scientific reports

OPEN

In-depth analysis of *Klebsiella aerogenes* resistome, virulome and plasmidome worldwide

Sergio Morgado[1]✉, Érica Fonseca[1], Fernanda Freitas[1], Raquel Caldart[2] & Ana Carolina Vicente[1]

Klebsiella aerogenes is an emergent pathogen associated with outbreaks of carbapenem-resistant strains. To date, studies focusing on *K. aerogenes* have been small-scale and/or geographically restricted. Here, we analyzed the epidemiology, resistome, virulome, and plasmidome of this species based on 561 genomes, spanning all continents. Furthermore, we sequenced four new strains from Brazil (mostly from the Amazon region). Dozens of STs occur worldwide, but the pandemic clones ST93 and ST4 have prevailed in several countries. Almost all genomes were clinical, however, most of them did not carry ESBL or carbapenemases, instead, they carried chromosomal alterations (*omp*36, *amp*D, *amp*G, *amp*R) associated with resistance to β-lactams. Integrons were also identified, presenting gene cassettes not yet reported in this species (*bla*IMP, *bla*VIM, *bla*GES). Considering the virulence loci, the yersiniabactin and colibactin operons were found in the ICEKp10 element, which is disseminated in genomes of several STs, as well as an incomplete salmochelin cluster. In contrast, the aerobactin hypervirulence trait was observed only in one ST432 genome. Plasmids were common, mainly from the ColRNAI replicon, with some carrying resistance genes (*mcr*, *bla*TEM, *bla*NDM, *bla*IMP, *bla*KPC, *bla*VIM) and virulence genes (EAST1, *sen*B). Interestingly, 172 genomes of different STs presented putative plasmids containing the colicin gene.

The bacteria formerly known as *Enterobacter aerogenes* has recently been renamed *Klebsiella aerogenes* based on whole-genome sequence phylogenetics[1,2]. This species is a ubiquitous member of the *Enterobacteriaceae* family, and although it has always been considered an opportunistic pathogen, recently, interest in this organism has increased, mainly due to the emergence of multidrug-resistant and carbapenem-resistant strains[3]. This may be due to the acquisition of antibiotic resistance genes carried by mobile elements, such as plasmids and conjugative integrative elements, which have already been identified in *K. aerogenes*[4–6]. However, in this species, the main mechanisms of resistance to carbapenems are attributed to chromosomal overexpression of AmpC β-lactamase (e.g., *bla*DHA-1 and *bla*CMY-2) and mutations that affect membrane permeability[4,7,8]. Furthermore, reports show that polymyxin resistance is occurring in strains of this species due to mutations in some housekeeping genes (*mgr*B, *pmr*A, *pmr*B, and *ept*A)[9,10] and, eventually, by plasmids carrying the *mcr* gene.

There are continuous reports of the high frequency of outbreaks of this emergent pathogen, some of them associated with high mortality rates[11,12], in different clinical settings around the world, such as neonatal[13–15], geriatric[16], and intensive care units[8,17]. Thus, *K. aerogenes* was included in the ESKAPE group, which encompasses important pathogens associated with antimicrobial resistance[5]. Among the known *K. aerogenes* lineages, comparative genomics showed that type 4 (ST4) and ST93 lineages likely represent the dominant lineages associated with human infections worldwide[4]. In addition, some lineages have been associated with virulence determinants, such as siderophores and toxins, which potentiate their persistence and favor the emergence of nosocomial outbreaks[4,5,10,14].

So far, studies focusing on *K. aerogenes* have been small-scale and/or with a restricted geographic perspective[4,5,18]. Here, in addition to generating new genomes of this species in Brazil and contextualizing them in the global scenario, we performed an in-depth analysis of the resistome, virulome, and plasmidome of this emergent pathogen.

[1]Laboratório de Genética Molecular de Microrganismos, Instituto Oswaldo Cruz, Rio de Janeiro, RJ 21040-360, Brazil. [2]Centro de Ciências da Saúde, Universidade Federal de Roraima, Boa Vista, RR 69300-000, Brazil. ✉email: sergio.morgado@ioc.fiocruz.br

Results

K. aerogenes sequence typing and epidemiology

In this study, we analyzed four strains of *K. aerogenes* from nosocomial cases from the Amazon and Southeast Brazil: Ka-01RR (subclavian vein catheter tip), Ka-02RR (tracheal secretion), Ka-04RR (unknown), Ka-06RJ (sputum). These strains were subjected to whole genome sequencing using the Illumina platform. The assembly statistics of these four genomes are presented in Table 1. The size of the genomes ranged from 5,043,146 to 5,175,129 bp with a GC content of ~ 55% and encoding 4728 to 4815 protein-coding sequences (CDS), 7–9 rRNAs, and 78–80 tRNAs (Table 1). Although the coverage depth of the genomes is not as high as commonly in Illumina runs, all genomes had size and number of CDS compatible with the species, as well as high genomic completeness (BUSCO > 95%).

In total, 561 *K aerogenes* were analyzed, where most (n = 477) were associated with human infections, and others were associated with the environment (n = 15) and animals (n = 13), covering the period from 1955 to 2022. These genomes were spread across 37 countries, with a bias towards the US, which corresponds to 44% (n = 247) of the genomes in the dataset. Only eight countries have more than a dozen genomes (Australia, Germany, Brazil, Canada, UK, China, Singapore, USA) (Table S1). For all sources, several STs were identified, where ST93 (n = 121) and ST4 (n = 57) were the most prevalent (Table S1). Other STs were present in no more than 20 genomes, showing that the prevalence of ST93 and ST4 is indeed much higher (Table S1). A maximum likelihood tree was constructed based on the core genes (n = 2319) of the 561 genomes (Fig. 1) and showed the division of the genomes into four main clusters, with the main one encompassing most of the genomes (493/561) into several subclusters, including ST4 and ST93 (Fig. 1). 41 genomes did not have any ST assigned, showing the genetic diversity of this species. Therefore, we submitted these allelic patterns to pubMLST database (https://pubmlst.org/bigsdb?db=pubmlst_kaerogenes_seqdef), where 41 new ST profiles were generated and included in this study (ST398-ST437, and ST440). The ST4 genomes are from 2004 to 2022 and occur in several countries in South America, North America, Asia, and, only in the United Kingdom in Europe. Most of this ST is associated with human sources; however, one genome is from the environment (Table S1). ST93 is more widespread than ST4 genomes as it also occurs in several European countries and Central America. Most ST93 genomes are from humans, occurring from 1997 to 2022 (Table S1).

In Brazil, 16 *K aerogenes* genomes were associated with human infections from 2006 to 2016 in three states in the Southeast region (Minas Gerais, Paraná and São Paulo), where three STs were observed: ST93 (11/20), ST16 (2 /20) and ST4 (3/20) (Table S1). Furthermore, the four new genomes sequenced here came from other states in the Southeast region, Rio de Janeiro (ST296), and from the North region, Roraima (ST93 and ST440). ST296 and ST440 represent two new STs in Brazil.

Brazilian *K. aerogenes* antibiotic susceptibility

The four Brazilian strains of *K. aerogenes* were tested for various antibiotics, including carbapenems and cephalosporins (Table 2). The Ka-04RR strain showed resistance to a greater number of antibiotics, mainly cephalosporins and β-lactams. The punctual pattern of resistance to carbapenems, cephalosporins and β-lactams suggests the absence of an enzymatic resistance mechanism. Even so, all strains would be considered multidrug resistant (MDR)[19].

K. aerogenes resistome analysis

The resistome of *K. aerogenes* genomes is heterogeneous, with 245 genomes presenting several genes associated with resistance to aminoglycosides, quinolones, macrolides, phenicol, sulfonamide, trimethoprim and β-lactams, while we did not observe acquired resistance genes in 316 genomes (Table S1). Among the β-lactamases genes, only 42 and 150 genomes, respectively, presented extended-spectrum β-lactamases (ESBL) and carbapenemases genes (Table S1). Therefore, although most genomes are of clinical origin, most of them do not carry carbapenemases. The most prevalent carbapenemases were bla_{KPC} (− 2 and − 3), bla_{OXA-48} and bla_{NDM} (− 1, − 5, − 7 and − 9) (Table S1). Most bla_{KPC} and bla_{NDM} genes were identified in the context of insertion sequences and transposases, where IS*26* and Tn*4401* (a/b/d isoforms) surrounded bla_{KPC}, and IS*5*, IS*630*, IS*Ab125*, and Tn*3*-like element IS*3000* surrounded bla_{NDM}. Several genomes (n = 169; 30%) carried genes associated with three or more classes of antibiotics, characterizing them as MDR, where most of these belonged to ST93 (n = 43), ST4 (n = 14) and ST116 (n = 11) (Table S1). In contrast, there are also genomes of pandemic clones ST4 (n = 39) and ST93 (n = 71) that lack acquired resistance genes (Table S1). Likewise, the Brazilian genomes sequenced here also did not show acquired resistance genes, including the ST93 (Ka-04RR), although this strain is MDR. Only 12 genomes were predicted to harbor the *mcr* gene (Ka37751, GCA_007632255.1; Ka.44724, GCA_007558485.1; Ka.42825, GCA_007558675.1; Ec.42059, GCA_007558345.1; Ka.37756, GCA_007558625.1; Ka.42442, GCA_007558645.1; Ka.43407, GCA_007558665.1; Ka.44253, GCA_007571405.1; Ka.46782, GCA_007558425.1; E845, GCA_023040435.1;

	Complete BUSCO (%)	Coverage depth	# Contigs	Largest contig (bp)	Total length (bp)	N50	L50	GC (%)	CDS	rRNA	tRNA
Ka-01RR	95.95	19	391	85,021	5,045,871	22,918	65	55.12	4758	7	79
Ka-02RR	95.95	23	364	97,818	5,043,146	25,629	62	55.12	4728	8	78
Ka-04RR	95.95	23	373	93,200	5,175,129	25,357	61	55.06	4815	7	78
Ka-06RJ	97.3	24	238	319,657	5,097,888	41,592	33	55.03	4752	9	80

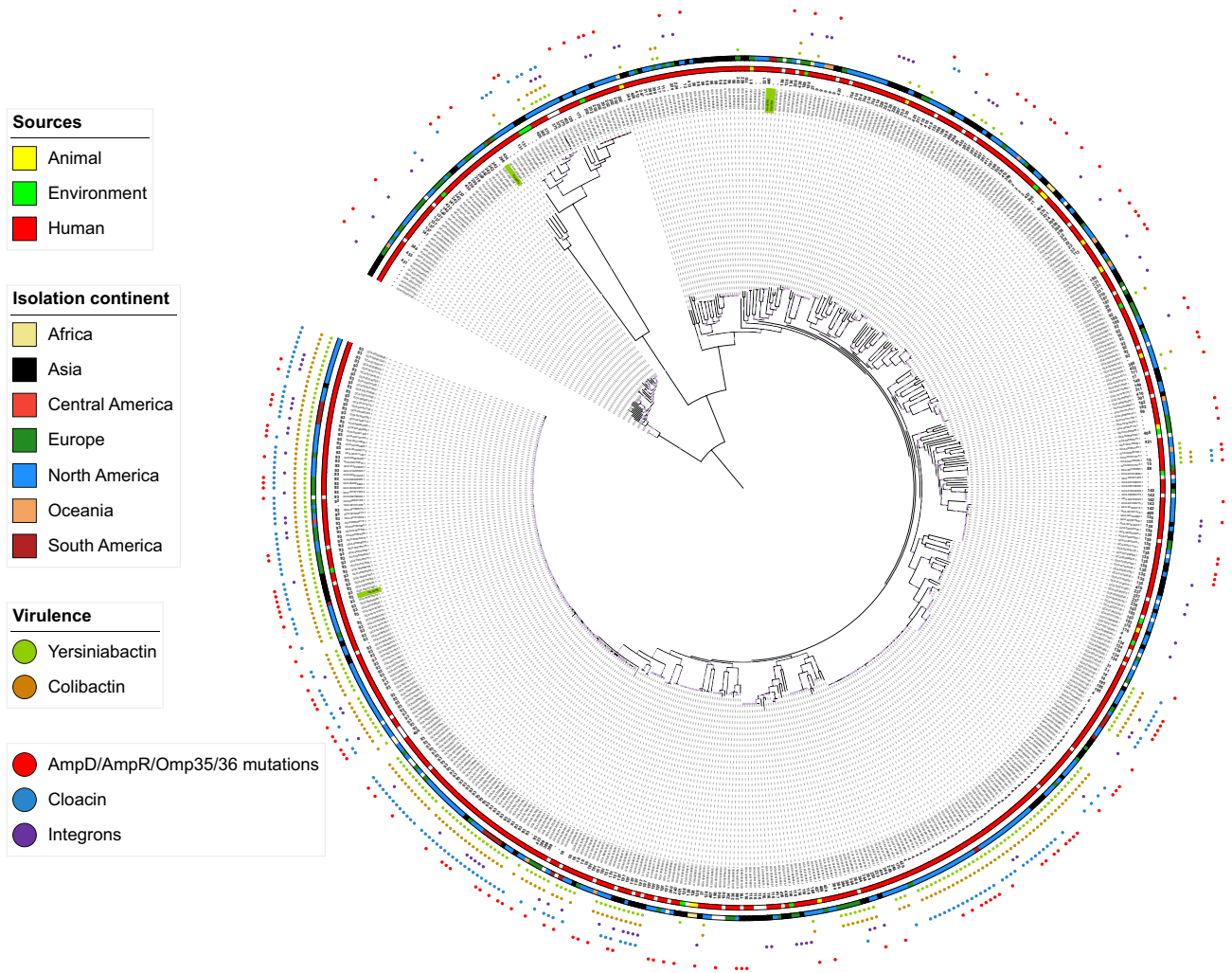**Table 1.** Assembly information of sequenced Brazilian genomes.

Figure 1. Maximum-likelihood tree based on the core gene alignment of the 561 K. *aerogenes* genomes. The ST number of each genome is next to the accession number. There are two orbits of colored blocks, where the innermost represents the source of isolation, and the outermost represents the regions of the genomes. Next, five colored circles indicate genomes with yersiniabactin (green), colibactin (beige), integron (purple), cloacin (blue), nonsense mutations in AmpD, AmpR, Omp35/36 (red)[4,20,21]. The genomes sequenced here are in green background color. Red circles at branches indicate bootstrap values > 80%.

201414107U, GCA_025071235.1; 4417, GCA_024604175.1), where nine were related to ST56 in China (Table S1). Moreover, analyzing PmrA, three Dutch ST364 genomes (isolate 25, GCA_024074185.1; isolate 26, GCA_024074205.1; isolate 27, GCA_024073885.1) were observed carrying the G53S amino acid substitution (Table S2), which has been associated with colistin resistance[9]. The GCA_000950145.1 genome (strain GCSL-DIFS-311) showed a large block of amino acid replacements from position 85 to 109, including insertions and deletions (Table S2), which could also affect its role. Interestingly, the L162M amino acid replacement was observed in almost all genomes (551/561), even in some colistin-susceptible strains (Table S2). Furthermore, three genomes had nonsense mutations in MgrB (N1431, GCA_021837555.1; N676, GCA_021837885.1; CK-00624, GCA_022016415.1), also being associated with colistin resistance[20].

In vitro analyses showed that Ka-04RR was resistant to meropenem and to most cephalosporins and β-lactams tested (Table 2). However, this strain lacked acquired ESBL and carbapenemases genes (Table S1). Thus, we explored amino acid replacements at chromosomal loci associated with β-lactam resistance in this strain, in addition to the entire genomic dataset. We focused on AmpC and outer membrane porins (Omp35 and Omp36), including their associated regulatory proteins (AmpD, AmpG, AmpR, and OmpR). Comparisons were evaluated using wild-type sequences from carbapenem-susceptible strains. All respective genes (*amp*C, *omp*35, *omp*36, *amp*D, *amp*G, *amp*R, and *omp*R) were identified in most genomes (Table S3). They showed different patterns of single nucleotide polymorphisms (SNPs), some leading to premature translation stops in their respective proteins. The AmpD protein was truncated in 44 genomes, including Ka-04RR. Mutation analysis, considering all genomes, showed that 65% (120/187 sites) of amino acid sites were conserved and several replacements were common, such as V134A, V135A, Q138R (Table S4). Furthermore, replacements associated with carbapenem resistance (P39S, A94T, W95L, S112L, I113S, I160V, R161H[4]) were observed in several genomes, most of them belonging to ST4

Antibiotic classes	Antibiotics	Strains			
		Ka-01RR	Ka-02RR	Ka-04RR	Ka-06RJ
Aminoglycosides	AMK	S	I	I	S
	GEN	S	I	S	S
	STR	S	I	R	S
	TOB	S	R	S	S
	NEO	S	I	S	S
	KAN	S	I	S	I
β-Lactams: carbapenems	IPM	R	I	S	S
	MEM	I	I	R	R
	ERT	S	S	S	S
β-Lactams: cephalosporins	CAZ	R	S	R	S
	CRO	S	S	R	S
	CTX	I	I	R	S
	FEP	I	S	I	S
	FOX	S	S	R	R
	CXM	I	S	R	S
β-Lactams: others	SAM	R	I	R	S
	TZP	S	S	R	S
	TIM	S	S	R	S
Quinolones	CIP	I	R	I	I
	LEV	S	I	S	S
Sulfonamides/antifolate	SXT	S	R	S	S
Tetracyclines	MIN	R	R	S	S
	TET	S	R	S	S
Polymyxins	COL	S	S	S	S

Table 2. Antibiotic susceptibility of Brazilian *K. aerogenes*. *AMK* amikacin, *GEN* gentamicin, *STR* streptomycin, *TOB* tobramycin, *NEO* neomycin, *KAN* kanamycin, *IPM* imipenem, *MEM* meropenem, *ERT* ertapenem, *CAZ* ceftazidime, *CRO* ceftriaxone, *CTX* cefotaxime, *FEP* cefepime, *FOX* cefoxitin, *CXM* cefuroxime, *SAM* ampicillin/sulbactam, *TZP* piperacillin/tazobactam, *TIM* ticarcillin/clavulanic acid, *CIP* ciprofloxacin, *LEV* levofloxacin, *SXT* trimethoprim/sulfamethoxazole, *MIN* minocycline, *TET* tetracycline, *COL* colistin. Resistance profile: *S* susceptible, *R* resistant, *I* intermediate.

and ST93 (Table S4). The AmpG sequences showed 91% (451/491 sites) of conserved amino acid sites (Table S5), where four genomes had truncated proteins, including the Ka01-RR genome, while the Ka02-RR genome lacked *amp*G. Also, several mutational patterns were observed, such as L89I, I93V, T424I, L430V, G461A and I479V (Table S5). Even so, all known AmpG activation motif residues (G25, A122, Q124, A181[21]) were conserved, except for GCA_021937655.1 (strain 122,288) (A181V). AmpR was found truncated in two genomes (presenting premature stop codons in their *amp*R genes) and absent in four genomes (strain UCI 45, GCA_000534135.1; GCSL-DIFS-311, GCA_000950145.1; 035, GCA_011604725.1; NY1688, GCA_022759585.1), which also lacked *amp*C (Table S3). Overall, the genomes showed 87% (255/292 sites) of conserved amino acid sites (Table S6). AmpR sites associated with its role (R86, G102, D135[22,23] were also conserved except for D135A/N substitutions in four genomes (strain 2485STDY5438452, GCA_900558125.1; 20174220 M, GCA_025115605.1; CK-00624, GCA_022016415.1; 95009, GCA_021936135.1) (Table S6). Regarding the porins, Omp35 was truncated in 21 genomes, but well conserved in the remaining genomes (93% of amino acid sites conserved; 337/359 sites) (Table S7). Most ST202, ST296, ST300, ST375 genomes had common mutations: S182V, D225N, D259N, H260Y (Table S7). Omp36 was the analyzed protein with the highest rates of nonsense mutations, where 98 genomes had premature stop codons. Among the Omp36 positive genomes, several sites had mutations, which resulted in 72% (270/375 sites) of conserved sites (Table S8). Most mutations were concentrated in the final half of the protein (Table S8). Some replacements were frequently found in ST-specific genomes, such as I59V in ST4, ST231, ST237; D189G in ST4, ST16, ST93; D205E in ST93 and ST228 (Table S8). The OmpR showed the highest conservation among the proteins explored with 97% (233/239 sites) of conserved amino acid sites (Table S9). Two important sites (G63 and R150[24,25]) were conserved in all genomes.

Among the proteins analyzed, amino acid replacements and/or premature translation stops in AmpD, AmpR, Omp35, and Omp36, which could lead to a phenotype of greater resistance to β-lactam, were present in 148 genomes spread across the phylogeny, regardless the ST (Fig. 1). Most of these genomes (117/148; ~ 80%) did not contain ESBL (19 of these 148 genomes had ESBL) and/or carbapenemase (12 of these 148 genomes had carbapenemase) genes, suggesting that these mutations could be selection mechanisms in the absence of enzymatic resistance genes. Even more so because most of these genomes (100/148; ~ 67%) also lacked acquired resistance

genes (Table S1). Thus, contrasting to the low prevalence of acquired genes, a large proportion of genomes carrying chromosomal alterations associated with resistance to β-lactam was identified.

Integrons (n = 139) could be identified in 96 genomes, mainly from ST93 (n = 29) (Table S10), where almost all were class 1 and only one class 3 integrons (strain 24A19CPO031, DAJACA010000050.1). As integrons are known to capture resistance genes, we searched for these genes in the identified integrons (Table S10). Different combinations of gene cassettes related to resistance were present, where most integrons presented the *dfr*A gene, while a smaller proportion presented *qac*E, *ant*, *aac*, *aph*, *arr*, bla_{OXA}, and bla_{GES}, some of them presenting several alleles. Furthermore, bla_{VIM} and bla_{IMP} carbapenemase genes were also present in gene cassettes (Table S10). Based on the gene cassette sequences of the integrons, 97 of them could be grouped into 22 clusters, with several of them present in genomes of different STs (Table S10).

K. aerogenes virulence analysis

Among the virulence factors identified in the genus *Klebsiella*, the yersiniabactin, colibactin, and aerobactin loci represent clinically important features. Thus, considering the 561 *K aerogenes* genomes, 222 genomes co-carried the *ybt* and *clb* loci, and 16 carried the *ybt* loci, while only one genome presented the aerobactin loci (strain 2022LN-00016, GCA_023553355.1) (Table S1). Furthermore, among the major STs, these loci were prevalent in the genomes of pandemic clones ST93 (n = 114) and ST4 (n = 47) (Table 3). Almost all *ybt* and *clb* loci were in the context of ICEKp10 (208/238 *ybt*+ genomes), regardless of ST, with the majority of ST4 having the *ybt* type 17 (n = 45), and majority of ST93 having the *ybt* type 17 (n = 59) and *ybt* type 20 (n = 52) (Table S1), while only the *clb* type 3 was found associated with these loci. ICEKp other than ICEKp10 were identified in a few genomes and carried only the *ybt*, lacking *clb* (Table S1). Only one environmental genome (ST93) presented the *ybt* 17 and *clb* 3 loci in the context of ICEKp10, while other animal and environmental genomes lacked these virulent loci. Curiously, the salmochelin (*iro*) virulence trait is present in almost all *K. aerogenes* genomes (517/561; ~ 92%) (Table S1), of which 88 did not present the entire cluster (*iro*BCDEN). In addition, although lacking the *ybt* and *clb* loci, most animal and environmental genomes carried the *iro* loci (19/27; ~ 70%). Another prevalent virulence factor among the genomes was the Type VI Secretion System (T6SS), present in 553/561 (~ 98%). Regarding toxins, two factors (EAST1 and *sen*B) were found in the genomes of ST240, ST15, ST93, ST331, ST364. In short, ICEKp10 carrying the *ybt* and *clb* is widespread in different STs, prevailing in ST4 and ST93.

In four genomes (GCA_020982565.1, GCA_001631645.1, GCA_003952125.1, GCA_021902315.1), the *ybt* type 4 was predicted, this type being associated with plasmid origin (Table S1). However, analyzing these genomes, we were unable to identify the plasmid replication genes. This could be due to genome assembly fragmentation, however, the *ybt* loci of these genomes were close to integrases and flanked by tRNA–Asn (in 3/4 genomes), which is a feature associated with chromosomal insertion. Thus, if *ybt* type 4 is on a mobile element, it appears to have the ability to integrate into the chromosome.

The US clinical genome GCA_023553355.1 (ST432) was the only one presenting the aerobactin loci (*iuc* 1; *iuc*A-D and *iut*A), in addition to the hypermucoviscosity-associated gene *rmp*A2 (Table S1). These loci likely lie within a genomic island (contig ABGJKZ010000015.1), in a region of approximately 75 kb length (7374 bp–82,380 bp) flanked by repetitive regions (93 bp repeats with 98% identity) and three transposase genes (ISL3-like element, ISNCY-like element). Furthermore, this region has 99% identity and 90% coverage with several *Klebsiella pneumoniae* plasmids (e.g. pFQ61_ST383_NDM-5, pAP855, pVir-CR-hvKP-C789), some of them designated as hypervirulent plasmids. Curiously, although only this genome contained the *iuc*A-D loci, most genomes (n = 552; 98%) harbored the *iut*A gene, which encodes the aerobactin receptor.

K. aerogenes plasmid resistome/virulome

We explored the 561 genomes for potential plasmids based on plasmid replicons, and identity and coverage with previously reported plasmids. Contig sequences presenting plasmid replicons and with high identity and coverage to previously reported complete plasmids were found in 317 genomes (317/561; 56%), which totaled 541 putative plasmid sequences (Table S11). Although it is not possible to state whether these contigs represent entire plasmids or fragments thereof, the presence of the plasmid replicon associated with high identity and coverage to previously reported plasmids suggests that these contigs belong to elements with an extra-chromosomal nature. The putative 541 plasmids had a median size of 10 kb and 52% GC content; ranging from 1 to 6 per genome (median of one plasmid); and ColRNAI was the most common replicon type. Based on the mobility markers present in these putative plasmids, we predicted that 51 would be conjugative (44.9 kb median size, 47% GC content), as they carried relaxase, *vir*D4, and *vir*B4 homologues genes[26]; 339 would be mobilizable (9.4 kb median size, 55%

ST	#	*ybt*	*clb*	*iro*	T6SS
93	121	114 (94%)	114 (94%)	121 (100%)	118 (97%)
4	57	47 (82%)	47 (82%)	57 (100%)	55 (96%)
135	16	0	0	16 (100%)	16 (100%)
116	13	1 (7%)	1 (7%)	13 (100%)	13 (100%)
197	12	8 (66%)	8 (66%)	12 (100%)	12 (100%)
56	10	0	0	10 (100%)	10 (100%)

Table 3. Prevalence of virulence loci in the main STs of *K. aerogenes*.

GC content) due to the presence of oriT sequences and/or relaxase genes, without T4SS-like genes; and 151 would be non-mobilizable (16 kb median size, 49% GC content) (Table S11). These putative plasmids were mainly harbored in the ST93 (n = 93) and ST4 (n = 34) genomes but were also present in other STs in smaller numbers (n < 10). These genomes are from several continents, presenting a wide time scale (2002 to 2021).

Interestingly, most of the putative plasmids carrying the ColRNAI replicon (172/194; ~ 88%) had a median length of 9.2 kb and a similar structure, mainly comprising a cloacin gene (a type of colicin) and its immunity gene, $traM$, $mobC$, $relB/relE$ toxin/antitoxin, and rop genes (Fig. 2). Variations were observed in these mobilizable sequences, for example, the sequence DAFYDZ010000031.1 presented genes related to mercury resistance; FKIV01000020.1 presented the bla_{TEM} gene; and Tn3 transposase was present in sequence LULD01000059.1 (Fig. 2). Virulence genes were rare in the putative plasmids, where only five had one gene, including the $traJ$ (associated with invasion), EAST1 (exotoxin), and $senB$ (exotoxin) genes (Table S12).

Considering resistance genes, 111/541 putative plasmids (in 82 genomes), with a median size of 46.1 kb, carried 1 to 12 genes (median of two genes) (Table S13), associated with several classes of antibiotics, including cephalosporins (bla_{TEM}, bla_{CTX-M}, bla_{OXA}), carbapenems (bla_{NDM}, bla_{IMP}, bla_{KPC}, bla_{VIM}), aminoglycosides (aac, aad, ant, $armA$, $aphA$), phenicol (cat), sulfonamides (sul), macrolides ($mphA$, $msrE$), quinolones ($qnrS$). The most prevalent resistance genes in the putative plasmids were BRP (gene encoding bleomycin resistance protein) (present in 22 putative plasmids); bla_{NDM-1} (n = 17); bla_{TEM-1} (n = 16); bla_{KPC-2} (n = 12); $qnrS1$ (n = 12), where the majority were spread across genomes of different STs (Table S13). Importantly, the mcr gene (colistin resistance) was associated with putative plasmids (10/12 genomes with mcr). These genomes belonged to ST56/China (8/10; IncI2 replicons) and ST207/Singapore (2/10; IncX4 replicons), in which the Chinese ones co-harbored the $bla_{CTX-M-199}$ gene (Table S13). Moreover, most of these plasmids were predicted as conjugative. Interestingly, these plasmids had approximately 100% identity and coverage with various *E. coli* plasmids (e.g. pC6-2, pEC1188-MCR, pZE36, pMCR-1_Msc, pDIB-1, pMCR_1139_A1) from strains from various locations around the world, which suggests that these elements are subject to horizontal transmission. In addition, the
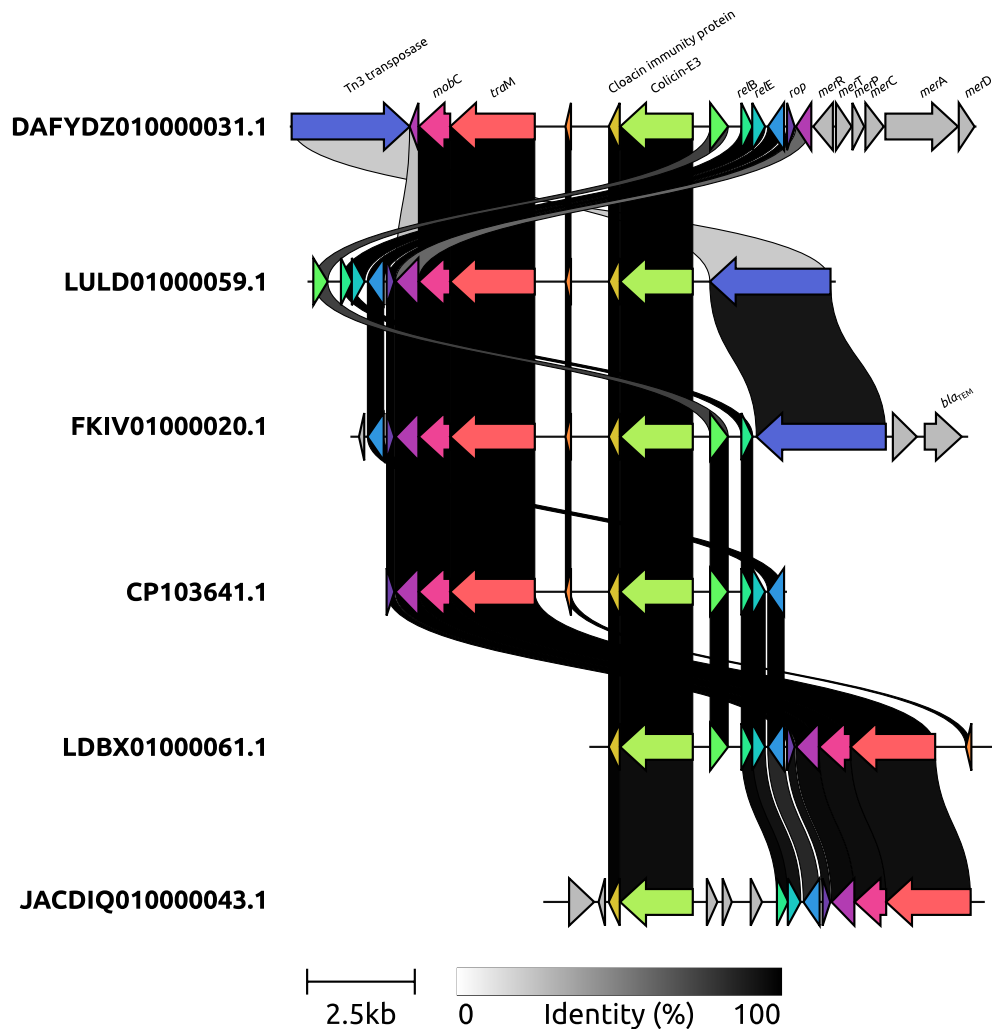

Figure 2. Synteny variations of putative *K. aerogenes* ColRNAI plasmids. A comparison of putative plasmids is shown, where the cloacin gene and its immunity gene, $traM$, $mobC$, $relB/relE$ toxin/antitoxin, and rop genes appear to represent a core segment.

bla_{NDM-1} gene was carried in 17 putative plasmids of 11/24 positive bla_{NDM-1} genomes, mainly in the ST197/Singapore genomes (featuring two plasmids, each carrying a copy of bla_{NDM-1}) that also co-harbored the BRP gene (Table S13). Interestingly, some integrons (n = 14) were present in plasmids, in which CP042532 plasmid harbored the bla_{IMP}, *aac*, *cat*B, *qac* cassette genes; and CP070519 plasmid harbored *qac*, *arr*, *cat*B, *aac*, bla_{OXA-1} cassette genes.

Although most genomes are of human origin, few putative plasmids from other sources could be observed, some of them carrying resistance genes (e.g. CP050069 from water, carrying *aph*, *mph*A, *qnr*, bla_{VIM}, *cat*, *dfr*A, *sul* genes; CP047668 from chicken manure, carrying *aph*, $bla_{CTX-M-65}$, *fos*A, bla_{NDM} genes).

Discussion

There is a historical gap in the epidemiology of *K. aerogenes*, as it has always been considered an opportunistic pathogen, but recently interest in this organism has increased, mainly due to the emergence of carbapenem-resistant isolates[3]. Furthermore, this species was recently renamed, being previously known as *Enterobacter aerogenes*, which still influences researchers to keep the old species name, even in recent publications. However, in genomic information databases such as GenBank, this has been updated. Until now, studies focused on this emerging pathogen were small-scale, considering few genomes and/or with a restricted geographic perspective. Here, in addition to generating new genomes of this species in Brazil and contextualizing them in the global scenario, we carried out an in-depth analysis of the resistome, virulome, and plasmidome in view of the emergence of *K. aerogenes* as a pathogen.

Previous genomic epidemiological analyzes have shown that *K. aerogenes* ST4 and ST93 have been implicated as the dominant global clones[4], as have occurred in a few countries in the Americas, Europe, and China. In fact, here, expanding this analysis, encompassing all *K. aerogenes* genomes, these STs continue to represent the dominant clones, causing infections/outbreaks in several other countries (e.g. Belgium, Germany, Japan, Qatar, Singapore, South Korea, Switzerland, Thailand, United Kingdom), thus characterizing them as pandemic clones that impact the clinics.

K. aerogenes genomes can be clustered by the presence/absence of two virulence determinants: the yersinia-bactin and colibactin[5]. In fact, it is observed that most STs lack these loci, while they are predominant in the ST4 and ST93 genomes. These virulence loci were found in several ICEKp, mainly in ICEKp10, and their distribution is not homogeneous among all genomes, even within the same ST, which shows their mobile characteristic. In *K aerogenes*, ICEKp10 had already been identified in a few strains of ST4 and ST93 from the USA[4]. The present analysis revealed a wide distribution of this virulence-associated element within this species, particularly in ST4 and ST93 genomes from several countries around the world. Moreover, colibactin, a genotoxin that can induce DNA damage in eukaryotic cells and tumor formation[27], raises concerns about its high prevalence and distribution in *Klebsiella* species.

In contrast to the *ybt* and *clb* loci, the salmochelin loci (*iro*), another virulence trait identified, was present in a high proportion of genomes (~ 92%), including those of animal and environmental origin. This result is similar to the findings of other studies considering smaller genome sets[5,28], and suggests that *K. aerogenes* may be a potential reservoir of virulence genes for other bacteria[28]. In fact, in *K. pneumoniae*, salmochelin is detected in low prevalence, being more identified in mobile elements, such as plasmids[29]. Although prevalent, 88 *K aerogenes* genomes had incomplete salmochelin loci. In avian pathogenic *Escherichia coli*, it was observed that the absence of *iro*C, *iro*DE, or *iro*N abrogated the virulence of the bacteria[30]. Thus, it can be hypothesized that the *iro* loci in these *K. aerogenes* would not act on virulence as in *E. coli*. Furthermore, the presence of incomplete loci may also suggest that other functions could be involved, such as colonization and survival linked to commensalism, as they can increase the fitness of strains within a specific niche.

A putative virulence island, carrying the aerobactin loci (*iuc*A-D and *iut*A) and *rmp*A2 gene, was found in a *K. aerogenes* genome (GCA_023553355.1). This island has also been observed in several *K. pneumoniae* plasmids. In fact, recently, a transposon harboring the aerobactin operon was identified in *K. pneumoniae* virulence plasmids[31]. Most of these plasmids harbor the *rmp*A/*rmp*A2 genes, aerobactin, and salmochelin loci, which greatly enhance the virulence of *K. pneumoniae* strains[32]. Although salmochelin is strongly associated with aerobactin in *K. pneumoniae*[33], this was not observed in *K. aerogenes* as *iro* was prevalent and *iuc* rare. Thus, we show here that a mobile element carrying *K. pneumoniae* hypervirulence marker genes is being transferred between species of the genus *Klebsiella*.

Although most genomes did not carry acquired antibiotic resistance genes, having only the intrinsic resistome, this does not mean an absence of resistance, since in vitro analyses of Ka-04RR showed its MDR profile, and further in silico analyses revealed the presence of alterations in chromosomally encoded factors associated with this MDR profile. Indeed, chromosomal alterations have been observed in several other genomes. Most genomes (74%; 411/561) do not carry carbapenemase genes (including Ka-04RR), however, almost all of them carry the *amp*C gene, which if overexpressed could lead to carbapenem resistance[3]. In fact, AmpC is chromosomally encoded by *K. aerogenes,* and alterations in its regulatory genes could affect AmpC translation[21]. Here we observed several amino acid replacements in AmpC regulatory proteins, some of them already associated with resistance to carbapenems (mainly in AmpD), in addition to a large set of other replacements yet to be characterized in further studies. Moreover, ~ 21% of the genomes showed truncated Omp35 and Omp36 porins, which could also increase resistance to β-lactam agents[3,34]. Therefore, regardless of the absence of β-lactamase genes in most genomes, we raised evidence of other possible mechanisms that could confer some level of resistance to β-lactams in *K. aerogenes* circulating the world. Unlike *K. pneumoniae*, reports of integrons in *K. aerogenes* are scarce[35–37]. Here, this genetic element was identified in ~ 17% of the analyzed genomes, presenting several combinations of gene cassettes, some not yet reported in this species, such as bla_{OXA-1}, bla_{OXA-10}, bla_{IMP}, bla_{VIM}, bla_{GES}. Thus, carbapenemase genes can also be captured by these genetic platforms in this species.

In Brazil, studies on *K. aerogenes* have focused on molecular analyses of resistance genes from carbapenem-resistant isolates. These studies cover isolates from three Brazilian regions (Minas Gerais, Paraná, and Pernambuco states), most of them co-harboring the bla_{KPC-2} and bla_{TEM} genes[38–40]. In fact, we observed that several Brazilian of ST4 and ST93 genomes presented these genes (Table S1), and plasmids carrying carbapenemase genes (bla_{KPC-2}, bla_{NDM-1}) have already been observed among *K. aerogenes* in the country[41,42]. In addition, we provided new genomes from other regions, expanding regional epidemiology. Although data on resistance genes in *K. aerogenes* in Brazil are available, most studies have not defined the ST of these strains, with only reports of ST93 and ST16[10,43]. Thus, the current epidemiological scenario of *K. aerogenes* in Brazil is mainly driven by ST93, ST16 and also by ST4, as is the case worldwide. In addition, a clinical bla_{NDM}-producing *K. aerogenes* was recently identified in the country belonging to ST128[44]. Here, ST128 was also associated with environmental (Asia) and animal (Africa) strains, therefore revealing the One Health trait of this species, particularly of ST128. Furthermore, this ST is widespread and evolving in the context of the acquisition of resistance and virulence genes.

In China, bla_{NDM} alleles were associated with plasmids in *K. aerogenes* ST4[45,46]. Here, the bla_{NDM} gene was also found in several putative plasmids from Singaporean genomes. Interestingly, both countries were the only ones to present genomes with putative plasmids carrying the *mcr* gene (colistin resistance). Indeed, the initial report of *mcr* in plasmids occurred in China (2016)[47], and our analysis showed that the presence of the *mcr* gene in putative plasmids of *K. aerogenes* still seems to be restricted to Asia. Even though these putative plasmids are already widespread in *E. coli* worldwide. Although plasmids drive the exchange of antibiotic resistance in other countries, we could only observe five Brazilian ST93 genomes harboring putative plasmids with bla_{KPC-2}. Therefore, other transfer mechanisms may be acting in other acquired genes. Nevertheless, if plasmids are also driving *K. aerogenes* adaptation in Brazil, more sequences should be made available to determine this question, in addition to more data for epidemiology.

The most common type of plasmid replicon identified was the ColRNAI, as observed by Passarelli–Araujoa et al.[5] on a smaller set of genomes. Interestingly, 172/194 of ColRNAI replicon putative plasmids have been observed carrying colicin E3 (cloacin-like; rRNase activity) and its immunity gene. Colicins are bacteriocins produced by some gram-negative bacteria, showing antibacterial activity against closely related species, and being highly found in natural populations of *E. coli*[48,49]. Indeed, blasting cloacin-bearing *K. aerogenes* putative plasmids, several *E. coli* and *K. pneumoniae* plasmids showed high coverage and identity of > 75% and > 99%, respectively. It has been reported that Cloacin-like exhibits a weaker effect and narrower spectrum of activity against several species, but exhibits high inhibitory activity against *K. aerogenes*, which suggests that these plasmids are associated with the ecology of this species[49]. Interestingly, most *K. aerogenes* genomes contained the *iut*A gene, which encodes a receptor for aerobactin, a rare virulence factor in the current dataset. However, curiously, this receptor is also used by cloacins[49], suggesting that the main function of this receptor in *K. aerogenes* may not be related to virulence, but perhaps to ecological functions.

In conclusion, *K. aerogenes* is a species capable of acquiring resistance genes from other organisms, but it seems to be more prone to becoming resistant by mutations in housekeeping genes, something that could only be assessed by in-depth analysis since most of the in silico resistome survey platforms are not up to date regarding this information. Furthermore, there are *K. aerogenes* lineages circulating around the world carrying virulence genes that may be determinants for its emergence as a clinical pathogen.

Methods

Public data set

All reported *Klebsiella aerogenes* genomes (n = 557), including complete and draft ones, were obtained from the Genome database of the National Center for Biotechnology Information (NCBI) in October 2022. The accession numbers are supplied in Table S1.

Isolates and genome sequencing and assembly

In this study we analyzed four *Klebsiella aerogenes* strains from nosocomial cases in the Amazonic and Southeast regions of Brazil: Ka-01RR (subclavian vein catheter tip), Ka-02RR (tracheal secretion), Ka-04RR (unknown), Ka-06RJ (sputum). The genomic DNA extraction was done using the NucleoSpin Microbial DNA kit (Macherey–Nagel), and the genomic libraries were constructed using Nextera paired-end library. The sequencing was performed using Illumina Hiseq 2500, generating reads of 150 bp length. The raw reads were filtered and trimmed using NGS QC Toolkit v.2.3.3[50] with a Phred score ≥ 20. The genomes were de novo assembled using SPAdes assembler v3.14.1[51].

Antimicrobial susceptibility testing

The four *Klebsiella aerogenes* strains were grown on Mueller Hinton Agar at a temperature of 35–37 °C for antimicrobial susceptibility testing by the disk-diffusion method using 24 antibiotics (amikacin, gentamicin, streptomycin, tobramycin, neomycin, kanamycin, imipenem, meropenem, ertapenem, ceftazidime, ceftriaxone, cefotaxime, cefepime, cefoxitin, cefuroxime, ampicillin/sulbactam, piperacillin/tazobactam, ticarcillin/clavulanic acid, ciprofloxacin, levofloxacin, trimethoprim/sulfamethoxazole, minocycline, tetracycline, colistin) considered for *Enterobacteriaceae* resistance classification[19] and interpreted according to the Clinical and Laboratory Standards Institute (31st ed.)[52]. The minimum inhibitory concentration of colistin was determined by the broth microdilution method and interpreted according to the European Committee for Antimicrobial Susceptibility Testing (EUCAST) guidelines (MIC breakpoint for resistance > 2 mg/L)[53].

Genome characterization

The *K. aerogenes* genomes were submitted to the Kleborate v2.1.0[33] pipeline to sequence typing, and identification of acquired virulence and antibiotic resistance genes. Integrons were survey with Integron Finder[54]. Type VI Secretion System (T6SS) was identified using the T6SS prediction tool[55]. Insertion Sequences and transposases associated with antibiotic resistance genes were identified using TnFinder[56].

Plasmid identification was based on the characterization of contigs carrying a plasmid replicon marker (*rep*), using PlasmidFinder database (v2.1.1)[57], and presenting identity (=> 70%) and coverage (=> 50%) to known plasmids, using the PLSDB (v.2023_11_03_v2), a curated database of complete bacterial plasmids[58]. The putative plasmids were further characterized for virulence and antibiotic resistance genes using ABRicate v1.0.1 (https://github.com/tseemann/abricate) with the VFDB (accessed in Jun/2023)[59] and CARD (v3.2.7)[60] databases, respectively (identity and coverage => 60%). Plasmid clustering was done using CD-HIT-EST at 90% identity and 70% coverage[61]. The mobility of the putative plasmids was predicted based on the presence of gene markers, such as relaxase, Type IV Secretion System (T4SS) genes (e.g. VirB4 and VirD4), and OriT sequences (origin sites of DNA transfer), as described[26,62]. The putative plasmids were classified as: conjugative, if encoding relaxase, *vir*B4, and *vir*D4 homologues; mobilizable, if encoding relaxase and/or oriT sequences, and lacking the T4SS; and non-mobilizable, if not encoding relaxase and/or oriT sequences. These searches were performed using hmm profiles as previous described[62], in addition to the *vir*B4 homologue CagE_TrbE_VirB (Pfam PF03135), present in some *Proteobacteria*.

Phylogeny

The genomes were annotated by Prokka v1.12[63]. *K. aerogenes* core genome was estimated using Roary v3.13 with default parameters (considering a minimum percentage identity of 95%)[64], and single nucleotide polymorphisms (SNPs) were extracted from the concatenated core genes with snp-sites v2.5.1[65]. Next, phylogenetic analysis was performed by IQTree v1.6.12[66] using the GTR+ F+ ASC+ R5 model to obtain a maximum likelihood tree with 1000 ultrafast bootstrap replicates[67]. The tree was designed using iTOL platform[68].

Data availability

References

1. Chavda, K. D. *et al.* Comprehensive genome analysis of carbapenemase-producing *Enterobacter* spp.: New insights into phylogeny, population structure, and resistance mechanisms. *mBio* **7**(6), e02093-e2116. https://doi.org/10.1128/mBio.02093-16 (2016).
2. Tindall, B. J., Sutton, G. & Garrity, G. M. *Enterobacter aerogenes* Hormaeche and Edwards 1960 (approved lists 1980) and *Klebsiella mobilis* Bascomb et al. 1971 (approved lists 1980) share the same nomenclatural type (ATCC 13048) on the approved lists and are homotypic synonyms, with consequences for the name *Klebsiella mobilis* Bascomb et al. 1971 (approved lists 1980). *Int. J. Syst. Evol. Microbiol.* **67**(2), 502–504. https://doi.org/10.1099/ijsem.0.001572 (2017).
3. Merhi, G., Amayri, S., Bitar, I., Araj, G. F. & Tokajian, S. Whole genome-based characterization of multidrug resistant enterobacter and *Klebsiella aerogenes* isolates from Lebanon. *Microbiol. Spectr.* **11**(1), e0291722. https://doi.org/10.1128/spectrum.02917-22 (2023).
4. Malek, A. *et al.* Next-generation-sequencing-based hospital outbreak investigation yields insight into *Klebsiella aerogenes* population structure and determinants of carbapenem resistance and pathogenicity. *Antimicrob. Agents Chemother.* **63**(6), e02577-e2618. https://doi.org/10.1128/AAC.02577-18 (2019).
5. Passarelli-Araujoa, H. *et al.* Genomic analysis unveils important aspects of population structure, virulence, and antimicrobial resistance in *Klebsiella aerogenes*. *FEBS J.* **286**(19), 3797–3810. https://doi.org/10.1111/febs.15005 (2019).
6. Mazumder, R. *et al.* Case report: A successfully treated case of community-acquired urinary tract infection due to *Klebsiella aerogenes* in Bangladesh. *Front. Med.* **10**, 1206756. https://doi.org/10.3389/fmed.2023.1206756 (2023).
7. Jacoby, G. A. AmpC beta-lactamases. *Clin. Microbiol. Rev.* **22**(1), 161–182. https://doi.org/10.1128/CMR.00036-08 (2009).
8. Meini, S., Tascini, C., Cei, M., Sozio, E. & Rossolini, G. M. AmpC β-lactamase-producing Enterobacterales: What a clinician should know. *Infection* **47**(3), 363–375. https://doi.org/10.1007/s15010-019-01291-9 (2019).
9. Diene, S. M. *et al.* The rhizome of the multidrug-resistant *Enterobacter aerogenes* genome reveals how new "killer bugs" are created because of a sympatric lifestyle. *Mol. Biol. Evol.* **30**(2), 369–383. https://doi.org/10.1093/molbev/mss236 (2013).
10. da Silva, K. E. *et al.* Genetic diversity of virulent polymyxin-resistant *Klebsiella aerogenes* isolated from intensive care units. *Antibiotics* **11**(8), 1127 (2022).
11. Loiwal, V., Kumar, A., Gupta, P., Gomber, S. & Ramachandran, V. G. *Enterobacter aerogenes* outbreak in a neonatal intensive care unit. *Pediatr. Int.* **41**(2), 157–161. https://doi.org/10.1046/j.1442-200x.1999.4121033.x (1999).
12. Hao, M. *et al.* Outbreak of *Klebsiella pneumoniae* carbapenemase-producing *Klebsiella aerogenes* strains in a tertiary hospital in China. *Infect. Drug Resist.* **12**, 3283–3290. https://doi.org/10.2147/IDR.S221279 (2019).
13. Narayan, S. A. *et al.* Investigation and control of an outbreak of *Enterobacter aerogenes* bloodstream infection in a neonatal intensive care unit in Fiji. *Infect. Control Hosp. Epidemiol.* **30**(8), 797–800. https://doi.org/10.1086/598240 (2009).
14. Baier-Grabner, S. *et al.* A Yersiniabactin-producing *Klebsiella aerogenes* strain causing an outbreak in an Austrian neonatal intensive care unit. *Pediatr. Infect. Dis. J.* **41**(7), 593–599. https://doi.org/10.1097/INF.0000000000003553 (2022).
15. Hallbäck, E. T. *et al.* Outbreak of OXA-48-producing Enterobacteriaceae in a neonatal intensive care unit in western Sweden. *Eur. J. Clin. Microbiol. Infect. Dis.* **42**(5), 597–605. https://doi.org/10.1007/s10096-023-04584-y (2023).
16. Piagnerelli, M., Kennes, B., Brogniez, Y., Deplano, A. & Govaerts, D. Outbreak of nosocomial multidrug-resistant *Enterobacter aerogenes* in a geriatric unit: Failure of isolation contact, analysis of risk factors, and use of pulsed-field gel electrophoresis. *Infect. Control Hosp. Epidemiol.* **21**(10), 651–653. https://doi.org/10.1086/501704 (2000).
17. Neuwirth, C., Siebor, E., Lopez, J., Pechinot, A. & Kazmierczak, A. Outbreak of TEM-24-producing *Enterobacter aerogenes* in an intensive care unit and dissemination of the extended-spectrum beta-lactamase to other members of the family enterobacteriaceae. *J. Clin. Microbiol.* **34**(1), 76–79. https://doi.org/10.1128/jcm.34.1.76-79.1996 (1996).

18. Cheng, Q. *et al.* Whole-genome sequencing analysis of *Klebsiella aerogenes* among men who have sex with men in Guangzhou, China. *Front. Microbiol.* **14**, 1102907. https://doi.org/10.3389/fmicb.2023.1102907 (2023).

19. Magiorakos, A. P. *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: An international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**(3), 268–281. https://doi.org/10.1111/j.1469-0691.2011.03570.x (2012).

20. Bedenić, B. *et al.* Emergence of colistin resistance in *Enterobacter aerogenes* from Croatia. *J. Chemother.* **30**(2), 120–123. https://doi.org/10.1080/1120009X.2017.1382121 (2018).

21. D'Souza, R. *et al.* Role of AmpG in the resistance to β-lactam agents, including cephalosporins and carbapenems: Candidate for a novel antimicrobial target. *Ann. Clin. Microbiol. Antimicrob.* **20**(1), 45. https://doi.org/10.1186/s12941-021-00446-7 (2021).

22. Kuga, A., Okamoto, R. & Inoue, M. ampR gene mutations that greatly increase class C beta-lactamase activity in *Enterobacter cloacae*. *Antimicrob. Agents Chemother.* **44**(3), 561–567. https://doi.org/10.1128/AAC.44.3.561-567.2000 (2000).

23. Caille, O. *et al.* Structural and functional characterization of *Pseudomonas aeruginosa* global regulator AmpR. *J. Bacteriol.* **196**(22), 3890–3902. https://doi.org/10.1128/JB.01997-14 (2014).

24. Dupont, H. *et al.* Structural alteration of OmpR as a source of ertapenem resistance in a CTX-M-15-producing *Escherichia coli* O25b:H4 sequence type 131 clinical isolate. *Antimicrob. Agents Chemother.* **61**(5), e00014-17. https://doi.org/10.1128/AAC.00014-17 (2017).

25. Ching, C. & Zaman, M. H. Identification of multiple low-level resistance determinants and coselection of motility impairment upon sub-MIC ceftriaxone exposure in *Escherichia coli*. *mSphere* **6**(6), e0077821. https://doi.org/10.1128/mSphere.00778-21 (2021).

26. Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. & de la Cruz, F. Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74**(3), 434–452. https://doi.org/10.1128/MMBR.00020-10 (2010).

27. Wernke, K. M. *et al.* Structure and bioactivity of colibactin. *Bioorg. Med. Chem. Lett.* **30**(15), 127280 (2020).

28. Spadar, A., Perdigão, J., Campino, S. & Clark, T. G. Large-scale genomic analysis of global *Klebsiella pneumoniae* plasmids reveals multiple simultaneous clusters of carbapenem-resistant hypervirulent strains. *Genome Med.* **15**(1), 3. https://doi.org/10.1186/s13073-023-01153-y (2023).

29. Lam, M. M. C. *et al.* Tracking key virulence loci encoding aerobactin and salmochelin siderophore synthesis in *Klebsiella pneumoniae*. *Genome Med.* **10**(1), 77. https://doi.org/10.1186/s13073-018-0587-5 (2018).

30. Caza, M., Lépine, F., Milot, S. & Dozois, C. M. Specific roles of the iroBCDEN genes in virulence of an avian pathogenic *Escherichia coli* O78 strain and in production of salmochelins. *Infect. Immun.* **76**(8), 3539–3549. https://doi.org/10.1128/IAI.00455-08 (2008).

31. Kaspersen, H. *et al.* Highly conserved composite transposon harbouring aerobactin iuc3 in *Klebsiella pneumoniae* from pigs. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000960 (2023).

32. Fan, L. P. *et al.* Genetic characterization and passage instability of a novel hybrid virulence plasmid in a ST23 hypervirulent *Klebsiella pneumoniae*. *Front. Cell. Infect. Microbiol.* https://doi.org/10.3389/fcimb.2022.870779 (2022).

33. Lam, M. M. C. *et al.* A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat. Commun.* **12**(1), 4188. https://doi.org/10.1038/s41467-021-24448-3 (2021).

34. Davin-Regli, A. & Pagès, J. M. *Enterobacter aerogenes* and *Enterobacter cloacae*; Versatile bacterial pathogens confronting antibiotic treatment. *Front. Microbiol.* **6**, 392. https://doi.org/10.3389/fmicb.2015.00392 (2015).

35. Ploy, M. C., Courvalin, P. & Lambert, T. Characterization of In40 of Enterobacter aerogenes BM2688, a class 1 integron with two new gene cassettes, cmlA2 and qacF. *Antimicrob. Agents Chemother.* **42**(10), 2557–2563. https://doi.org/10.1128/AAC.42.10.2557 (1998).

36. Wu, K. *et al.* Class 1 integron gene cassettes in multidrug-resistant gram-negative bacteria in southern China. *Int. J. Antimicrob. Agents* **40**(3), 264–267. https://doi.org/10.1016/j.ijantimicag.2012.05.017 (2012).

37. Cheng, C. *et al.* New structures simultaneously harboring class 1 integron and ISCR1-linked resistance genes in multidrug-resistant gram-negative bacteria. *BMC Microbiol.* **16**, 71. https://doi.org/10.1186/s12866-016-0683-x (2016).

38. Tuon, F. F. *et al.* KPC-producing *Enterobacter aerogenes* infection. *Braz. J. Infect. Dis.* **19**(3), 324–327. https://doi.org/10.1016/j.bjid.2015.01.003 (2015).

39. Pereira, R. S. *et al.* Physiological and molecular characteristics of carbapenem resistance in *Klebsiella pneumoniae* and *Enterobacter aerogenes*. *Infect. Dev. Ctries.* **10**(6), 592–599. https://doi.org/10.3855/jidc.6821 (2016).

40. Cabral, A. B. *et al.* Clonal spread and accumulation of β-lactam resistance determinants in *Enterobacter aerogenes* and *Enterobacter cloacae* complex isolates from infection and colonization in patients at a public hospital in Recife, Pernambuco, Brazil. *J. Med. Microbiol.* **66**(1), 70–77 (2017).

41. Bispo Beltrão, E. M. *et al.* Multidrug-resistant *Klebsiella aerogenes* clinical isolates from Brazil carrying IncQ1 plasmids containing the blaKPC-2 gene associated with non-Tn4401 elements (NTEKPC-IId). *J. Glob. Antimicrob. Resist.* **22**, 43–44. https://doi.org/10.1016/j.jgar.2020.05.001 (2020).

42. Soares, C. R. P., Oliveira-Júnior, J. B. & Firmo, E. F. First report of a blaNDM-resistant gene in a *Klebsiella aerogenes* clinical isolate from Brazil. *Rev. Soc. Bras. Med. Trop.* **54**, e02622020. https://doi.org/10.1590/0037-8682-0262-2020 (2020).

43. Passarelli-Araujob, H. *et al.* Molecular epidemiology of 16S rRNA methyltransferase in Brazil: RmtG in *Klebsiella aerogenes* ST93 (CC4). *An. Acad. Bras. Cienc.* **91**(suppl 1), e20180762. https://doi.org/10.1590/0001-376520182018762 (2019).

44. Camargo, C. H. *et al.* Genomic diversity of NDM-producing *Klebsiella* species from Brazil, 2013–2022. *Antibiotics (Basel)* **11**(10), 1395 (2022).

45. Pan, F., Xu, Q. & Zhang, H. Emergence of NDM-5 producing carbapenem-resistant *Klebsiella aerogenes* in a pediatric hospital in Shanghai,China. *Front. Public Health* **9**, 621527 (2021).

46. Li, Y., Wang, Q., Xiao, X., Li, R. & Wang, Z. Emergence of blaNDM-9-bearing tigecycline-resistant *Klebsiella aerogenes* of chicken origin. *J. Glob. Antimicrob. Resist.* **26**, 66–68 (2021).

47. Liu, Y. Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: A microbiological and molecular biological study. *Lancet Infect. Dis.* **16**(2), 161–168. https://doi.org/10.1016/S1473-3099(15)00424-7 (2016).

48. Simons, A., Alhanout, K. & Duval, R. E. Bacteriocins, antimicrobial peptides from bacterial origin: Overview of their biology and their impact against multidrug-resistant bacteria. *Microorganisms* **8**(5), 639 (2020).

49. Le, M. N. *et al.* Comprehensive analysis of bacteriocins produced by the hypermucoviscous *Klebsiella pneumoniae* species complex. *Microbiol. Spectr.* **11**(3), e008632310 (2023).

50. Patel, R. K. & Jain, M. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS ONE* **7**, e30619. https://doi.org/10.1371/journal.pone.0030619 (2012).

51. Bankevich, A. *et al.* 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 (2012).

52. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing* 31st edn. M100 (2021).

53. EUCAST. *Breakpoint Tables for Interpretation of MICs and Zone Diameters (Version 11.0. 2021)* https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_11.0_Breakpoint_Tables.pdf. Accessed 13 Dec 2022.

54. Néron, B. *et al.* IntegronFinder 2.0: Identification and analysis of integrons across bacteria, with a focus on antibiotic resistance in *Klebsiella*. *Microorganisms* **10**(4), 700. https://doi.org/10.3390/microorganisms10040700 (2022).

55. Zhang, J. *et al.* SecReT6 update: A comprehensive resource of bacterial type VI secretion systems. *Sci. China Life Sci.* **66**(3), 626–634. https://doi.org/10.1007/s11427-022-2172-x (2023).

56. Ross, K. *et al.* TnCentral: A prokaryotic transposable element database and web portal for transposon analysis. *mBio* **12**(5), e0206021. https://doi.org/10.1128/mBio.02060-21 (2021).
57. Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**(7), 3895–3903. https://doi.org/10.1128/AAC.02412-14 (2014).
58. Galata, V., Fehlmann, T., Backes, C. & Keller, A. PLSDB: A resource of complete bacterial plasmids. *Nucleic Acids Res.* **47**(D1), D195–D202. https://doi.org/10.1093/nar/gky1050 (2019).
59. Chen, L. *et al.* VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* **33**, D325–D328. https://doi.org/10.1093/nar/gki008 (2005).
60. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**(D1), D517–D525. https://doi.org/10.1093/nar/gkz935 (2020).
61. Li, W. & Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**(13), 1658–1659. https://doi.org/10.1093/bioinformatics/btl158 (2006).
62. Morgado, S. M. & Vicente, A. C. P. Comprehensive in silico survey of the *Mycolicibacterium* mobilome reveals an as yet underexplored diversity. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000533 (2021).
63. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153 (2014).
64. Page, A. J. *et al.* Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421 (2015).
65. Page, A. J. *et al.* SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016).
66. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. https://doi.org/10.1093/molbev/msu300 (2015).
67. Minh, B. Q., Nguyen, M. A. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195. https://doi.org/10.1093/molbev/mst024 (2013).
68. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259. https://doi.org/10.1093/nar/gkz239 (2019).

Acknowledgements

Author contributions

Conceptualization, ACV, SM; methodology, SM, ACV; formal analysis, SM; investigation, FF; Resources, RC; writing—original draft preparation, SM and ACV; writing—review and editing, SM, ACV, and EF; supervision, ACV. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-57245-1.

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

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