

RESEARCH ARTICLE

Open Access



Molecular characterization of carbapenem-resistant *Klebsiella pneumoniae* isolates with focus on antimicrobial resistance

Xiaoling Yu¹, Wen Zhang², Zhiping Zhao¹, Chengsong Ye³, Shuyan Zhou⁴, Shaogui Wu⁴, Lifan Han¹, Zhaofang Han^{5,6*} and Hanhui Ye^{1*}

Abstract

Background: The enhancing incidence of carbapenem-resistant *Klebsiella pneumoniae* (CRKP)-mediated infections in Mengchao Hepatobiliary Hospital of Fujian Medical University in 2017 is the motivation behind this investigation to study gene phenotypes and resistance-associated genes of emergence regarding the CRKP strains. In current study, seven inpatients are enrolled in the hospital with complete treatments. The carbapenem-resistant *K. pneumoniae* whole genome is sequenced using MiSeq short-read and Oxford Nanopore long-read sequencing technology. Prophages are identified to assess genetic diversity within CRKP genomes.

Results: The investigation encompassed eight CRKP strains that collected from the patients enrolled as well as the environment, which illustrate that *bla*_{KPC-2} is responsible for phenotypic resistance in six CRKP strains that *K. pneumoniae* sequence type (ST11) is informed. The plasmid with IncR, ColRNAI and pMLST type with IncF[F33:A-B-] co-exist in all ST11 with KPC-2-producing CRKP strains. Along with carbapenemases, all *K. pneumoniae* strains harbor two or three extended spectrum β -lactamase (ESBL)-producing genes. *fosA* gene is detected amongst all the CRKP strains. The single nucleotide polymorphisms (SNP) markers are indicated and validated among all CRKP strains, providing valuable clues for distinguishing carbapenem-resistant strains from conventional *K. pneumoniae*.

Conclusions: ST11 is the main CRKP type, and *bla*_{KPC-2} is the dominant carbapenemase gene harbored by clinical CRKP isolates from current investigations. The SNP markers detected would be helpful for characterizing CRKP strain from general *K. pneumoniae*. The data provides insights into effective strategy developments for controlling CRKP and nosocomial infection reductions.

Keywords: *Klebsiella pneumoniae*, Carbapenem-resistant, Whole-genome sequencing, *bla*_{KPC-2}

Background

Antibiotic resistance is amongst the extremely severe public health challenges nowadays. Carbapenem-resistant *Enterobacteriaceae* (CRE) is reported as a consequence mainly due to acquisition of carbapenemase genes, and CRE is inferred as an urgent threat to human health by the Centers for Disease Control and Prevention (CDC), USA in 2013

[1]. Carbapenems such as imipenem, meropenem, and biapenem represent the first-line treatment of serious infections caused by multi-resistant Enterobacteriaceae including *Klebsiella pneumoniae* (*K. pneumoniae*) and *Escherichia coli* (*E. coli*) [2]. Whereas carbapenems can be hydrolyzed by carbapenemase in carbapenem-resistant *K. pneumoniae* (CRKP) [3], which results in resistance to β -Lactam antibiotics including carbapenem. Carbapenemases can be divided into Ambler class A β -lactamases (e.g. *Klebsiella pneumoniae* carbapenemases (KPC)), class B metallo- β -lactamases (MBLs), verona integrin-encoded metallo- β -lactamase (VIM), New Delhi metallo- β -lactamase (NDM) type, and Class D Enzymes of the OXA-48 type [4]. Among Ambler class A β -lactamases, plasmid-mediated KPC has

* Correspondence: zhaofang_han@foxmail.com; 15960102808@163.com

⁵State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, Fujian, People's Republic of China

¹Department of Infectious Diseases, Mengchao Hepatobiliary Hospital of Fujian Medical University, Xihong Road 312, Fuzhou 350025, Fujian, People's Republic of China

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



been identified in all gram-negative members of the ESKAPE pathogens [5], and KPC is the most clinically indispensable enzyme due to its prevalence in Enterobacteriaceae [6]. Moreover, pathogens harboring KPC-2 are resistant to all β -lactams and β -lactamase inhibitors except ceftazidime/avibactam, which extremely limit treatment options as well as lead to high mortality rates [7]. Additionally, NDM has become a serious threat to public health due to the rapid global dissemination of NDM-bearing pathogens and the presence on mobile genetic elements in an extensive series of species [8]. Consequently, it is imperative and urgent to investigate the CRKP characteristics for better controlling pathogens and diagnosing as well as treating patients.

In current investigation, seven CRKP strains are extracted from patients during their hospitalizations and another one CRKP strain is obtained from the dining car in Mengchao Hepatobiliary Hospital of Fujian Medical University (Additional file 1: Table S1). The whole genome of CRKP is sequenced using MiSeq short-reads and Oxford Nanopore long-reads sequencing technology. We conduct surveillance of the CRKP-mediated infection prevalence in Mengchao Hepatobiliary Hospital of Fujian Medical University, investigate the molecular characterization of the strains that obtained, and identify gene phenotypes as well as resistance-associated genes of the strain emergence. The detected single nucleotide

polymorphisms (SNP) markers would be helpful for recognizing CRKP strain from general *K. pneumoniae*. Data of this study provide essential insights into effective strategy developments for controlling CRKP and nosocomial infection reductions.

Results

Antimicrobial susceptibilities of the CRKP strains

The source of isolates is supplied in Table 1, which denotes the infectious type and the result of susceptibility testing during the patients' hospitalization. All eight strains involved in the study are confirmed to be *K. pneumoniae*, with five strains from sputum, one from bile, one from blood, and one from the environment (Additional file 1: Table S1). Clinical data demonstrate that seven of the eight patients are referred due to pulmonary infection, and another one is referred due to abdominal infection. The susceptibility testing data in Table 1 reveals that all the *K. pneumoniae* strains are resistant to almost all antibiotics, such as cephalosporins, penicilins, quinolones and carbapenems (imipenem with MICs ≥ 16 $\mu\text{g/ml}$). For aminoglycosides antibiotics, except that 1567D isolate is sensitive to amikacin and tobramycin, all other isolates are resistant to aminoglycosides antibiotics. The strains including 1566D, 2038D, 2039D and 2040D are resistant to sulfamethoxazole/trimethoprim with MICs ≥ 320 $\mu\text{g/ml}$, and the other strains

Table 1 Antibiotic susceptibility profiles of *K. pneumoniae*. The results of antimicrobial susceptibility testing - antibiotics MIC (mg/L) and breakpoint interpretation or epidemiological cut-off value

Isolates	1566D	1567D	2035D	2036D	2037D	2038D	2039D	2040D
source	sputum	bile	sputum	sputum	blood	sputum	sputum	environment
Infection	Pulmonary	Abdominal	Pulmonary	Pulmonary	Pulmonary	Pulmonary	Pulmonary	N/A
ampicillin	≥ 32 (R)							
ampicillin/sulbactam	≥ 32 (R)							
piperacillin/tazobactam	≥ 128 (R)							
cefazolin	≥ 64 (R)							
cefotetan	≥ 64 (R)							
ceftazidime	≥ 64 (R)							
ceftriaxone	≥ 64 (R)							
cefepime	≥ 64 (R)							
aztreonam	≥ 64 (R)							
imipenem	≥ 16 (R)							
amikacin	≥ 64 (R)	≤ 4 (S)	≥ 64 (R)					
gentamicin	≥ 16 (R)							
tobramycin	≥ 16 (R)	≤ 2 (S)	≥ 16 (R)					
ciprofloxacin	≥ 4 (R)							
levofloxacin	≥ 8 (R)							
macroclant	256(R)	256(R)	≥ 512 (R)					
sulfamethoxazole/trimethoprim	≥ 320 (R)	≤ 20 (S)	≤ 20 (S)	≤ 20 (S)	≤ 20 (S)	≥ 320 (R)	≥ 320 (R)	≥ 320 (R)

S susceptible, I intermediate, R resistant

(1567D, 2035D, 2036D, 2037D) are sensitive to sulfamethoxazole/trimethoprim with MICs ≤ 20 $\mu\text{g/ml}$.

Genome assembly and annotation

The short-read sequenced seven CRKP strains are assembled into contigs. As listed in Table 2, the assembled genome size of all strains ranged from 5.4 Mb to 5.8 Mb, with mean length of 5.7 Mb and average contigs numbering 199. The N50 length of genomes is from 176.6 kb to 251.6 kb with an average N50 length of 220.4 kb and mean GC content of 57.2%. To obtain a more complete genome, the 1567D strain is resequenced via long-read sequencing technology and assembled into three contigs with size of 5.6 Mb (Additional file 1: Figure S1). A total of 5841 protein-coding genes are predicted with length between 37 to 1649 bp (Additional file 1: Figure S2). Totals of 4657, 5097, 4714, 3179 and 3099 predicted genes are functionally annotated in NR, COG, Swiss-Prot, GO and KEGG databases, respectively (Additional file 1: Figures. S3, S4, S5).

Characteristics of the CRKP isolates

The isolated eight CRKP bacteria are sequenced through Illumina MiSeq platform and assembled into whole genomes. To understand genetic diversity, mobile genetic elements of 24 prophages are identified in eight CRKP genomes, with sizes ranging from 8.4 kb to 98.9 kb (Fig. 1). According to the criterion that the length of an intact prophage should be more than 20 kb [9]. Prophages detected in most strains (except for 2036D) are complete with a size of at least 20.2 kb with an average GC percentage of 52.7%. Additionally, three prophages are respectively identified in 3 strains at the same time, revealing the genomic sequence homology among all isolates. The 2036D strain is comprised of just one prophage probably because of the small genome size and distinct sequence characteristics, which is expected to have less neutral targets for prophage integration [9].

Furthermore, multilocus-sequence typing (MLST) analysis reveals that there are two unrelated sequence type (ST) in *K. pneumoniae* strains isolated from different patients. 2036D *K. pneumoniae* strain correlates with ST2632,

and the other six strains are relevant to ST11 (Table 3). pMLST analysis reveals that all of the six ST11 *K. pneumoniae* strains are associated with IncF[F33:A-b-] and the ST2632 *K. pneumoniae* strain is relevant to IncHI1 and IncF.

Plasmid analysis [10] shows different circular plasmids carried by the individual strains. All strains harbored IncR and ColRNAI plasmids with no virulence genes but contain several resistance-associated genes that cause resistance to carbapenems, which is demonstrated in Table 3. The IncR plasmid is identified as multidrug-resistant plasmids and has variable copy numbers of certain resistance genes among *K. pneumoniae* isolates.

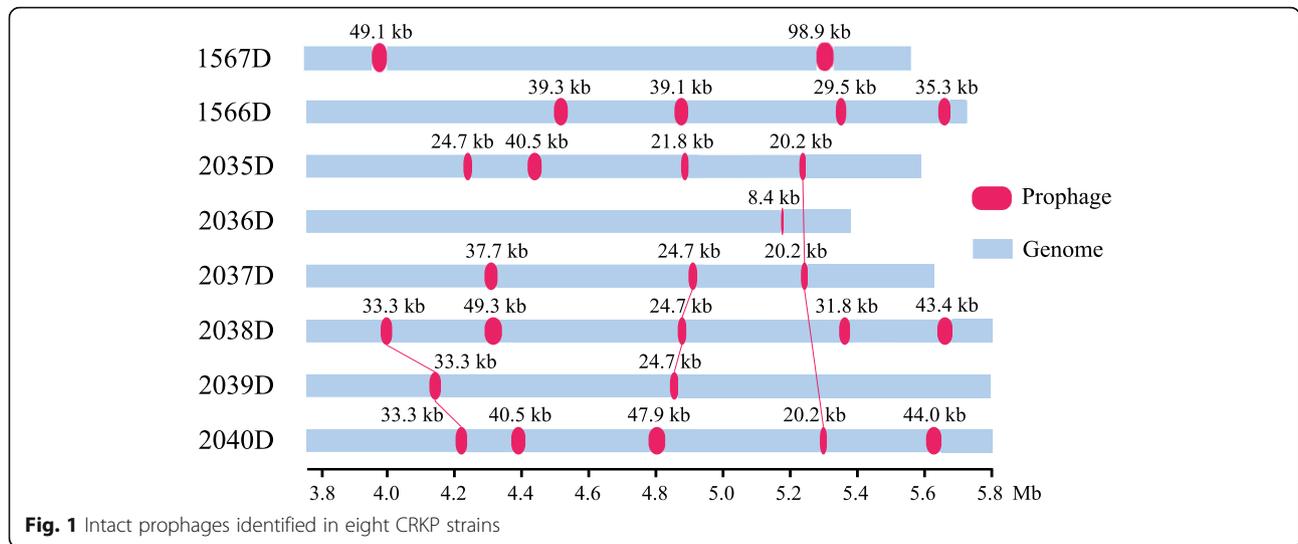
Detection of antibiotic resistance genes of CRKP isolates

The antibiotic resistance-associated genes of seven CRKP bacteria (Table 3) are sequenced on Illumina MiSeq platform among the patient and environmental isolates. As illustrated in Table 3, some antimicrobial resistance genes are mediated by plasmid such as β -lactamase correlative genes ($bla_{\text{CTX-M}}$, bla_{KPC} , bla_{LEN} , bla_{TEM}) and those genes which encoded aminoglycoside [$aac(3)\text{-IIa}$, $rmtB$], chloramphenicol ($catA1$, $catA2$), trimethoprim ($dfrA1$, $dfrA17$), and fluoroquinolone [$QnrS1$]. The other antimicrobial resistance genes are encoded by chromosome including bla_{SHV} (narrow-spectrum β -lactamase in *K. pneumoniae*), $oqxA$ (1176 bp), $oqxB$ (3153 bp) (efflux pumps), and $fosA$ (420 bp, fosfomycin resistance) genes.

Except 2036D, all the other *K. pneumoniae* strains harbor $bla_{\text{KPC-2}}$ which is associated with carbapenems resistance. Extended-spectrum β -lactamases (ESBLs) resistance genes such as $bla_{\text{CTX-M}}$, bla_{TEM} , bla_{LEN} and bla_{SHV} are also informed. bla_{TEM} is one of the genes that produce ESBL. $bla_{\text{CTX-M}}$ with different types ($bla_{\text{CTX-M-14}}$, $bla_{\text{CTX-M-3}}$, $bla_{\text{CTX-M-55}}$ and $bla_{\text{CTX-M-65}}$) is found among all the *K. pneumoniae* strains. $bla_{\text{CTX-M-3}}$ is observed in 2036D strains. $bla_{\text{CTX-M-55}}$ is observed in 1567D strain and $bla_{\text{CTX-M-14}}$ is observed in 1566D and 2040D strains. $bla_{\text{CTX-M-65}}$ is detected in the other four (2035D, 2037D, 2038D, 2039D) *K. pneumoniae* strains. bla_{LEN12} gene is exclusively found in 1566D strain, and there is no bla_{SHV} gene in it. Nevertheless, $bla_{\text{SHV-93}}$ and $bla_{\text{SHV-11}}$ genes are

Table 2 Assembly statistics of seven CRKP strains via short-read sequencing

Assembly	1566D	2035D	2036D	2037D	2038D	2039D	2040D
Contig number	195	163	215	242	184	196	195
Total length (bp)	5,758,754	5,633,502	5,432,179	5,670,795	5,833,697	5,831,354	5,835,044
Largest contig (bp)	380,781	381,132	929,110	380,795	381,132	381,056	381,056
GC (%)	57.33	57.39	57.18	57.35	57.21	57.21	57.2
N50	176,606	196,479	251,620	196,479	183,862	190,289	183,862
L50	12	11	7	11	11	11	11
Total number of Ns	20	30	20	30	130	130	30



detected in 2036D strain and the other five *K. pneumoniae* strains, respectively. Except for the 2036D strain, *bla*_{TEM-1B} gene is observed in all the other six *K. pneumoniae* strains. *Aac(3)-IId* and *rmtB* encoding fluoroquinolone resistance are observed among all strains. *oqx*A and *oqx*B with the resistance to fluoroquinolones are exclusively detected in 2036D strain. *fos*A resulting in fosfomycin resistance [11] is also informed among all CRKP strains.

Characterizing CRKP SNPs and phylogeny

The SNP markers are identified for all strains that sequenced using the short-read MiSeq data. The data demonstrate that 33,716 markers are detected in the 2036D strain, which is largely more than the other strains with an average of 8289 SNPs. The cSNPs located

in protein-coding regions are in slightly higher amounts among all detected SNPs of a minimal ratio of 85.5% (Additional file 1: Table S11). In addition, the pairwise comparison analysis reveals that 2036D isolate is disparate with the other strains based on clusters of sequence similarities using subprogram of Trinity [12] (Fig. 2a). Furthermore, the 2036D strain share few SNP loci with the others, which coincides with strain clusters (Fig. 2b).

For validations, all strains have a high detection rate in that approximately 153 out of 200 SNPs (76.4%) that have amplifications, which demonstrate the analysis accuracy (Additional file 1: Table S11). After filtering SNP loci that are not located in exome regions, containing no-alleles locus, and comprising all-wild SNP loci in each isolate, we eventually obtain 92 SNPs among 200

Table 3 Resistance genes among the patient and environmental isolates

Isolates	1566D	2035D	2036D	2037D	2038D	2039D	2040D
MLST	ST11	ST11	ST2632	ST11	ST11	ST11	ST11
pMLST	IncF[F33:A-B-]	IncF[F33:A-B-]	IncH11, IncF	IncF[F33:A-B-]	IncF[F33:A-B-]	IncF[F33:A-B-]	IncF[F33:A-B-]
Plasmids	IncR, ColRNAI	IncR, ColRNAI	IncR, ColRNAI	IncR, ColRNAI	IncR, ColRNAI	IncR, ColRNAI	IncR, ColRNAI
Penicillins: Ampicillin/ Narrow-Spectrum	<i>bla</i> CTX-M-14	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-3	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-14
Cephalosporins: cefazolin and cefotetan	<i>bla</i> KPC-2 <i>bla</i> LEN12 <i>bla</i> TEM-1B	<i>bla</i> KPC-2 <i>bla</i> SHV-11 <i>bla</i> TEM-1B	<i>bla</i> SHV-93	<i>bla</i> KPC-2 <i>bla</i> SHV-11 <i>bla</i> TEM-1B			
β-lactam inhibitors/ Carbapenems	<i>bla</i> KPC-2	<i>bla</i> KPC-2	-	<i>bla</i> KPC-2	<i>bla</i> KPC-2	<i>bla</i> KPC-2	<i>bla</i> KPC-2
Extended-Spectrum Cephalosporins/ Monobactam	<i>bla</i> CTX-M-14	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-3	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-65	<i>bla</i> CTX-M-65
Aminoglycosides	<i>rmtB</i>	<i>rmtB</i>	<i>aac(3)-IId</i>	<i>rmtB</i>	<i>aac(3)-IId</i> <i>rmtB</i>	<i>aac(3)-IId</i> <i>rmtB</i>	<i>aac(3)-IId</i> <i>rmtB</i>
Fluoroquinolones	<i>QnrS1</i>	-	<i>oqx</i> A <i>oqx</i> B	-	<i>QnrS1</i>	<i>QnrS1</i>	<i>QnrS1</i>
Phosphonic Acid	<i>fos</i> A	<i>fos</i> A	<i>fos</i> A	<i>fos</i> A	<i>fos</i> A	<i>fos</i> A	<i>fos</i> A
Phenicol	-	<i>cat</i> A2	<i>cat</i> A1	<i>cat</i> A2	<i>cat</i> A2	<i>cat</i> A2	<i>cat</i> A2
Folate-pathway Inhibitors	<i>dfr</i> A1	-	<i>dfr</i> A17	-	<i>dfr</i> A1	<i>dfr</i> A1	<i>dfr</i> A1

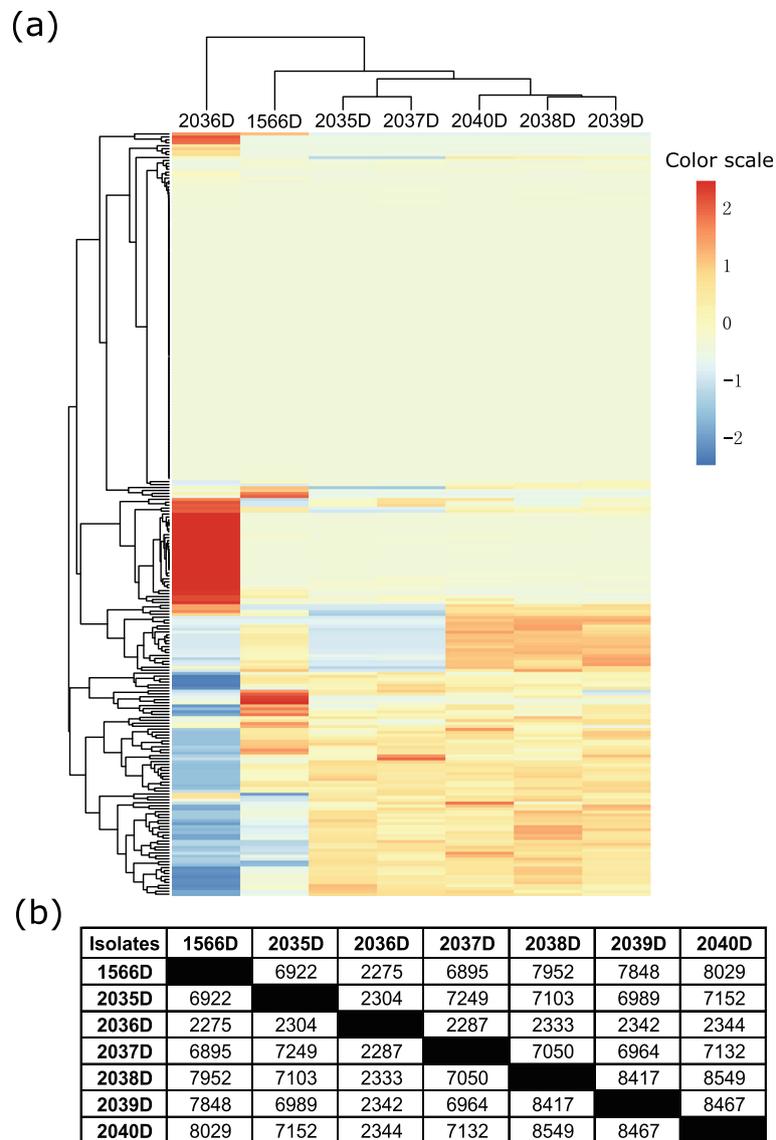


Fig. 2 Assessing the genetic relatedness of the CRKP by WGS. **a** Clustering of all isolates based on sequence similarity. **b** Communal SNP markers detected by pairwise comparison analysis

validated loci. A total of 40 out of 92 SNPs are all-variation loci in all isolates, which could be utilized for recognizing CRKP strain from ordinary *K. pneumoniae* (Additional file 1: Table S12). In addition, 24 SNPs of strain's unique loci, including strains of 2036D (18 loci), 2035D (3 loci), 1566D (2 loci) and 2037D (1 loci), would be helpful resources for specific strain identification of clinical analysis.

Previous 5 CRKP strains that isolated in Hangzhou [13] are downloaded from GenBank, and we conduct comparisons with strains in our study. The comparison result suggests that CRKP strains in Hangzhou are different from that in Fuzhou, presenting geographical difference (Additional file 1: Figure S6). The phylogenetic

tree shows that 1566D strain is most distantly related to other strains, and 2036D is more different from other strains, which is not even included in the phylogenetic tree (Additional file 1: Figure S6).

GWAS analysis

To further identify significant SNPs and genes, we perform genome-wide association study (GWAS) analysis. The patients' body temperature and counts of leukocyte are selected as phenotypic character. The short-sequencing reads of six strains (Fig. 3) are aligned to the 1567D genome using BWA v0.7.17 software. We call SNPs using Platypus v0.8.1 [14], and then filter the SNPs through plink v1.9 according to the following conditions:

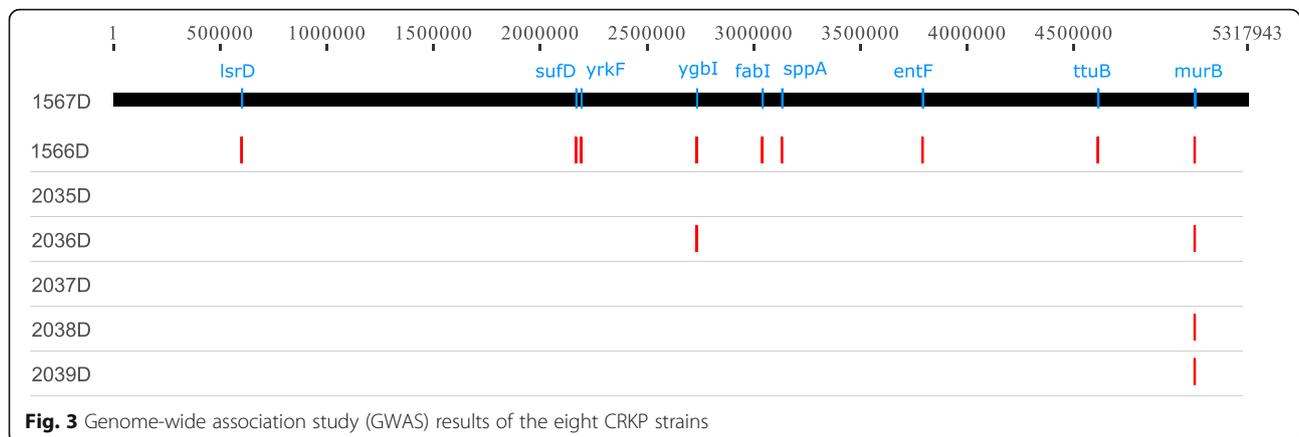


Fig. 3 Genome-wide association study (GWAS) results of the eight CRKP strains

(i) missing loci, (ii) minor allele frequency (MAF) < 0.05 and (iii) significant deviation from the Hardy-Weinberg equilibrium (HWE) ($P < 0.01$). A total of 698 SNP markers are remained and utilized for GWAS analysis. As a result, 9 loci are identified ($P < 0.05$). Two loci (*ygbI* and *murB*) are related with temperature and the other seven loci (*IsrD*, *SufD*, *yrkF*, *fabI*, *sppA*, *entF* and *ttuB*) are relevant to leukocyte (Fig. 3).

Discussion

Data of current study confirm that all CRKP strains hold two types of plasmids with no virulence gene whereas harbor an abundance of associated resistance genes such as ESBLs and carbapenemases. One genotype of carbapenemases with *bla*_{KPC-2} and two ST types with ST11 and ST2632 are identified in the study, and the ST11 with KPC-2-positive is a prevalent strain accounting in all the six strains. The plasmid with *IncR*, *ColRNAI* and *pMLST* type with *IncF*[F33:A-B-] co-exist in all ST11 with KPC-2-producing CRKP strains. The initial detection of a KPC-2-producing *K. pneumoniae* isolate from a hospital in China is reported in 2007 [15]. Since then, *bla*_{KPC-2}-bearing *K. pneumoniae* isolates have become more prevalent and reported in China as well as other countries and areas [16]. Recently, one patient is found to have susceptible *K. pneumoniae* bacteraemia in US [15]. While that case is relatively specific since the patient might be affected during the visit and hospitalization in India, which would add more complex environmental factors to confound the results. CRKP of ST11 associated with *bla*_{KPC-2} is disseminated widely across China [17, 18], which is concordant with the results of our study. These findings suggest that the CRKP-mediated infections in our hospital result from ST11 with KPC-2-positive *K. pneumoniae* isolates. Continuous monitoring will be necessary to prevent further dissemination of carbapenemase-resistance genes.

Besides carbapenemases, a variety of ESBLs such as *bla*_{CTX-M}, *bla*_{SHV}, *bla*_{LEN}, *bla*_{TEM} are present in CRKP

strains of this study. *K. pneumoniae* is one of the most indispensable infectious agents in the ICU [19]. There are “classic” and hypervirulent strains of *K. pneumoniae* [20–22]. The “classic” non-virulent strain of *K. pneumoniae* (C-KP) can produce ESBLs related to nosocomial infectious outbreaks especially in the ICU of a hospital. C-KP more easily acquires antimicrobial resistance such as ESBLs. In our investigation, *bla*_{CTX-M} with different type is found among all the CRKP strains. Chromosome-mediated *bla*_{SHV} and plasmid-mediated *bla*_{TEM} are also positive for ESBLs production and are observed in six *K. pneumoniae* strains. Co-occurrence of *bla*_{CTX-M}, *bla*_{KPC-2}, *bla*_{SHV-11} and *bla*_{TEM-1B} are observed among five *K. pneumoniae* strains. All *K. pneumoniae* strains harbor two or three ESBLs-producing genes (*bla*_{CTX-M}, *bla*_{SHV} and *bla*_{TEM}), which indicate all isolates contained multiple ESBLs resistance genes. Previous reports noted consistent results that co-occurrence of *bla*_{TEM}, *bla*_{SHV} and *bla*_{CTX-M} (any two or all three) was observed among *Klebsiella* isolates [23].

fosA is frequently identified in the *E. coli* and *K. pneumoniae* genomes [24, 25]. The *fosA5* gene is first found in *E. coli* in 2014 [26]. In 2017, it was reported that all of 73 carbapenem-resistant *K. pneumoniae* isolates were positive for *fosA5* in one Chinese area: Zhejiang Province [27]. Antimicrobial susceptibility testing about fosfomycin is not conducted in this study though *fosA* is also found among all the CRKP strains, which might indicate that fosfomycin-modifying enzymes account for a majority of the fosfomycin resistance, and that fosfomycin is resistant to CRKP strains. As reported, a clinical *Escherichia coli* strain HS102707 isolate and an *Enterobacter aerogenes* strain HS112625 isolate are resistant to carbapenem and fosfomycin and positive for the *bla*_{KPC-2} and *fosA3* genes [25], and *fosA* exists in all CRKP strains with *bla*_{KPC-2} in our study. Continuous monitoring will be necessary to prevent further dissemination of fosfomycin-resistant bacteria together with prudent use of fosfomycin in clinical settings.

OqxA and *oqxB* genes are relevant to efflux pumps, which means that antibiotics such as cephalosporins, carbapenems and fluoroquinolones are almost completely expelled from *K. pneumoniae* through its cell membrane [28]. To our knowledge, these two genes are mainly reported to be responsible for the resistance to fluoroquinolones. They do have been previously reported to be associated with the nitrofurantoin resistance.

The genome sequences of the seven strains include massive contigs which are highly fragmented. Upon further investigation, we sequence the 1567D strain using long-read sequencing platform, which could help us assemble the genome with considerable improvement in completeness and contiguity. The carbapenem-resistant genes including *fosA*, *oqxA* and *oqxB* and 40 all-variation SNP loci are also identified in the above genome demonstrating the high-quality assembly. In comparison with previous study revealing 12.3 substitutions in average [29], we identify more SNP markers in each isolate due to loose threshold. The method in Yang et al. can filter large number of SNPs with low frequency or depth and ensure the quality of SNPs, however, those isolate-specific markers might also be filtered, which would not provide many enough markers for GWAS and downstream analysis for current study. As *Klebsiella pneumoniae* is an emerging nosocomial pathogen with extended antibiotic resistance, online resources, such as BacWGSTdb [30], offering rapid typing and phylogenetic relatedness linked to antibiotic resistance genes and clinical data would be increasingly indispensable in a globalized community. The assembly and annotation information will be beneficial in understanding the whole genomic characterization of CRKP strain for future study.

Conclusions

In conclusion, ST11 is the main CRKP type, and *bla*_{KPC-2} is the dominant carbapenemase gene harbored by clinical CRKP isolates of current investigation. The plasmid with IncR, ColRNAI and pMLST type with IncF[F33:A:-B-] exist in all ST11 with KPC-2-producing CRKP strains. Besides carbapenemases, all *K. pneumoniae* strains harbor two or three ESBLs-producing genes (*bla*_{CTX-M}, *bla*_{SHV} and *bla*_{TEM}), which indicate that all isolates contain multiple ESBLs resistance genes. *fosA* genes are also found among all the CRKP strains, which may infer that fosfomycin-modifying enzymes account for a majority of the fosfomycin resistance and that CRKP strains are resistant to fosfomycin. The 40 all-variation SNP loci in all isolates could be employed and referred for distinguishing CRKP strain from ordinary *K. pneumoniae*. The detected SNP markers would be helpful for characterizing CRKP strain from general *K. pneumoniae*. This study provides insights into effective strategy developments for controlling CRKP and nosocomial infection reductions.

Methods

Patient clinical information

In total, seven patients received treatments during their hospitalizations and the data of them were completely classified and studied. One bacterium was extracted from the dining car in the hospital and since the carrier was not human, there was no clinical data relating to it. All patients, except patient 1567P that was diagnosed as abdominal infection, were diagnosed as severe pneumonia or suffered lung infections (Additional file 1: Table S1). We further give Additional file 1: Tables S2-S8 to in detail provide all patients' treatment records as well as the phenotype measurement results and data.

All patients received systematic medical examinations such as whole blood cell test, blood routine test, blood electrolyte test, blood clotting, fungal D-glucan detection, galactomannan detection, etc. All the records are archived in detail for further investigations.

Bacterial isolates, identification and antimicrobial resistance

Single patient isolates are obtained from specimens that received from inpatients admitted to Mengchao Hepatobiliary Hospital of Fujian Medical University (Fuzhou, China) in 2017. From April, 2017 to December, 2017, a total of eight CRKP isolates (Additional file 1: Table S1), which are resistance to all the antibiotics tested, such as cephalosporins, penicilins, quinolones, aminoglycosides and carbapenems (Imipenem with MICs ≥ 16 $\mu\text{g/ml}$) (Table 1), were processed following standard operating procedures: the isolates are extracted according to the aseptic operating procedures and cultured in the bacterial culture medium with Columbia Agar + 5% sheep blood. The study has been performed in accordance with the Institutional Ethical Committee of the Faculty of Medicine, Mengchao Hepatobiliary Hospital of Fujian Medical University, which approved this study (No. 2017_036_01).

K. pneumoniae isolates are confirmed by Matrix-assisted Laser Desorption Ionization-time of Flight Mass Spectrometry (MALDI-TOF-MS) (BioMerieux SA, BioMerieux Inc., France). The resistance of pathogenic bacteria is identified by Automatic Microbial Identification & Drug Sensitivity Analysis System (VITEK-2 Compact, BioMerieuxInc., France) with Gram-Negative identification card (VITEK2 AST-GN13, BioMerieuxInc., France). The results of antimicrobial susceptibility testing are interpreted based upon Clinical and Laboratory Standards Institute (CLSI) M100-S24 [31]. The standard strain under quality control is *K. pneumoniae* isolates ATCC700603 (American Type Culture Collection, ATCC).

Whole genome sequencing (WGS) and assembly

The isolated seven CRKP bacteria are sequenced on Illumina MiSeq (Illumina, San Diego, CA, USA) platform.

MiSeq short-read sequencing library is generated with 1 ng purified DNA. Inserting a phosphate to 5' UTR end and "A" to 3' UTR end produces end-repair, and PCR fragments (300 ~ 600 bp) are collected from bar-coded adapter ligation. The library is purified via AMPure XP (Beckman Coulter), which is then sequenced on MiSeq platform. In sum, a total of 40.5 million reads (2 × 300 bp) with a size of 1.36 Gb data are yielded (Additional file 1: Table S9). All short reads are first filtered for the low-quality sequences and then assembled into contigs using SPAdesv3.11.1 software [32].

Subsequently, we select an isolate of 1567D to perform long-read sequencing on Oxford Nanopore MinION (Oxford, UK) platform to easily sequence across repeat regions. The sequencing library is constructed with 1.5 μg purified DNA using the LSK-108 Oxford Nanopore Technologies (ONT) ligation protocol, and the prepared library is sequenced following the standard protocol of Oxford Nanopore MinION. A total of 7.48 Gb ultra-long reads are generated with N50 length of 25,890 bp (Additional file 1: Table S10). The long reads that 'passed' during the Nanopore base calling are used to assemble into complete genomic sequences via Canu software [33]. The long-read sequencing data of the same individual are used to correct base errors of assembled genome using Nanopolish (<https://github.com/jts/nanopolish>).

Detecting Prophages in the CRKP genomes

The putative prophages within contigs of the CRKP genomic sequences are identified using the PHAST web server (PHAge Search Tool) [34]. The prophage completeness and categorization (intact, incomplete, or questionable) are presented applying over sequences to check homology, and to detect, annotate, and graphically display prophages.

Carbapenemase-resistance gene identifications

To predict the protein-coding genes and functional proteins in the CRKP genomes, all assembled sequences are annotated by a web-based package RAST (Rapid Annotations using Subsystems Technology) [35]. The antibiotic resistance and virulence genes, plasmids, phenotyping and genotyping of CRKP genomes are scanned using the Bacterial Analysis Pipeline [36]. Carbapenemase-resistance genes are further identified from above annotated sequences according to Simner et al. [37].

The protein-coding genes of long-read assembled genome are predicted using GLIMMER (Gene Locator and Interpolated Markov ModelER) v3.02 [38]. To functionally annotate the predicted genes and perform the pathway analysis, we align them to NR, COG, Swiss-Prot, GO and KEGG databases using blastX (E-value: 10⁻⁵). The annotated genes serve to improve the completeness of some important carbapenemase-resistance genes.

Comparisons of strain similarity are performed using the Harvest Tools Suite [39] (version 1.1.2). For all of the isolates sequenced on a particular platform, parsnp is utilized to compare all the assembled isolates against each other and known reference strain. Results are visualized using EvolView.

SNP identification and validation

We download *K. pneumoniae* genome from NCBI as the reference (Accession No. PRJNA78789) to identify SNP markers [18]. All high-quality data (Q value > 20, reads length > 50 bp, number of uncertain bases < 5%) of eight CRKP strains are aligned to the reference genome sequences using BWA v0.7.17 [40], and aligned reads are sorted by coordinates via SAMTOOLS v1.4 [41]. The GATK (Genome Analysis Tool Kit) software v3.8.0 [42] is utilized to detect SNPs, which is described as following: (1) duplicated reads are removed; (2) reads around insertions/deletions are realigned; (3) base quality is recalibrated using default parameters; (4) all variants are identified using HaplotypeCaller method in GATK with emitting and calling standard confidence thresholds at 10.0 and 30.0, respectively. To validate the detected SNPs in the seven CRKPs, we select 20 loci within each sample that are located in protein-coding regions and sequence them with high read depth. All chosen markers are designed primers for amplification using Sequenom MassARRAY iPLEX platform.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-6225-9>.

Additional file 1: Figure S1. Circle diagram of *K. pneumoniae* genome sequenced via Oxford Nanopore sequencing technology. **Figure S2.** Distribution of protein-coding genes predicted in 1567D strain. **Figure S3.** COG classification 1567D strain for the carbapenem-resistant *K. pneumoniae*. **Figure S4.** Distribution of *K. pneumoniae* genes annotated in GO term. **Figure S5.** Annotation of KEGG pathways in the carbapenem-resistant *K. pneumoniae*. **Figure S6.** Phylogenetic tree assessing the relatedness of the carbapenem-resistant *K. pneumoniae* strains in Fuzhou (purple) and in Hangzhou (green) to the reference genome database (blue). **Table S1.** Information of strains and patient diagnosis. **Table S2.** Phenotypes of 1566D, a.k.a., medical records of 1566P. **Table S3.** Phenotypes of 1567D, a.k.a., medical records of 1567P. **Table S4.** Phenotypes of 2035D, a.k.a., medical records of 2035P. **Table S5.** Phenotypes of 2036D, a.k.a., medical records of 2036P. **Table S6.** Phenotypes of 2037D, a.k.a., medical records of 2037P. **Table S7.** Phenotypes of 2038D, a.k.a., medical records of 2038P. **Table S8.** Phenotypes of 2039D, a.k.a., medical records of 2039P. **Table S9.** Illumina MiSeq sequencing yields. **Table S10.** Oxford Nanopore sequencing yields. **Table S11.** Detection and validation of SNPs in seven strains. **Table S12.** A total of 92 all-variation SNPs in seven strains. Red refers to 40 all-variation loci; Bold stands for 24 strain's unique SNP loci.

Abbreviations

CDC: Centers for Disease Control and Prevention; C-KP: The "classic" non-virulent strain of *K. pneumoniae*; *Coli*: *Escherichia coli*; CRE: Carbapenem-resistant *Enterobacteriaceae*; CRKP: Carbapenem-resistant *Klebsiella pneumoniae*; *K. pneumoniae*: *Klebsiella pneumoniae*; KPC: *Klebsiella*

pneumoniae carbapenemases; MALDI-TOF: Matrix-assisted laser desorption ionization-time of flight; MBLs: Metallo- β -lactamases; NDM: New Delhi metallo- β -lactamase; SNP: Single nucleotide polymorphisms; VIM: Verona integrin-encoded metallo- β -lactamase

Acknowledgements

Authors express their gratitude's to: (1) Lijuan Liu (Department of Microbiology, Mengchao Hepatobiliary Hospital of Fujian Medical University, China) for support with bacterial isolates, and Baorong Liu (Department of ICUs, Mengchao Hepatobiliary Hospital of Fujian Medical University, China) for support with clinical data; (2) Rebecca Lahniche from the University of Pittsburgh English Language Institute for the proofreading assistance; (3) All voluntary participants donating bacterial samples; (4) Anonymous reviewers for their suggestive comments, which helped improve this presentation; (5) Associate editor for kindly dealing with our submissions.

Authors' contributions

XY, ZZ, LH and HY conceived of the method. HY supervised the study. SZ, SW and CY implemented the bacteria culture. XY, WZ and ZH performed the bioinformatics analysis. XY and ZH optimized and performed the sequencing. XY, WZ and ZH drafted the article with inputs and feedbacks from all the other authors. All authors read and approved the final manuscript.

Funding

This study is sponsored by Key Clinical Specialty Discipline Construction Program of Fuzhou, P.R. China (Grant No. 201510301), Clinical Medicine Center Construction Program of Fuzhou, Fujian, P.R.C. (Grant No. 2018080306), Health Research Innovation Team Cultivation Project of Fuzhou, P.R.C. (Grant No. 2019-S-wt4) and Key Clinical Specialty Discipline Construction Program of Fujian, P.R. China.

Availability of data and materials

The genome shotgun sequencing data and long reads of Oxford Nanopore data are deposited at NCBI/GenBank as BioProject of PRJNA506754.

Ethics approval and consent to participate

All participants provided their written informed consent and ethical approval was provided by the Institutional Ethical Committee of the Faculty of Medicine, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fujian, China (No. 2017_036_01).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Infectious Diseases, Mengchao Hepatobiliary Hospital of Fujian Medical University, Xihong Road 312, Fuzhou 350025, Fujian, People's Republic of China. ²Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ³Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, Fujian, People's Republic of China. ⁴Department of Microbiology, Mengchao Hepatobiliary Hospital of Fujian Medical University, Xihong Road 312, Fuzhou 350025, Fujian, People's Republic of China. ⁵State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, Fujian, People's Republic of China. ⁶Xiamen Cingene Science and Technology co., LTD, Xiamen 361021, Fujian, People's Republic of China.

Received: 8 August 2019 Accepted: 25 October 2019

Published online: 07 November 2019

References

1. Queenan AM, Bush K. Carbapenemases: the versatile beta-lactamases. *Clin Microbiol Rev.* 2007;20:440–58.
2. Lee C, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global dissemination of Carbapenemase-producing *Klebsiella pneumoniae*: epidemiology, genetic

context, treatment options, and detection methods. *Front Microbiol.* 2016;7: 895.

3. Xu L, Sun X, Ma X. Systematic review and meta-analysis of mortality of patients infected with carbapenem-resistant *Klebsiella pneumoniae*. *Ann Clin Microbiol Antimicrob.* 2017;16:18.
4. Nordmann P, Naas T, Poirel L. Global spread of Carbapenemase-producing Enterobacteriaceae. *Emerg Infect Dis.* 2011;17:1791–8.
5. Pendleton JN, Gorman SP, Gilmore B. Clinical relevance of the ESKAPE pathogens. *Expert Rev Anti-Infect Ther.* 2013;11:297–308.
6. Nordmann P, Cuzon G, Naas T. The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. *Lancet Infect Dis.* 2009;9:228–36.
7. Tzouveleki LS, Markogiannakis A, Psychogiou M, Tassios PT, Daikos GL. Carbapenemases in *Klebsiella pneumoniae* and other Enterobacteriaceae: an evolving crisis of global dimensions. *Clin Microbiol Rev.* 2012;25:682–707.
8. Göttig S, Hamprecht A, Christ S, Kempf VA, Wichelhaus T. Detection of NDM-7 in Germany, a new variant of the New Delhi metallo-lactamase with increased carbapenemase activity. *J Antimicrob Chemother.* 2013;68:1737–40.
9. Vale F, Nunes A, Oleastro M, Gomes JA, Sampaio D, Rocha R, Vitor J, Engstrand L, Pascoe B, Berthenet EK, Sheppard S, Hitchings M, Megraud F, Vadivelu J, Lehours P. Genomic structure and insertion sites of *Helicobacter pylori* prophages from various geographical origins. *Sci Rep.* 2017;7:42471.
10. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, Ou HY. oriFinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Res.* 2018;46(W1):W229–34.
11. Ito R, Mustapha MM, Tomich AD, Callaghan JD, McElheny CL, Mettus RT, Shanks RMQ, Sluiscremer N, Doi Y. Widespread Fosfomycin Resistance in Gram-Negative Bacteria Attributable to the Chromosomal *fosA* Gene. *Mbio.* 2017;8(4):e00749–17.
12. Grabherr M, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29:644–52.
13. Gu D, Dong N, Zheng Z, Lin D, Huang M, Wang L, Chan EW-C, Shu L, Yu J, Zhang R, et al. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. *Lancet Infect Dis.* 2018;18(1):37–46.
14. Rimmer AJ, Phan H, Mathieson I, et al. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* 2014;46:912–8.
15. Wei Z, Du X, Yu Y, Shen P, Chen Y, Li L. Plasmid-mediated KPC-2 in a *Klebsiella pneumoniae* isolate from China. *Antimicrob Agents Chemother.* 2007;51:763–5.
16. Zhang R, Liu L, Zhou H, Chan EW, Li J, Fang Y, Li Y, Liao K, Chen S. Nationwide surveillance of clinical Carbapenem-resistant Enterobacteriaceae (CRE) strains in China. *EBioMedicine.* 2017;19:98–106.
17. Liu J, Yu J, Chen F, Yu J, Simner P, Tamma P, Liu Y, Shen L. Emergence and establishment of KPC-2-producing ST11 *Klebsiella pneumoniae* in a general hospital in Shanghai, China. *Eur J Clin Microbiol Infect Dis.* 2018;37:293–9.
18. Bi D, Jiang X, Sheng ZK, Ngmenterebo D, Tai C, Wang M, Deng Z, Rajakumar K, Ou HY. Mapping the resistance-associated mobilome of a carbapenem-resistant *Klebsiella pneumoniae* strain reveals insights into factors shaping these regions and facilitates generation of a 'resistance-disarmed' model organism. *J Antimicrob Chemother.* 2015;70(10):2770–74.
19. Meropol SB, Haupt AA, Debanne SM. Incidence and outcomes of infections caused by multidrug-resistant Enterobacteriaceae in children, 2007–2015. *J Pediatric Infect Dis Soc.* 2018;7:36–45.
20. Patel PK, Russo TA, Karchmer AW. Hypervirulent *Klebsiella pneumoniae*. *Open Forum Infectious Diseases.* 2014;1.
21. Xie Y, Tian L, Li G, Qu H, Sun J, Liang W, Li X, Wang X, Deng Z, Liu J, Ou HY. Emergence of the third-generation cephalosporin-resistant hypervirulent *Klebsiella pneumoniae* due to the acquisition of a self-transferable blaDHA-1-carrying plasmid by an ST23 strain. *Virulence.* 2018;9(1):838–44.
22. Wang X, Xie Y, Li G, Liu J, Li X, Tian L, Sun J, Ou HY, Qu H. Whole-genome-sequencing characterization of bloodstream infection-causing hypervirulent *Klebsiella pneumoniae* of capsular serotype K2 and ST374. *Virulence.* 2018 Jan 1;9(1):510–21.
23. Hossain Mondal A, Siddiqui MT, Sultan I, Mohd, Rizwanul Haq Q. Prevalence and diversity of bla TEM, bla SHV and bla CTX-M variants among multidrug resistant *Klebsiella* spp. from an urban riverine environment in India. *Int J Environmental Health Res.* 2018; 6:1–13.
24. Ito R, Mustapha MD, Tomich AD, Callaghan J, McElheny CT, Mettus R, Shanks R, Sluis-Cremer N, Doi Y. Widespread Fosfomycin Resistance in

- Gram-Negative Bacteria Attributable to the Chromosomal fosA Gene. *mBio*. 2017;8:e00749–17.
25. Li G, Zhang Y, Bi D, Shen P, Ai F, Liu H, Tian Y, Ma Y, Wang B, Rajakumar K, Ou HY, Jiang X. First report of a clinical, multidrug-resistant Enterobacteriaceae isolate coharboring fosfomycin resistance gene fosA3 and carbapenemase gene blaKPC-2 on the same transposon, Tn1721. *Antimicrob Agents Chemother*. 2015;59(1):338–43.
 26. Ma Y, Xu X, Guo Q, Wang P, Wang W, Wang M. Characterization of fosA5, a new plasmid-mediated fosfomycin resistance gene in *Escherichia coli*. *Letts Appl Microbiol*. 2015;60:259–64.
 27. Huang L, Yan Hu Y, Zhang R. Prevalence of fosfomycin resistance and plasmid-mediated fosfomycin-modifying enzymes among carbapenem-resistant Enterobacteriaceae in Zhejiang. *China Journal of Medical Microbiology*. 2017;66:1332–4.
 28. Zheng J-X, Lin Z-W, Sun X, Lin W-H, Chen Z, Wu Y, Qi G-B, Deng Q-W, Qu D, Yu Z-J. Overexpression of OqxAB and MacAB efflux pumps contributes to eravacycline resistance and heteroresistance in clinical isolates of *Klebsiella pneumoniae*. *Emerg Microbes Infect*. 2018;7:139.
 29. Yang S, Hemarajata P, Hindler J, Li F, Adisetiyo H, Aldrovandi G, Sebra R, Kasarskis A, MacCannell D, Didelot X, et al. Evolution and transmission of Carbapenem-resistant *Klebsiella pneumoniae* expressing the blaOXA-232 gene during an institutional outbreak associated with endoscopic retrograde Cholangiopancreatography. *Clin Infect Dis*. 2017;64(7):894–901.
 30. Ruan Z, Feng Y. BacWGSTdb, a database for genotyping and source tracking bacterial pathogens. *Nucleic Acids Res*. 2016;44(D1):D682–7.
 31. Clinical and Laboratory Standards Institute C. Performance standards for antimicrobial susceptibility testing; twenty-fourth informational supplement. 2014;34:51–59.
 32. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77.
 33. Koren S, Walenz B, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 2017;27:722–36.
 34. Zhou Y, Liang Y, Lynch KH, Dennis J, Wishart DS. PHAST: a fast Phage search tool. *Nucleic Acids Res*. 2011;39:347–52.
 35. Brettin T, Davis JJ, Disz T, Edwards R, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD. RASTk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep*. 2015;5:8365.
 36. Thomsen MCF, Ahrenfeldt J, Cisneros J, Jurtz VI, Larsen MV, Hasman H, Aarestrup FM, Lund O. A bacterial analysis platform: an integrated system for Analysing bacterial whole genome sequencing data for clinical diagnostics and surveillance. *PLoS One*. 2016;11.
 37. Simner PJ, Antar AAR, Hao S, Gurtowski J, Tamma PD, Rock C, Opene BNA, Tekle T, Carroll KC, Schatz MC. Antibiotic pressure on the acquisition and loss of antibiotic resistance genes in *Klebsiella pneumoniae*. *J Antimicrob Chemother*. 2018;73:1796–803.
 38. Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Res*. 2012;40.
 39. Treangen TJ, Ondov BD, Koren S, et al. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol*. 2014;15:524.
 40. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25:1754–60.
 41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth GT, Abecasis GR, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.
 42. Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky AM, Garimella K, Altshuler D, Gabriel SB, Daly MJ. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–303.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

