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

Xiaoling Yu[1], Wen Zhang[2], Zhiping Zhao[1], Chengsong Ye[3], Shuyan Zhou[4], Shaogui Wu[4], Lifen Han[1], Zhaofang Han[5,6*] and Hanhui Ye[1*]

[1] Department of Infectious Diseases, Mengchao Hepatobiliary Hospital of Fujian Medical University, Xihong Road 312, Fuzhou 350025, Fujian, P.R. China.

[2] Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.

[3] Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, Fujian, P.R. China.

[4] Department of Microbiology, Mengchao Hepatobiliary Hospital of Fujian Medical University, Xihong Road 312, Fuzhou 350025, Fujian, P.R. China.

[5] State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, 361102, Fujian, PR China.

[6] Xiamen Cingene Science and Technology co., LTD, Xiamen 361021, Fujian, PR China.

Corresponding authors:

Zhaofang Han, E-mail: zhaofang_han@foxmail.com; Hanhui Ye, E-mail: 15960102808@163.com .

Abstract

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. The investigation encompassed eight CRKP strains that collected from the patients enrolled as well as the environment, which illustrate that $bla_{\text{KPC-2}}$ 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*. In conclusion, ST11 is the main CRKP type, and $bla_{\text{KPC-2}}$ is the dominant carbapenemase gene harbored by clinical CRKP isolates from current investigations.

Keywords: *Klebsiella pneumoniae*; Carbapenem-resistant; Prophage; whole-genome sequencing

38 **Introduction**

39 Antibiotic resistance is amongst the extremely severe public health challenges nowadays.
40 Carbapenem-resistant *Enterobacteriaceae* (CRE) is reported as a consequence mainly due to
41 acquisition of carbapenemase genes, and CRE is inferred as an urgent threat to human health by
42 the Centers for Disease Control and Prevention (CDC), USA in 2013[1]. Carbapenems such as
43 imipenem, meropenem, and biapenem represent the first-line treatment of serious infections
44 caused by multi-resistant Enterobacteriaceae including *Klebsiella pneumoniae* (*K. pneumoniae*)
45 and *Escherichia coli* (*E. coli*)[2]. Whereas carbapenems can be hydrolyzed by carbapenemase in
46 carbapenem-resistant *K. pneumoniae* (CRKP)[3], which results in resistance to β-Lactam
47 antibiotics including carbapenem. Carbapenemases can be divided into Ambler class A β-
48 lactamases (e.g. *Klebsiella pneumoniae* carbapenemases (KPC)), class B metallo-β-lactamases
49 (MBLs), verona integrin-encoded metallo-β-lactamase (VIM), New Delhi metallo-β-lactamase
50 (NDM) type, and Class D Enzymes of the OXA-48 type[4]. Among Ambler class A β-lactamases,
51 plasmid-mediated KPC has been identified in all gram-negative members of the ESKAPE
52 pathogens[5], and KPC is the most clinically indispensable enzyme due to its prevalence in
53 Enterobacteriaceae[6]. Moreover, pathogens harboring KPC-2 are resistant to all β-lactams and β-
54 lactamase inhibitors except ceftazidime/avibactam, which extremely limit treatment options as
55 well as lead to high mortality rates[7]. Additionally, NDM has become a serious threat to public
56 health due to the rapid global dissemination of NDM-bearing pathogens and the presence on
57 mobile genetic elements in an extensive series of species[8]. Consequently, it is imperative and
58 urgent to investigate the CRKP characteristics for better controlling pathogens and diagnosing as
59 well as treating patients.
60

61 In current investigation, seven CRKP strains are extracted from patients during their
62 hospitalizations and another one CRKP strain is obtained from the dining car in Mengchao
63 Hepatobiliary Hospital of Fujian Medical University (**Supplementary Table 1**). The whole
64 genome of CRKP is sequenced using MiSeq short-reads and Oxford Nanopore long-reads
65 sequencing technology. We conduct surveillance of the CRKP-mediated infection prevalence in
66 Mengchao Hepatobiliary Hospital of Fujian Medical University, investigate the molecular
67 epidemiological characteristics of the strains that obtained, and identify gene phenotypes as well
68 as resistance-associated genes of the strain emergence. The detected single nucleotide
69 polymorphisms (SNP) markers would be helpful for recognizing CRKP strain from general *K.*
70 *pneumoniae*. Data of this study provide essential insights into effective strategy developments
71 for controlling CRKP and nosocomial infection reductions.

72 **Materials & Methods**

73 **Patient clinical information**

74 In total, seven patients received treatments during their hospitalizations and the data of them
75 were completely classified and studied. One bacterium was extracted from the dining car in the
76 hospital and since the carrier was not human, there was no clinical data relating to it. All patients,
77 except patient 1567P that was diagnosed as abdominal infection, were diagnosed as severe
78 pneumonia or sufferred lung infections (**Supplementary Table 1**). We further give
79 **Supplementary Table 2-8** to in detail provide all patients' treatment records as well as the
80 phenotype measurement results and data.

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

84 **Bacterial Isolates and Antimicrobial Resistance**

85 Single patient isolates are obtained from specimens that received from inpatients admitted to
86 Mengchao Hepatobiliary Hospital of Fujian Medical University in 2017. A total of eight CRKP
87 isolates (**Supplementary Table 1**), which are resistance to all the antibiotics tested, such as
88 cephalosporins, penicilins, quinolones, aminoglycosides and carbapenems (Imipenem with MICs
89 ≥16 μg/ml) (Table 1), were processed following standard operating procedures: the isolates are
90 extracted according to the aseptic operating procedures and cultured in the bacterial culture
91 medium with Columbia Agar + 5% sheep blood. The study has been performed in accordance
92 with the Institutional Ethical Committee of the Faculty of Medicine, Mengchao Hepatobiliary
93 Hospital of Fujian Medical University, which approved this study.
94
95 *K. pneumoniae* isolates are confirmed by Matrix-assisted Laser Desorption Ionization-time of
96 Flight Mass Spectrometry (MALDI-TOF-MS) mass spectrometry (BioMerieux SA, BioMerieux
97 Inc., France). The resistance of pathogenic bacteria is identified by Automatic Microbial
98 Identification & Drug Sensitivity Analysis System (VITEK-2 Compact, BioMerieuxInc., France)
99 with Gram-Negative identification card (VITEK2 AST-GN13, BioMerieuxInc., France). The
100 results of antimicrobial susceptibility testing are interpreted based upon Clinical and Laboratory
101 Standards Institute (CLSI) M100-S24[9]. The standard strain under quality control is *K.*
102 *pneumoniae* isolates ATCC700603 (American Type Culture Collection, ATCC).
103

104 **Whole Genome Sequencing (WGS) and Assembly**

105 The isolated seven CRKP bacteria are sequenced on Illumina MiSeq (Illumina, San Diego, CA,
106 USA) platform. MiSeq short-read sequencing library is generated with 1 ng purified DNA.
107 Inserting a phosphate to 5' UTR end and "A" to 3' UTR end produces end-repair, and PCR
108 fragments (300 ~ 600 bp) are collected from bar-coded adapter ligation. The library is purified
109 via AMPure XP (Beckman Coulter) , which is then sequenced on MiSeq platform. In sum, a total
110 of 40.5 million reads (2 × 300 bp) with a size of 1.36 Gb data are yielded (**Supplementary**
111 **Table 9**). All short reads are first filtered for the low-quality sequences and then assembled into
112 contigs using SPAdesv3.11.1 software[10].

113

114 Subsequently, we select an isolate of 1567D to perform long-read sequencing on Oxford
115 Nanopore MinION (Oxford, UK) platform to easily sequence across repeat regions. The
116 sequencing library is constructed with 1.5 μg purified DNA using the LSK-108 Oxford
117 Nanopore Technologies (ONT) ligation protocol, and the prepared library is sequenced following
118 the standard protocol of Oxford Nanopore MinION. A total of 7.48 Gb ultra-long reads are
119 generated with N50 length of 25,890 bp (**Supplementary Table 10**). The long reads that
120 'passed' during the Nanopore base calling are used to assemble into complete genomic
121 sequences via Canu software[11]. The long-read sequencing data of the same individual are used to
122 correct base errors of assembled genome using Nanopolish (https://github.com/jts/nanopolish).

123 **Detecting Prophages in the CRKP Genomes**

124 The putative prophages within contigs of the CRKP genomic sequences are identified using the
125 PHAST web server (PHAge Search Tool)[12]. The prophage completeness and categorization

126 (intact, incomplete, or questionable) are presented applying over sequences to check homology,
127 and to detect, annotate, and graphically display prophages.

128 **Carbapenemase-resistance Gene Identifications**

129 To predict the protein-coding genes and functional proteins in the CRKP genomes, all assembled
130 sequences are annotated by a web-based package RAST (Rapid Annotations using Subsystems
131 Technology)[13]. The antibiotic resistance and virulence genes, plasmids, phenotyping and
132 genotyping of CRKP genomes are scanned using the Bacterial Analysis Pipeline[14].
133 Carbapenemase-resistance genes are further identified from above annotated sequences
134 according to Simner et al.[15].

135 The protein-coding genes of long-read assembled genome are predicted using GLIMMER (Gene
136 Locator and Interpolated Markov ModelER) v3.02[16]. To functionally annotate the predicted
137 genes and perform the pathway analysis, we align them to NR, COG, Swiss-Prot, GO and KEGG
138 databases using blastX (E-value: 10^{-5}). The annotated genes serve to improve the completeness
139 of some important carbapenemase-resistance genes.

140 Comparisons of strain similarity are performed using the Harvest Tools Suite[17] (version 1.1.2).
141 For all of the isolates sequenced on a particular platform, parsnp is utilized to compare all the
142 assembled isolates against each other and known reference strain. Results are visualized using
143 EvolView.

144 **SNP Identification and Validation**

145 We download *K. pneumoniae* genome from NCBI as the reference to identify SNP markers[18].
146 All high-quality data (Q value >20, reads length > 50 bp, number of uncertain bases < 5%) of
147 eight CRKP strains are aligned to the reference genome sequences using BWA v0.7.17[19], and
148 aligned reads are sorted by coordinates via SAMTOOLS v1.4[20]. The GATK (Genome Analysis

149 Tool Kit) software v3.8.0[21] is utilized to detect SNPs, which is described as following: (1)
150 duplicated reads are removed; (2) reads around insertions/deletions are realigned; (3) base
151 quality is recalibrated using default parameters; (4) all variants are identified using
152 HaplotypeCaller method in GATK with emitting and calling standard confidence thresholds at
153 10.0 and 30.0, respectively. To validate the detected SNPs in the seven CRKPs, we select 20 loci
154 within each sample that are located in exonic regions and sequence them with high read depth.
155 All chosen markers are designed primers for amplification using Sequenom MassARRAY
156 iPLEX platform.

157 **Results**

158 **Antimicrobial Susceptibilities of the CRKP Strains**

159 The source of isolates is supplied in **Table 1**, which denotes the infectious type and the result of
160 susceptibility testing during the patients' hospitalization. All eight strains involved in the study
161 are confirmed to be *K. pneumoniae*, with five strains from sputum, one from bile, one from blood,
162 and one from the environment (**Supplementary Table 1**). Clinical data demonstrate that seven
163 of the eight patients are referred due to pulmonary infection, and another one is referred due to
164 abdominal infection. The susceptibility testing data in **Table 1** reveals that all the *K. pneumoniae*
165 strains are resistant to almost all antibiotics, such as cephalosporins, penicilins, quinolones and
166 carbapenems (imipenem with MICs ≥16 μg/ml). For aminoglycosides antibiotics, except that
167 1567D isolate is sensitive to amikacin and tobramycin, all other isolates are resistant to
168 aminoglycosides antibiotics. The strains including 1566D, 2038D, 2039D and 2040D are
169 resistant to sulfamethoxazole/trimethoprim with MICs ≥320, and the other strains (1567D,
170 2035D, 2036D, 2037D) are sensitive to sulfamethoxazole/trimethoprim with MICs ≤ 20.

171 **Genome assembly and annotation**

172 The short-read sequenced seven CRKP strains are assembled into contigs. As listed in **Table 2**,
173 the assembled genome size of all trains ranged from 5.4 Mb to 5.8 Mb, with mean length of 5.7
174 Mb and average contigs numbering 199. The N50 length of genomes is from 176.6 kb to 251.6
175 kb with an average N50 length of 220.4 kb and mean GC content of 57.2%. To obtain a more
176 complete genome, the 1567D strain is resequenced via long-read sequencing technology and
177 assembled into three contigs with size of 5.6 Mb (**Supplementary Figure 1**). A total of 5,841
178 protein-coding genes are predicted with length between 37 to 1,649 bp (**Supplementary Figure**
179 **2**). Totals of 4,657, 5,097, 4,714, 3,179 and 3,099 predicted genes are functionally annotated in
180 NR, COG, Swiss-Prot, GO and KEGG databases, respectively (**Supplementary Figures 3, 4, 5**).

181 **Characteristics of the CRKP Isolates**

182 The isolated seven CRKP bacteria are sequenced through Illumina MiSeq platform and
183 assembled into whole genomes. To understand genetic diversity, mobile genetic elements of 24
184 prophages are identified in seven CRKP genomes, with sizes ranging from 8.4 kb to 49.3 kb
185 (**Figure 1**). According to the criterion that the length of an intact prophage should be more than
186 20 kb[22]. Prophages detected in most strains (except for 2036D) are complete with a size of at
187 least 20.2 kb with an average GC percentage of 52.7%. Additionally, three prophages are
188 respectively identified in 3 strains at the same time, revealing the genomic sequence homology
189 among all isolates. The 2036D strain is comprised of just one prophage probably because of the
190 small genome size and distinct sequence characteristics, which is expected to have less neutral
191 targets for prophage integration[22].

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

194 *pneumoniae* strain correlates with ST2632, and the other six strains are relevant to ST11 (**Table**
195 **3**). pMLST analysis reveals that all of the six ST11 *K. pneumoniae* strains are associated with
196 IncF[F33:A-:b-] and the ST2632 *K. pneumoniae* strain is relevant to IncHI1 and IncF. Majority
197 of *K. pneumoniae* strains are ST11, with IncF [F33:A-:b-] type.
198 Plasmid analysis[23] shows different circular plasmids carried by the individual strains. All strains
199 harbored IncR and ColRNAI plasmids with no virulence genes but contain several resistance-
200 associated genes that cause resistance to carbapenems, which is demonstrated in **Table 3**. The
201 IncR plasmid is identified as multidrug-resistant plasmids and has variable copy numbers of
202 certain resistance genes among *K. pneumoniae* isolates.

203 **Detection of Antibiotic Resistance Genes of CRKP Isolates**

204 The resistance-associated genes of seven CRKP bacteria (**Table 3**) are sequenced on Illumina
205 MiSeq platform among the patient and environmental isolates. As illustrated in **Table 3**, some
206 antimicrobial resistance genes are mediated by plasmid such as β-lactamase correlative genes
207 ($bla_{\text{CTX-M}}$, bla_{KPC}, bla_{LEN}, bla_{TEM}) and those genes which encoded aminoglycoside [*aac(3)-IId,*
208 *rmtB*], chloramphenicol (*catA1, catA2*), trimethoprim (*dfrA1,dfrA17*), and fluoroquinolone
209 [*QnrS1*]. The other antimicrobial resistance genes are encoded by chromosome including bla_{SHV}
210 (narrow-spectrum β-lactamasein *K. pneumoniae*), *oqxA* (1,176 bp), *oqxB* (3,153 bp) (efflux
211 pumps), and *fosA* (420 bp, fosfomycin resistance) genes.
212 Except 2036D, all the other *K. pneumoniae* strains harbor the associated carbapenemases-
213 producing resistance gene, $bla_{\text{KPC-2}}$. Extended-spectrum β-lactamases (ESBLs) resistance genes
214 such as $bla_{\text{CTX-M}}$, bla_{TEM}, bla_{LEN} and bla_{SHV} are also informed. bla_{TEM} is one of the genes that
215 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 $bla_{\text{CTX-M-65}}$)
216 is found among all the *K. pneumoniae* strains. $bla_{\text{CTX-M-3}}$ is observed in 2036D strains. $bla_{\text{CTX-M-55}}$

217 is observed in 1567D strain and *bla*$_{CTX-M-14}$ is observed in 1566D and 2040D strains. *bla*$_{CTX-M-65}$
218 is detected in the other four (2035D, 2037D, 2038D, 2039D) *K. pneumoniae* strains. *bla*$_{LEN12}$
219 gene is exclusively found in 1566D strain, and there is no *bla*$_{SHV}$ gene in it. Nevertheless, *bla*$_{SHV-}$
220 $_{93}$ and *bla*$_{SHV-11}$ genes are detected in 2036D strain and the other five *K. pneumoniae* strains,
221 respectively. Except for the 2036D strain, *bla*$_{TEM-1B}$ gene is observed in all the other six *K.*
222 *pneumoniae* strains. *Aac(3)-IId* and *rmtB* encoding fluoroquinolone resistance are observed
223 among all strains. *fosA* resulting in fosfomycin resistance[24] is also informed among all strains.
224
225 **Characterizing CRKP SNPs**
226 The SNP markers are identified for all strains that sequenced using the short-read MiSeq data.
227 The data demonstrate that 33,716 markers are detected in the 2036D strain, which is more
228 significant than the other strains with an average of 8,289 SNPs. The cSNPs located in exonic
229 regions are in slightly higher amounts among all detected SNPs of a minimal ratio of 85.5%
230 (**Supplementary Table 11**). In addition, the pairwise comparison analysis reveals that 2036D
231 isolate is disparate with the other strains based on clusters of sequence similarities using
232 subprogram of Trinity[25] (**Figure 2a**). Furthermore, the 2036D strain share few SNP loci with the
233 others, which coincides with strain clusters (**Figure 2b**).
234
235 For validations, all strains have a high detection rate in that approximately 153 out of 200 SNPs
236 (76.4%) that have amplifications, which demonstrate the analysis accuracy (**Supplementary**
237 **Table 11**). After filtering SNP loci that are not located in exome regions, containing no-alleles
238 locus, and comprise all-wild SNP loci in each isolate, we eventually obtain 92 SNPs among 200
239 validated loci. A total of 40 out of 92 SNPs are all-variation loci in all isolates, which could be

240 utilized for recognizing CRKP strain from ordinary *K. pneumoniae* (**Supplementary Table 12**).
241 In addition, 24 SNPs of strain's unique loci, including strains of 2036D (18 loci), 2035D (3 loci),
242 1566D (2 loci) and 2037D (1 loci), would be helpful resources for specific strain identification of
243 clinical analysis.
244
245 **Phylogeny analysis results**
246 The phylogenic tree shows that 1567D strain is most distantly related to the other strains, and
247 2036D is more closely related to the reference strain comparing with all the other strains
248 (**Supplementary Figure 6**).
249
250 **GWAS analysis**
251 To further identify significant SNPs and genes, we perform genome-wide association study
252 (GWAS) analysis. The patients' body temperature and counts of leukocyte are selected as
253 phenotypic character. The short-sequencing reads of six strains (**Figure 3**) are aligned to the
254 1567D genome using BWA v0.7.17 software. We call SNPs using Platypus v0.8.1[26], and then
255 filter the SNPs through plink v1.9 according to the following conditions: (i) missing loci, (ii)
256 minor allele frequency (MAF) < 0.05 and (iii) significant deviation from the Hardy-Weinberg
257 equilibrium (HWE) ($P < 0.01$). A total of 698 SNP markers are remained and utilized for GWAS
258 analysis. As a result, 9 loci are identified ($P < 0.05$). Two loci (ygbI and murB) are related with
259 temperature and the other seven loci (IsrD, SufD, yrkF, fabI, sppA, entF and ttuB) are relevant to
260 leukocyte (**Figure 3**).

261 **Discussion**

262 Data of current study confirm that all CRKP strains hold two types of plasmids with no virulence
263 gene whereas harbor an abundance of associated resistance genes such as ESBLs and
264 carbapenemases. One genotype of carbapenemases with *bla*$_{KPC-2}$ and two ST types with ST11
265 and ST2632 are identified in the study, and the ST11 with KPC-2-positive is a prevalent strain
266 accounting in all the six strains. The plasmid with IncR, ColRNAI and pMLST type with
267 IncF[F33:A-:B-] co-exist in all ST11 with KPC-2-producing CRKP strains. The initial detection
268 of a KPC-2-producing *K. pneumoniae* isolate from a hospital in China is reported in 2007[27].
269 Since then, *bla*$_{KPC-2}$-bearing *K. pneumoniae* isolates have become more prevalent and reported in
270 China as well as other countries and areas [28]. Recently, one patient is found to have susceptible *K.*
271 *pneumoniae* bacteraemia in US[15]. While that case is relatively specific since the patient might be
272 affected during the visit and hospitalization in India, which would add more complex
273 environmental factors to confound the results. CRKP of ST11 associated with *bla*$_{KPC-2}$ is
274 disseminated widely across China[18,29], which is concordant with the results of our study. These
275 findings all suggest that the CRKP-mediated infections in our hospital result from ST11 with
276 KPC-2-positive *K. pneumoniae* isolates. Continuous monitoring will be necessary to prevent
277 further dissemination of carbapenemase-resistance genes.
278 Besides carbapenemases, a variety of ESBLs such as *bla*$_{CTX-M}$, *bla*$_{SHV}$, *bla*$_{LEN}$, *bla*$_{TEM}$ are present
279 in CRKP strains of this study. *K. pneumoniae* is one of the most indispensable infectious agents
280 in the ICU[30]. There are "classic" and hypervirulent strains of *K. pneumoniae*[31,32,33]. The "classic"
281 non-virulent strain of *K. pneumoniae* (C-KP) can produce ESBLs related to nosocomial
282 infectious outbreaks especially in the ICU of a hospital. C-KP more easily acquires antimicrobial
283 resistance such as ESBLs. *bla*$_{CTX-M}$ with different type is found among all the CRKP strains.
284 Chromosome-mediated *bla*$_{SHV}$ and plasmid-mediated *bla*$_{TEM}$ are also positive for ESBLs

285 production and are observed in six *K. pneumoniae* strains. Co-occurrence of bla_{CTX-M}, bla_{KPC-2},
286 bla_{SHV-11} and bla_{TEM-1B} are observed among five *K. pneumoniae* strains. All *K. pneumoniae*
287 strains harbor two or three ESBLs-producing genes (bla_{CTX-M}, bla_{SHV} and bla_{TEM}), which indicate
288 all isolates contained multiple ESBLs resistance genes. Previous reports noted consistent results
289 that co-occurrence of bla_{TEM}, bla_{SHV} and bla_{CTX-M} (any two or all three) was observed among
290 *Klebsiella* isolates[34].
291
292 *fosA* is frequently identified in the *E. coli* and *K. pneumoniae* genomes [35,36]. The *fosA5* gene is
293 first found in *E. coli* in 2014[37]. In 2017, it was reported that all of 73 carbapenem-resistant *K.*
294 *pneumoniae* isolates were positive for *fosA5* in one Chinese area: Zhejiang Province[38].
295 Antimicrobial susceptibility testing about fosfomycin is not conducted in this study though *fosA*
296 is also found among all the CRKP strains, which might indicate that fosfomycin-modifying
297 enzymes account for a majority of the fosfomycin resistance, and that fosfomycin is resistant to
298 CRKP strains. It is reported that *fosA* gene is transferred from *E. coli* to *K. pneumoniae* through
299 whole plasmid transmission or mobile genetic element transmission, which raise doubts whether
300 fosfomycin can be used as a supplementary drug for urinary tract infection caused by
301 carbapenem-resistant *E. coli* in the hospital, as *fosA* exists in all CRKP strains from our study.
302 Continuous monitoring will be necessary to prevent further dissemination of fosfomycin-
303 resistant bacteria together with prudent use of fosfomycin in clinical settings.
304
305 *OqxA* and *oqxB* genes are relevant to efflux pumps, which means that antibiotics such as
306 cephalosporins, carbapenems and fluoroquinolones are almost completely expelled from *K.*
307 *pneumoniae* through its cell membrane[39]. Although there are no carbapenemases that observed in

308 2036D strain, *oqxA* and *oqxB* genes are identified in it. To our knowledge, these two genes are
309 mainly reported to be responsible for the resistance to fluoroquinolones. They do have been
310 previously reported to be associated with the nitrofurantoin resistance.
311
312 The genome sequences of the seven strains include massive contigs which are highly fragmented.
313 Upon further investigation, we sequence the 1567D strain using long-read sequencing platform,
314 which could help us assemble the genome with considerable improvement in completeness and
315 contiguity. The carbapenem-resistant genes including *fosA*, *oqxA* and *oqxB* and 40 all-variation
316 SNP loci are also identified in the above genome demonstrating the high-quality assembly. The
317 assembly and annotation information will be beneficial in understanding the whole genomic
318 characterization of CRKP strain for future study.

319 **Conclusions**

320 In conclusion, ST11 is the main CRKP type, and $bla_{\text{KPC-2}}$ is the dominant carbapenemase gene
321 harbored by clinical CRKP isolates from present study. The plasmid with IncR, ColRNAI and
322 pMLST type with IncF[F33:A-:B-] exist in all ST11 with KPC-2-producing CRKP strains.
323 Besides carbapenemases, all *K. pneumoniae* strains harbor two or three ESBLs-producing genes
324 ($bla_{\text{CTX-M}}$, bla_{SHV} and bla_{TEM}), which indicate all isolates contain multiple ESBLs resistance
325 genes. *fosA* genes are also found among all the CRKP strains, which may infer that fosfomycin-
326 modifying enzymes account for a majority of the fosfomycin resistance and that CRKP strains
327 are resistant to fosfomycin. The differential expressions of *oqxA* and *oqxB* in CRKP strain might
328 possibly result in carbapenem-resistant, but this presumption needs more solid experimental
329 evidences. The 40 all-variation SNP loci in all isolates could be employed and referred for
330 distinguishing CRKP strain from ordinary *K. pneumoniae*.

331 **Declarations**

332 **Ethics approval**

333 The study has been performed in accordance with the Institutional Ethical Committee of the
334 Faculty of Medicine, Mengchao Hepatobiliary Hospital of Fujian Medical University.

335 **Consent to participate**

336 Not applicable.

337 **Consent to publish**

338 All authors read and approved the final manuscript.

339 **Availability of data and materials**

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

342 **Competing interests**

343 The authors declare that they have no competing interests.

344 **Funding**

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

350 **Authors' Contributions**

351 X.Y., Z.Z., L.H. and H.Y. conceived of the method. H.Y. supervised the study. S.Z., S.W. and
352 C.Y. implemented the bacteria culture. X.Y., W.Z. and Z.H. performed the bioinformatics
353 analysis. X.Y. and Z.H. optimized and performed the sequencing. X.Y., W.Z. and Z.H. drafted

the article with inputs and feedbacks from all the other authors. All authors read and approved the final manuscript.

Acknowledgements

Authors express their gratitude's to Rebecca Lahniche from the University of Pittsburgh English Language Institute for the proofreading assistance.

361 **References**

362 1. Queenan AM, Bush K. 2007. Carbapenemases: the versatile beta-lactamases. Clinical microbiology
363 reviews 20:440-458.
364 2. Lee C, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. 2016. Global Dissemination of
365 Carbapenemase-Producing Klebsiella pneumoniae: Epidemiology, Genetic Context, Treatment
366 Options, and Detection Methods. Frontiers in Microbiology 7:895.
367 3. Xu L, Sun X, Ma X. 2017. Systematic review and meta-analysis of mortality of patients infected with
368 carbapenem-resistant Klebsiella pneumoniae. Annals of Clinical Microbiology and Antimicrobials
369 16:18.
370 4. Nordmann P, Naas T, Poirel L. 2011. Global Spread of Carbapenemase-producing Enterobacteriaceae.
371 Emerging Infectious Diseases 17:1791-1798.
372 5. Pendleton JN, Gorman SP, Gilmore B. 2013. Clinical relevance of the ESKAPE pathogens. Expert
373 Review of Anti-infective Therapy 11:297-308.
374 6. Nordmann P, Cuzon G, Naas T. 2009. The real threat of Klebsiella pneumoniae carbapenemase-
375 producing bacteria. Lancet Infectious Diseases 9:228-236.
376 7. Tzouvelekis LS, Markogiannakis A, Psichogiou M, Tassios PT, Daikos GL. 2012. Carbapenemases in
377 Klebsiella pneumoniae and Other Enterobacteriaceae: an Evolving Crisis of Global Dimensions.
378 Clinical Microbiology Reviews 25:682-707.
379 8. Göttig S, Hamprecht A, Christ S, Kempf V, A Wichelhaus T. 2013. Detection of NDM-7 in Germany,
380 a new variant of the New Delhi metallo- -lactamase with increased carbapenemase activity, vol 68.
381 9. Clinical and Laboratory Standards Institute C. 2014. Performance standards for antimicrobial
382 susceptibility testing; twenty-fourth informational supplement, vol 34.
383 10. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
384 Pham S, Prjibelski AD. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to
385 Single-Cell Sequencing. Journal of Computational Biology 19:455-477.
386 11. Koren S, Walenz B, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and
387 accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Research
388 27:722-736.
389 12. Zhou Y, Liang Y, Lynch KH, Dennis J, Wishart DS. 2011. PHAST: A Fast Phage Search Tool.
390 Nucleic Acids Research 39:347-352.
391 13. Brettin T, Davis JJ, Disz T, Edwards R, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch
392 GD. 2015. RASTtk: A modular and extensible implementation of the RAST algorithm for building
393 custom annotation pipelines and annotating batches of genomes. Scientific Reports 5:8365-8365.

394 14. Thomsen MCF, Ahrenfeldt J, Cisneros J, Jurtz VI, Larsen MV, Hasman H, Aarestrup FM, Lund O.
395 2016. A Bacterial Analysis Platform: An Integrated System for Analysing Bacterial Whole Genome
396 Sequencing Data for Clinical Diagnostics and Surveillance. PLOS ONE 11.
397 15. Simner PJ, Antar AAR, Hao S, Gurtowski J, Tamma PD, Rock C, Opene BNA, Tekle T, Carroll KC,
398 Schatz MC. 2018. Antibiotic pressure on the acquisition and loss of antibiotic resistance genes in
399 Klebsiella pneumoniae. Journal of Antimicrobial Chemotherapy 73:1796-1803.
400 16. Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL. 2012. Gene prediction with Glimmer for
401 metagenomic sequences augmented by classification and clustering. Nucleic Acids Research 40.
402 17. Treangen TJ, Ondov BD, Koren S et al. 2014. The Harvest suite for rapid core-genome alignment and
403 visualization of thousands of intraspecific microbial genomes. Genome Biol 15: 524.
404 18. Bi D, Jiang X, Sheng ZK, Ngmenterebo D, Tai C, Wang M, Deng Z, Rajakumar K, Ou HY. Mapping
405 the resistance-associated mobilome of a carbapenem-resistant Klebsiella pneumoniae strain reveals
406 insights into factors shaping these regions and facilitates generation of a 'resistance-disarmed' model
407 organism. J Antimicrob Chemother. 2015 Oct;70(10):2770-4.
408 19. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform.
409 Bioinformatics 25:1754-1760.
410 20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth GT, Abecasis GR, Durbin R.
411 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.
412 21. Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky AM, Garimella K, Altshuler
413 D, Gabriel SB, Daly MJ. 2010. The Genome Analysis Toolkit: A MapReduce framework for
414 analyzing next-generation DNA sequencing data. Genome Research 20:1297-1303.
415 22. Vale F, Nunes A, Oleastro M, Gomes J, A. Sampaio D, Rocha R, Vítor J, Engstrand L, Pascoe B,
416 Berthenet E, K. Sheppard S, Hitchings M, Megraud F, Vadivelu J, Lehours P. 2017. Genomic
417 structure and insertion sites of Helicobacter pylori prophages from various geographical origins.
418 Scientific Reports 7:42471.
419 23. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, Ou HY. oriTfinder: a web-based tool for the identification
420 of origin of transfers in DNA sequences of bacterial mobile genetic elements. Nucleic Acids Res.
421 2018 Jul 2;46(W1):W229-W234.
422 24. Ito R, Mustapha MM, Tomich AD, Callaghan JD, Mcelheny CL, Mettus RT, Shanks RMQ,
423 Sluiscremer N, Doi Y. 2017. Widespread Fosfomycin Resistance in Gram-Negative Bacteria
424 Attributable to the Chromosomal fosA Gene. Mbio 8.
425 25. Grabherr M, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
426 Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a
427 reference genome. Nature Biotechnology 29:644-652.

428 26. Rimmer, A. J., Phan, H., Mathieson, I., et al. 2014. Integrating mapping-, assembly- and haplotype-
429 based approaches for calling variants in clinical se- quencing applications. Nature Genetics, 46, 912-
430 918.
431 27. Wei Z, Du X, Yu Y, Shen P, Chen Y, Li L. 2007. Plasmid-mediated KPC-2 in a Klebsiella
432 pneumoniae isolate from China. Antimicrobial Agents and Chemotherapy 51:763-765.
433 28. Zhang R, Liu L, Zhou H, Chan EW, Li J, Fang Y, Li Y, Liao K, Chen S. 2017. Nationwide
434 Surveillance of Clinical Carbapenem-resistant Enterobacteriaceae (CRE) Strains in China.
435 EBioMedicine 19:98-106.
436 29. Liu J, Yu J, Chen F, Yu J, Simner P, Tamma P, Liu Y, Shen L. 2018. Emergence and establishment of
437 KPC-2-producing ST11 Klebsiella pneumoniae in a general hospital in Shanghai, China. European
438 Journal of Clinical Microbiology & Infectious Diseases 37:293-299.
439 30. Meropol SB, Haupt AA, Debanne SM. 2018. Incidence and Outcomes of Infections Caused by
440 Multidrug-Resistant Enterobacteriaceae in Children, 2007–2015. Journal of the Pediatric Infectious
441 Diseases Society 7:36-45.
442 31. Patel PK, Russo TA, Karchmer AW. 2014. Hypervirulent Klebsiella pneumoniae. Open Forum
443 Infectious Diseases 1.
444 32. 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
445 third-generation cephalosporin-resistant hypervirulent Klebsiella pneumoniae due to the acquisition of
446 a self-transferable blaDHA-1-carrying plasmid by an ST23 strain. Virulence. 2018 Dec 31;9(1):838-
447 844.
448
449 33. Wang X, Xie Y, Li G, Liu J, Li X, Tian L, Sun J, Ou HY, Qu H. Whole-Genome-Sequencing
450 characterization of bloodstream infection-causing hypervirulent Klebsiella pneumoniae of capsular
451 serotype K2 and ST374. Virulence. 2018 Jan 1;9(1):510-521.
452 34. Hossain Mondal A, Siddiqui MT, Sultan I, Mohd. Rizwanul Haq Q. 2018. Prevalence and diversity of
453 bla TEM, bla SHV and bla CTX-M variants among multidrug resistant Klebsiella spp. from an urban
454 riverine environment in India. International Journal of Environmental Health Research 6:1-13.
455 35. Ito R, Mustapha M, D. Tomich A, D. Callaghan J, McElheny C, T. Mettus R, Shanks R, Sluis-Cremer
456 N, Doi Y. 2017. Widespread Fosfomycin Resistance in Gram-Negative Bacteria Attributable to the
457 Chromosomal fosA Gene. mBio 8:e00749-17.
458 36. Li G, Zhang Y, Bi D, Shen P, Ai F, Liu H, Tian Y, Ma Y, Wang B, Rajakumar K, Ou HY, Jiang X.
459 First report of a clinical, multidrug-resistant Enterobacteriaceae isolate coharboring fosfomycin
460 resistance gene fosA3 and carbapenemase gene blaKPC-2 on the same transposon,
461 Tn1721.Antimicrob Agents Chemother. 2015 Jan;59(1):338-43.
462 37. Ma Y, Xu X, Guo Q, Wang P, Wang W, Wang M. 2015. Characterization of fosA5, a new plasmid-
463 mediated fosfomycin resistance gene in Escherichia coli. Letters in Applied Microbiology 60:259-264.

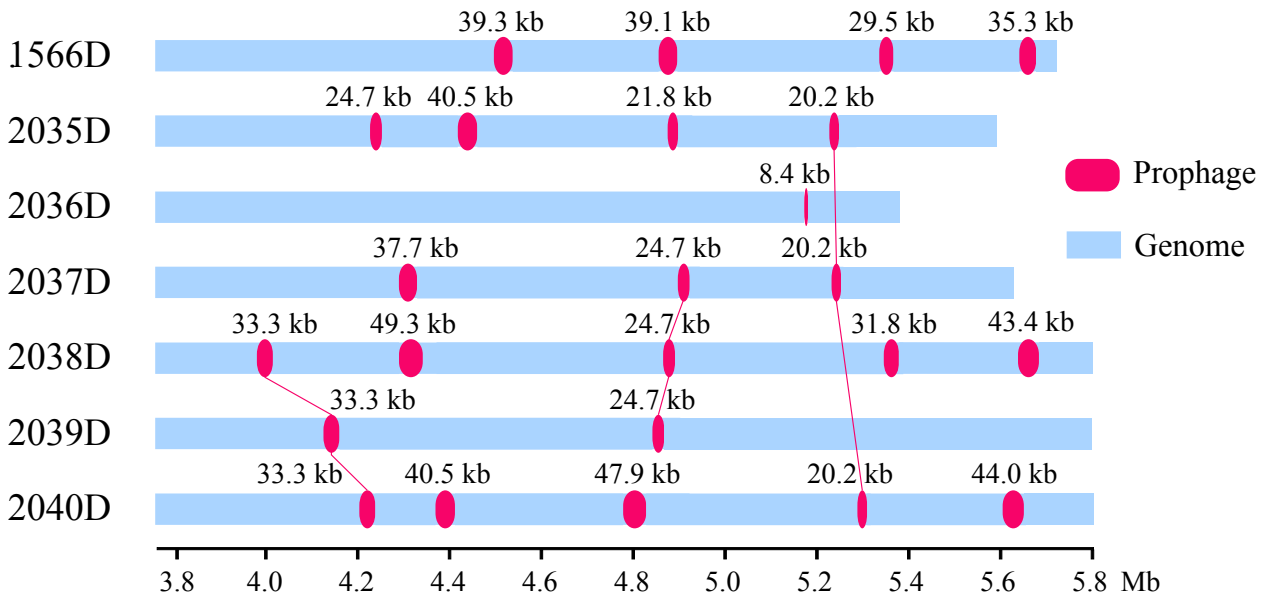
38. Huang L, Yan Hu Y, Zhang R. 2017. Prevalence of fosfomycin resistance and plasmid-mediated fosfomycin-modifying enzymes among carbapenem-resistant Enterobacteriaceae in Zhejiang, China. Journal of Medical Microbiology 66:1332-1334.

39. 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. 2018. Overexpression of OqxAB and MacAB efflux pumps contributes to eravacycline resistance and heteroresistance in clinical isolates of Klebsiella pneumoniae. Emerg Microbes Infect 7:139.

490 **Figures**

491 **Figure 1. Intact prophages identified in seven CRKP strains.**

492



493

494

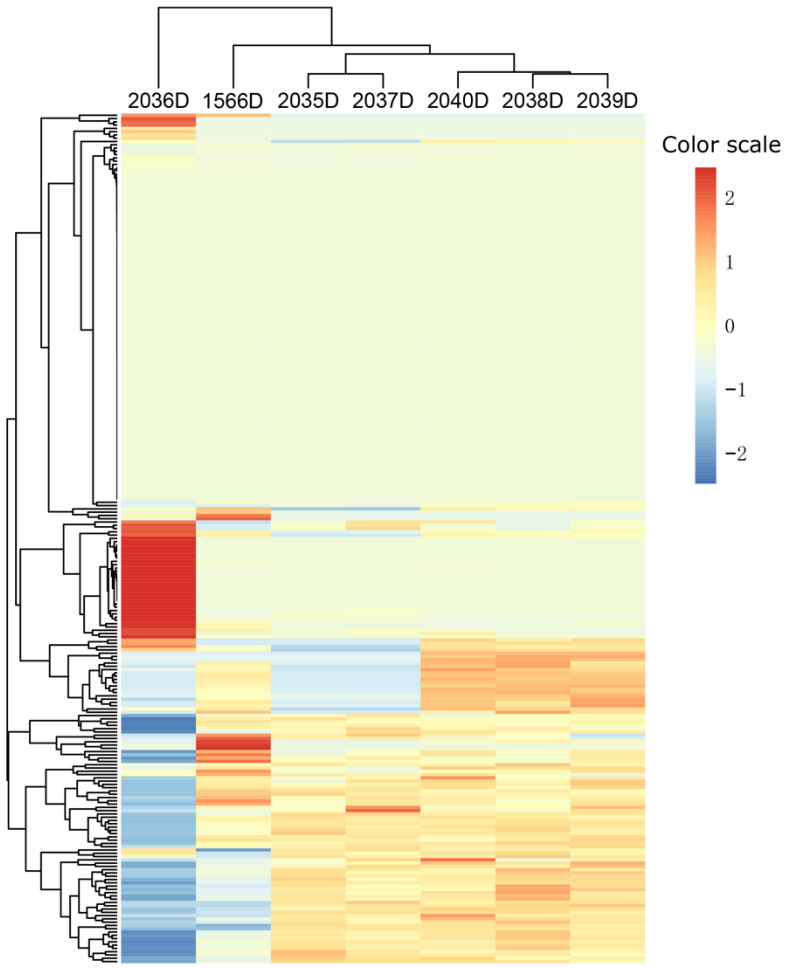
495

496 **Figure 2. Assessing the genetic relatedness of the CRKP by WGS.** (a) Clustering of all isolates based on

497 sequence similarity. (b) Communal SNP markers detected by pairwise comparison analysis.

498

(a)



(b)

Isolates	1566D	2035D	2036D	2037D	2038D	2039D	2040D
1566D		6922	2275	6895	7952	7848	8029
2035D	6922		2304	7249	7103	6989	7152
2036D	2275	2304		2287	2333	2342	2344
2037D	6895	7249	2287		7050	6964	7132
2038D	7952	7103	2333	7050		8417	8549
2039D	7848	6989	2342	6964	8417		8467
2040D	8029	7152	2344	7132	8549	8467	

499

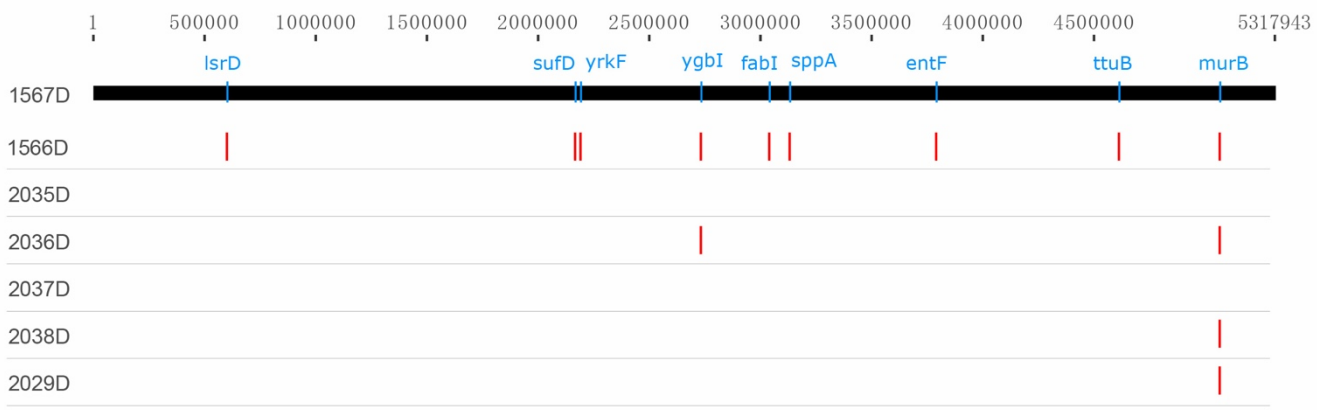
500

501

502

503 **Figure 3. GWAS results of the analysis.**

504



505

506

Tables

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. S: susceptible; I: intermediate; R: resistant.

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）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）
ampicillin/ sulbactam	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）	≥32(R）
piperacillin/ tazobactam	≥128(R）	≥128(R）	≥128(R）	≥128(R）	≥128(R）	≥128(R）	≥128(R）	≥128(R）
cefazolin	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
cefotetan	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
ceftazidime	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
ceftriaxone	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
cefepime	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
aztreonam	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）
imipenem	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）
amikacin	≥64(R）	≤4（S)	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）	≥64(R）

gentamicin	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）
tobramycin	≥16(R）	≤2（S）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）	≥16(R）
ciprofloxacin	≥4(R）	≥4(R）	≥4(R）	≥4(R）	≥4(R）	≥4(R）	≥4(R）	≥4(R）
levofloxacin	≥8(R）	≥8(R）	≥8(R）	≥8(R）	≥8(R）	≥8(R）	≥8(R）	≥8(R）
macrodantin	256(R）	256(R）	≥512(R）	≥512(R）	≥512(R）	≥512(R）	≥512(R）	≥512(R）
sulfamethoxazole/ trimethoprim	≥320(R）	≤20（S）	≤20（S）	≤20（S）	≤20（S）	≥320(R）	≥320(R）	≥320(R)

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

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-]	IncHI1, 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/	*blaCTX-M-14*	*blaCTX-M-65*		*blaCTX-M-65*	*blaCTX-M-65*	*blaCTX-M-65*	*blaCTX-M-14*
Narrow-Spectrum	*blaKPC-2*	*blaKPC-2*	*blaCTX-M-3*	*blaKPC-2*	*blaKPC-2*	*blaKPC-2*	*blaKPC-2*
Cephalosporins: cefazolin and	*blaLEN12*	*blaSHV-11*	*blaSHV-93*	*blaSHV-11*	*blaSHV-11*	*blaSHV-11*	*blaSHV-11*
cefotetan	*blaTEM-1B*	*blaTEM-1B*		*blaTEM-1B*	*blaTEM-1B*	*blaTEM-1B*	*blaTEM-1B*
β-lactam inhibitors/ Carbapenems	*blaKPC-2*	*blaKPC-2*	-	*blaKPC-2*	*blaKPC-2*	*blaKPC-2*	*blaKPC-2*
Extended-Spectrum Cephalosporins/ Monobactam	*blaCTX-M-14*	*blaCTX-M-65*	*blaCTX-M-3*	*blaCTX-M-65*	*blaCTX-M-65*	*blaCTX-M-65*	*blaCTX-M-65*
Aminoglycosides	*rmtB*	*rmtB*	*aac(3)-IId*	*rmtB*	*aac(3)-IId* *rmtB*	*aac(3)-IId* *rmtB*	*aac(3)-IId* *rmtB*
Fluoroquinolones	*QnrS1*	-	*oqxA* *oqxB*	-	*QnrS1*	*QnrS1*	*QnrS1*
Phosphonic Acid	*fosA*	*fosA*	*fosA*	*fosA*	*fosA*	*fosA*	*fosA*
Phenicol	-	*catA2*	*catA1*	*catA2*	*catA2*	*catA2*	*catA2*
Folate-pathway Inhibitors	*dfrA1*	-	*dfrA17*	-	*dfrA1*	*dfrA1*	*dfrA1*