampalavanar S, Karunakaran R, Velayuthan R, et al. The emergence of carbapenem resistant *Klebsiella pneumoniae* in Malaysia: correlation between microbiological trends

- with host characteristics and clinical factors. *Antimicrob Resist Infect Control*. 2017;6:5.
25. Poirel L, Abdelaziz MO, Bernabeu S, Nordmann P. Occurrence of OXA-48 and VIM-1 carbapenemase-producing Enterobacteriaceae in Egypt. *Int J Antimicrob Agents*. 2013;41:90–1.
 26. Cuzon G, Naas T, Bogaerts P, Glupczynski Y, Huang T-D, Nordmann P. Plasmid-encoded carbapenem-hydrolyzing β -lactamase OXA-48 in an imipenem-susceptible *Klebsiella pneumoniae* strain from Belgium. *Antimicrob Agents Chemother*. 2008;52:3463–4 (American Society for Microbiology).
 27. Hammerum AM, Larsen AR, Hansen F, Justesen US, Friis-Møller A, Lemming LE, et al. Patients transferred from Libya to Denmark carried OXA-48-producing *Klebsiella pneumoniae*, NDM-1-producing *Acinetobacter baumannii* and methicillin-resistant *Staphylococcus aureus*. *Int J Antimicrob Agents*. 2012;40:191–2.
 28. Kaase M, Schimanski S, Schiller R, Beyreiß B, Thürmer A, Steinmann J, et al. Multicentre investigation of carbapenemase-producing *Escherichia coli* and *Klebsiella pneumoniae* in German hospitals. *Int J Med Microbiol*. 2016;306:415–20.
 29. Morris D, O'Connor M, Izdebski R, Corcoran M, Ludden CE, McGrATH E, et al. Dissemination of clonally related multidrug-resistant *Klebsiella pneumoniae* in Ireland. *Epidemiol Infect Cambridge Univ Press*. 2016;144:443–8.
 30. De Laveleye M, Huang TD, Bogaerts P, Berhin C, Bauraing C, Sacré P, et al. Increasing incidence of carbapenemase-producing *Escherichia coli* and *Klebsiella pneumoniae* in Belgian hospitals. *Eur J Clin Microbiol Infect Dis*. 2017;36:139–46.
 31. Czobor I, Novais Á, Rodrigues C, Chifiriuc MC, Mihăescu G, Lazăr V, et al. Efficient transmission of IncFIY and IncL plasmids and *Klebsiella pneumoniae* ST101 clone producing OXA-48, NDM-1 or OXA-181 in Bucharest hospitals. *Int J Antimicrob Agents*. 2016;48:223–4.
 32. Löfmark S, Sjöström K, Mäkitalo B, Edquist P, Tegmark Wisell K, Giske CG. Carbapenemase-producing Enterobacteriaceae in Sweden 2007–2013: experiences from seven years of systematic surveillance and mandatory reporting. *Drug Resist Updates*. 2015;20:29–38.
 33. Adler A, Shklyar M, Schwaber MJ, Navon-Venezia S, Dhaher Y, Edgar R, et al. Introduction of OXA-48-producing Enterobacteriaceae to Israeli hospitals by medical tourism. *J Antimicrob Chemother*. 2011;66:2763–6.
 34. Skalova A, Chudejova K, Rotova V, Medvecký M, Studentova V, Chudackova E, et al. Molecular characterization of OXA-48-like-producing Enterobacteriaceae in the Czech Republic and evidence for horizontal transfer of pOXA-48-like plasmids. *Antimicrob Agents Chemother*. 2017;61:e01889-16 (American Society for Microbiology).
 35. Ageevets VA, Partina IV, Lisitsyna ES, Ilna EN, Lobzin YV, Shlyapnikov SA, et al. Emergence of carbapenemase-producing Gram-negative bacteria in Saint Petersburg, Russia. *Int J Antimicrob Agents*. 2014;44:152–5.
 36. Liapis E, Pantel A, Robert J, Nicolas-Chanoine M-H, Cavalieri L, van der Mee-Marquet N, et al. Molecular epidemiology of OXA-48-producing *Klebsiella pneumoniae* in France. *Clin Microbiol Infect*. 2014;20:O1121–3.
 37. Mairi A, Pantel A, Sotto A, Lavigne J-P, Touati A. OXA-48-like carbapenemases producing Enterobacteriaceae in different niches. *Eur J Clin Microbiol Infect Dis*. 2018;37:587–604.
 38. Yang J, Ye L, Guo L, Zhao Q, Chen R, Luo Y, et al. A nosocomial outbreak of KPC-2-producing *Klebsiella pneumoniae* in a Chinese hospital: dissemination of ST11 and emergence of ST37, ST392 and ST395. *Clin Microbiol Infect*. 2013;19:E509-515.
 39. Di Pilato V, Errico G, Monaco M, Giani T, Del Grosso M, Antonelli A, et al. The changing epidemiology of carbapenemase-producing *Klebsiella pneumoniae* in Italy: toward polyclonal evolution with emergence of high-risk lineages. *J Antimicrob Chemother*. 2021;76:355–61.
 40. Fursova NK, Astashkin EI, Gabrielyan NI, Novikova TS, Fedyukina GN, Kubanova MK, et al. Emergence of five genetic lines ST395NDM-1, ST13OXA-48, ST3346OXA-48, ST39CTX-M-14, and Novel ST3551OXA-48 of multidrug-resistant clinical *Klebsiella pneumoniae* in Russia. *Microb Drug Resist*. 2020;26:924–33.
 41. Lazareva I, Ageevets V, Sopova J, Lebedeva M, Starkova P, Likholotova D, et al. The emergence of hypervirulent blaNDM-1-positive *Klebsiella pneumoniae* sequence type 395 in an oncology hospital. *Infect Genet Evol*. 2020;85:104527.
 42. Kuzmenkov AY, Trushin IV, Vinogradova AG, Avramenko AA, Sukhorukova MV, Malhotra-Kumar S, et al. AMRmap: an interactive web platform for analysis of antimicrobial resistance surveillance data in Russia. *Front Microbiol*. 2021;12:620002.
 43. ISO 20776-1:2019 Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices — Part 1: Broth micro-dilution reference method for testing the in vitro activity of antimicrobial agents against rapidly growing aerobic bacteria involved in infectious diseases. ISO. (Cited 2022 Oct 27). Available from: <https://www.iso.org/standard/70464.html>
 44. European Committee on Antimicrobial Susceptibility Testing. EUCAST reading guide for broth microdilution. (Cited 2022 Oct 27). Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/2022_manuals/Reading_guide_BMD_v_4.0_2022.pdf
 45. European Committee on Antimicrobial Susceptibility Testing. EUCAST guidelines for detection of resistance mechanisms and specific resistances of clinical and/or epidemiological importance. (Cited 2022 Oct 27). Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Resistance_mechanisms/EUCAST_detection_of_resistance_mechanisms_170711.pdf
 46. Kozlov RS, Edelstein MV, Shaidullina Elvira R, Shapovalova VV, Shelonkov AA. Surveillance of carbapenemase-producing *Klebsiella pneumoniae* in Russia, Belarus and Kazakhstan. *Eur Nucleotide Arch*. 2022. Available from: <https://www.ebi.ac.uk/ena/browser/view/PRJEB49683>
 47. Bouganim R, Dykman L, Fakeh O, Motro Y, Oren R, Daniel C, et al. The clinical and molecular epidemiology of noncarbapenemase-producing carbapenem-resistant Enterobacteriaceae: a case-control matched analysis. *Open Forum Infect Dis*. 2020;7:ofaa299.
 48. Zautner AE, Bunk B, Pfeifer Y, Spröer C, Reichard U, Eiffert H, et al. Monitoring microevolution of OXA-48-producing *Klebsiella pneumoniae* ST147 in a hospital setting by SMRT sequencing. *J Antimicrob Chemother*. 2017;72:2737–44.
 49. Shaidullina ER, Schwabe M, et al. Genomic analysis of the international high-risk clonal lineage *Klebsiella pneumoniae* sequence type 395. *Eur Nucleotide Arch*. 2023. Available from: <https://www.ebi.ac.uk/ena/browser/view/PRJEB53512>
 50. Lebreton F, Ong AC, Kwak YI, Bennett JW, Mc Gann PT. Genomic surveillance of *K. pneumoniae* in the Military Health System. Available from: <https://www.ncbi.nlm.nih.gov/bioproject/849970>
 51. Seemann T. mlst. 2022 (Cited 2022 Oct 27). Available from: <https://github.com/tseemann/mlst>
 52. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010; Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
 53. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
 54. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20:257.
 55. Pribelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Curr Protoc Bioinformatics*. 2020;70: e102.
 56. Lanfear R, Schalamun M, Kainer D, Wang W, Schwessinger B. MinIONQC: fast and simple quality control for MinION sequencing data. *Bioinformatics*. 2019;35:523–5.
 57. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, et al. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol*. 2021;22:266.
 58. Medaka. Oxford Nanopore Technologies; 2022 (Cited 2022 Oct 27). Available from: <https://github.com/nanoporetech/medaka>
 59. Wick RR, Holt KE. Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol*. 2022;18:e1009802.
 60. Zimin AV, Salzberg SL. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol*. 2020;16:e1007981.
 61. Douglass AP, O'Brien CE, Offei B, Coughlan AY, Ortiz-Merino RA, Butler G, et al. Coverage-versus-length plots, a simple quality control step for de novo yeast genome sequence assemblies. *G3*. 2019;9:879–87.
 62. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol*. 2014;15:524.

63. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res.* 2019;47:5539–49.
64. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49:W293–6.
65. Seemann T. snippy: fast bacterial variant calling from NGS reads. 2022. Available from: <https://github.com/tseemann/snippy>
66. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 2015;43:e15.
67. Comandatore F, Sasser D, Bayliss SC, Scaltriti E, Gaiarsa S, Cao X, et al. Gene composition as a potential barrier to large recombinations in the bacterial pathogen *Klebsiella pneumoniae*. *Genome Biol Evol.* 2019;11:3240–51.
68. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, et al. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res.* 2020;48:D606–12.
69. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol.* 2009;26:1641–50.
70. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5:e9490 (Public Library of Science).
71. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30:2068–9.
72. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci Rep.* 2021;11:12728 (Nature Publishing Group).
73. Lam MMC, Wyres KL, Judd LM, Wick RR, Jenney A, Brisse S, et al. Tracking key virulence loci encoding aerobactin and salmochelin siderophore synthesis in *Klebsiella pneumoniae*. *Genome Med.* 2018;10:77.
74. Lam MMC, Wick RR, Wyres KL, Gorrie CL, Judd LM, Jenney AWJ, et al. Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in *Klebsiella pneumoniae* populations. *Microb Genom.* 2018;4(9):e000196.80.
75. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat Commun.* 2021;12:4188 (Nature Publishing Group).
76. Wick RR, Heinz E, Holt KE, Wyres KL. Kaptive Web: user-friendly capsule and lipopolysaccharide serotype prediction for *Klebsiella* genomes. *J Clin Microbiol.* 2018;56:e00197–18 (American Society for Microbiology).
77. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* 2018;3:124.
78. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 2006;34:D32–6.
79. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 2016;44:W16–21.
80. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom.* 2018;4(8):e000206.
81. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011;12:402.
82. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421.
83. Wyres KL, Wick RR, Judd LM, Froumine R, Tokolyi A, Gorrie CL, et al. Distinct evolutionary dynamics of horizontal gene transfer in drug resistant and virulent clones of *Klebsiella pneumoniae*. *PLOS Genetics.* 2019;15:e1008114 (Public Library of Science).
84. Shaidullina E, Shelenkov A, Yanushevich Y, Mikhaylova Y, Shagin D, Alexandrova I, et al. Antimicrobial resistance and genomic characterization of OXA-48- and CTX-M-15-Co-producing hypervirulent *Klebsiella pneumoniae* ST23 recovered from nosocomial outbreak. *Antibiotics.* 2020;9:862 (Multidisciplinary Digital Publishing Institute).
85. Beyrouthy R, Robin F, Delmas J, Gibold L, Dalmaso G, Dabboussi F, et al. IS1R-mediated plasticity of IncL/M plasmids leads to the insertion of blaOXA-48 into the *Escherichia coli* chromosome. *Antimicrob Agents Chemother.* 2014;58:3785–90 (American Society for Microbiology).
86. Carrèr A, Poirel L, Eraksoy H, Cagatay AA, Badur S, Nordmann P. Spread of OXA-48-positive carbapenem-resistant *Klebsiella pneumoniae* isolates in Istanbul. *Turkey Antimicrob Agents Chemother.* 2008;52:2950–4.
87. Starkova P, Lazareva I, Avdeeva A, Sulian O, Likholetova D, Ageevets V, et al. Emergence of hybrid resistance and virulence plasmids harboring New Delhi metallo- β -lactamase in *Klebsiella pneumoniae* in Russia. *Antibiotics.* 2021;10:691 (Multidisciplinary Digital Publishing Institute).
88. Fursova NK, Astashkin EI, Ershova ON, Aleksandrova IA, Savin IA, Novikova TS, et al. Multidrug-resistant *Klebsiella pneumoniae* causing severe infections in the neuro-ICU. *Antibiotics (Basel).* 2021;10:979.
89. Chen L, Mathema B, Pitout JDD, DeLeo FR, Kreiswirth BN. Epidemic *Klebsiella pneumoniae* ST258 is a hybrid strain. *mBio.* 2014;5:e01355–01314.
90. Shen J, Zhou J, Xu Y, Xiu Z. Prophages contribute to genome plasticity of *Klebsiella pneumoniae* and may involve the chromosomal integration of ARGs in CG258. *Genomics.* 2020;112:998–1010.
91. Potron A, Poirel L, Rondinaud E, Nordmann P. Intercontinental spread of OXA-48 beta-lactamase-producing Enterobacteriaceae over a 11-year period, 2001 to 2011. *Euro Surveill.* 2013;18:20549.
92. Castanheira M, Doyle TB, Collingsworth TD, Sader HS, Mendes RE. Increasing frequency of OXA-48-producing Enterobacteriales worldwide and activity of ceftazidime/avibactam, meropenem/vaborbactam and comparators against these isolates. *J Antimicrob Chemother.* 2021;76:3125–34.
93. Paganini JA, Plantinga NL, Arredondo-Alonso S, Willems RJL, Schürch AC. Recovering *Escherichia coli* plasmids in the absence of long-read sequencing data. *Microorganisms.* 2021;9:1613.
94. Robertson J, Bessonov K, Schonfeld J, Nash JHE. Universal whole-sequence-based plasmid typing and its utility to prediction of host range and epidemiological surveillance. *Microb Genom.* 2020;6(10):e000435.
95. Pitout JDD, Peirano G, Kock MM, Strydom K-A, Matsumura Y. The global ascendency of OXA-48-type carbapenemases. *Clin Microbiol Rev.* 2019;33:e00102–e119.
96. Dong H, Li Y, Cheng J, Xia Z, Liu W, Yan T, et al. Genomic epidemiology insights on NDM-producing pathogens revealed the pivotal role of plasmids on blaNDM transmission. *Microbiol Spectr.* 2022;10:e0215621.
97. Kim JO, Song SA, Yoon E-J, Shin JH, Lee H, Jeong SH, et al. Outbreak of KPC-2-producing Enterobacteriaceae caused by clonal dissemination of *Klebsiella pneumoniae* ST307 carrying an IncX3-type plasmid harboring a truncated Tn4401a. *Diagn Microbiol Infect Dis.* 2017;87:343–8.
98. Schwanbeck J, Bohne W, Hasdemir U, Groß U, Pfeifer Y, Bunk B, et al. Detection of a new resistance-mediating plasmid chimera in a blaOXA-48-positive *Klebsiella pneumoniae* strain at a German University Hospital. *Microorganisms.* 2021;9:720.
99. Cabanel N, Rosinski-Chupin I, Chiarelli A, Botin T, Tato M, Canton R, et al. Evolution of VIM-1-producing *Klebsiella pneumoniae* isolates from a hospital outbreak reveals the genetic bases of the loss of the urease-positive identification character. *mSystems.* 2021;6:e0024421.
100. Fasciana T, Gentile B, Aquilina M, Ciammaruconi A, Mascarella C, Anselmo A, et al. Co-existence of virulence factors and antibiotic resistance in new *Klebsiella pneumoniae* clones emerging in south of Italy. *BMC Infect Dis.* 2019;19:928.
101. Shankar C, Basu S, Lal B, Shanmugam S, Vasudevan K, Mathur P, et al. Aerobactin seems to be a promising marker compared with unstable RmpA2 for the identification of hypervirulent carbapenem-resistant *Klebsiella pneumoniae*: in silico and in vitro evidence. *Front Cell Infect Microbiol.* 2021;11: 709681.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.