

## Accepted Manuscript

Title: First genomic insights into carbapenem-resistant *Klebsiella pneumoniae* from Malaysia

Authors: Han Ming Gan, Wilhelm Wei Han Eng, Amreeta Dhanoa



PII: S2213-7165(19)30174-2  
DOI: <https://doi.org/10.1016/j.jgar.2019.07.008>  
Reference: JGAR 986

To appear in:

Received date: 29 May 2019  
Revised date: 3 July 2019  
Accepted date: 5 July 2019

Please cite this article as: Gan HM, Eng WWH, Dhanoa A, First genomic insights into carbapenem-resistant *Klebsiella pneumoniae* from Malaysia, *Journal of Global Antimicrobial Resistance* (2019), <https://doi.org/10.1016/j.jgar.2019.07.008>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# First genomic insights into carbapenem-resistant *Klebsiella pneumoniae* from Malaysia

Han Ming Gan<sup>a,b,c,d\*</sup>, Wilhelm Wei Han Eng<sup>c,d</sup>, Amreeta Dhanoa<sup>e, f\*</sup>

<sup>a</sup>Deakin Genomics Centre, Deakin University, Geelong, Victoria, Australia

<sup>b</sup>Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Geelong, Victoria, Australia

<sup>c</sup>Monash University Malaysia Genomics Facility, Tropical Medicine and Biology Multidisciplinary Platform, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

<sup>d</sup>School of Science, Monash University Malaysia, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

<sup>e</sup>Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

<sup>f</sup>Tropical Medicine and Biology Platform, Monash University Malaysia, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

**\*Corresponding Author:** Han Ming Gan, 75 Pigdons Road, Waurin Ponds, Victoria 3216, Australia, han.gan@deakin.edu

Amreeta Dhanoa, Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Bandar Sunway, Jalan Lagoon Selatan, Bandar Sunway 47500, Selangor, Malaysia, amreeta.dhanoa@monash.edu

## Highlights:

- First genomic characterization of carbapenem-resistant *Klebsiella pneumoniae* from Malaysia

- The eight sequenced strains exhibited high genetic diversity (six sequence types identified) despite similar isolation source
- Genomic potential for carbapenem-resistance was due to the presence of plasmid-localized *bla*<sub>NDM</sub> (*bla*<sub>NDM-1</sub>/*bla*<sub>NDM-5</sub>) or *bla*<sub>KPC</sub> (*bla*<sub>KPC-2</sub>/*bla*<sub>KPC-6</sub>) gene that was flanked mostly by repetitive sequences.
- We also report the first assembled genome of *bla*<sub>KPC-6</sub> -harboring *K. pneumoniae*

## ABSTRACT

### Objectives

Despite the increasing report of carbapenem-resistant Enterobacteriaceae in Malaysia, genomic resource for carbapenem-resistant clinical strains of *Klebsiella pneumoniae* remains unavailable. In this study, we aim to fill in this gap by sequencing the genomes of multiple carbapenem-resistant *K. pneumoniae* strains from Malaysia in addition to identifying the genetic basis for their resistance.

### Methods

Illumina whole-genome sequencing was performed on eight carbapenem-resistant *Klebsiella pneumoniae* isolated from a Malaysian hospital. Genetic diversity was inferred from the assembled genomes based on *in-silico* multi locus sequence typing (MLST). In addition, plasmid- and chromosome-derived contigs were predicted using machine learning approach. After genome annotation, genes associated with carbapenem resistance were identified based on similarity search against the ResFinder database.

## Results

The eight *K. pneumoniae* isolates were grouped into six different sequence types, some of which were only represented by a single isolate in the MLST database. Genomic potential for carbapenem-resistance was attributed to the presence of plasmid-localized *bla*<sub>NDM</sub> (*bla*<sub>NDM-1</sub>/*bla*<sub>NDM-5</sub>) or *bla*<sub>KPC</sub> (*bla*<sub>KPC-2</sub>/*bla*<sub>KPC-6</sub>) in these sequenced strains. A majority of these carbapenem resistance genes were flanked by repetitive (transposase or integrase) sequences, suggesting their potential mobility. We also report the first *bla*<sub>KPC-6</sub>-harboring plasmid contig to be assembled for *K. pneumoniae*, the second for the genus *Klebsiella*.

## Conclusion

We report the first genomic resources for carbapenem-resistant *K. pneumoniae* from Malaysia. The high diversity of carbapenem resistance genes and sequence types uncovered from only 8 isolates from the same hospital is worrying and indicates an urgent need to improve the genomic surveillance of clinical *K. pneumoniae* in Malaysia.

## KEYWORDS

*Klebsiella pneumoniae*, plasmid, carbapenem, Illumina sequencing, *bla*<sub>KPC-6</sub>

## INTRODUCTION

Carbapenems remain the first-line therapeutic antimicrobials for severe infections caused by organisms such as the extended-spectrum  $\beta$ -lactamase (ESBL)-producing multidrug-resistant Enterobacteriaceae<sup>1</sup>. Therefore, the increasing worldwide trend of carbapenem-resistant Enterobacteriaceae (CRE) represents a formidable threat to modern healthcare and is associated with high

morbidity and mortality <sup>2</sup>. In addition to the resistance to beta-lactams, CREs are frequently resistant to most other classes of antibiotics, diminishing and even eliminating the efficacy of the available antibiotic armamentarium <sup>1</sup>. Amongst the risk factors associated with carbapenem resistance include previous use of carbapenems, underlying comorbidities, longer hospital stay, mechanical ventilation, intensive care unit (ICU) stay, surgeries and transfer from healthcare settings with high rates of carbapenem resistance <sup>1,2</sup>.

Among Enterobacteriaceae, *Klebsiella* species have emerged as the most important pathogens causing a wide range of healthcare related infections including pneumonia, bacteraemia, urinary tract and wound infections <sup>3</sup>. Over the last decades, carbapenem resistance has been steadily escalating among the Enterobacteriaceae, especially amongst the *Klebsiella* species <sup>1</sup>. Multiple carbapenem-resistance mechanisms have been identified, including the production of carbapenemases with direct carbapenem-hydrolyzing activity, over-expression of efflux pumps, and reduced permeability of outer membrane mediated by porin mutations. Three types of carbapenemases (class A: bla<sub>KPC</sub>, bla<sub>GES</sub>; class B: bla<sub>NDM-1</sub>, bla<sub>IMP</sub>, and bla<sub>VIM</sub>; class D: bla<sub>OXA-48</sub>) hydrolyze carbapenems at varying degree <sup>3</sup>. Carbapenem genes are carried mainly on plasmids and pose a potential for widespread dissemination of carbapenem resistance which describes their inclination to cause outbreaks of infection within and between healthcare facilities <sup>2,4</sup>. Understanding the molecular epidemiology and genetic characteristics of carbapenem-resistant *Klebsiella* strains from hospital environment requires a discriminatory bacterial typing technique. However, methods widely used such as pulsed-field gel electrophoresis (PFGE) or multilocus sequence typing (MLST) lack this discriminatory resolution <sup>4</sup>.

Whole genome sequencing (WGS) elicits a level of discrimination on genetic relatedness that readily surpasses the previous typing methods <sup>4</sup>. However, there is a scarcity of reports on WGS of antibiotic resistant clinical

Enterobacteriaceae isolated from Malaysia and most have been confined mainly to genome sequencing of single isolates without in-depth comparative analysis<sup>5-8</sup>. We performed WGS of the 8 carbapenem resistant clinical *Klebsiella pneumoniae* strains isolated from a hospital in Malaysia, providing the first insight into the genetic diversity and antimicrobial resistance mechanisms of carbapenem-resistant *Klebsiella pneumoniae* in Malaysia. This data is especially important, since resistance to imipenem and meropenem is on the rise in Malaysia; from 2.3% and 2.6% respectively in 2016 to 4.2% and 4.7% in 2017<sup>9</sup>. WGS will provide important information to infer the origin and the spread of *Klebsiella pneumoniae* strains and their antibiotic resistance genes in the healthcare setting, facilitating future epidemiological surveillance and infection control efforts in Malaysia.

## MATERIALS AND METHODS

### **Bacterial isolates and the determination of minimal inhibitory concentration (MIC)**

Eight non-clonal clinical *Klebsiella pneumoniae* strains isolated from a hospital in Johor, Malaysia, were revived from -20°C glycerol stock cultures and grown on nutrient agar plate at 30°C for 48 hours. These isolates were anonymized prior to DNA extraction and sequencing and no clinical or demographic data was collected. These strains were sub-cultured onto the same medium and tested for an expanded range of antimicrobial sensitivities using the VITEK 2 system (BioMérieux, Marcy l'Etoile, France). Specifically, we used the VITEK 2 AST-GN87 cards to test for the following beta-lactam antibiotics (concentration range): ampicillin/sulbactam (2/1-32/16 µg/ml), cefazolin (4-64 µg/ml), ceftriaxone (1-64 µg/ml), cefepime (1-64 µg/ml), ceftazidime (1-64 µg/ml), piperacillin/tazobactam (4/4-128/4 µg/ml), imipenem (0.25-16 µg/ml), meropenem (0.25-16 µg/ml), ertapenem (0.5-8 µg/ml). Susceptibility category

was designated according to the 2017 Clinical and Laboratory Standards Institute (CLSI) M100-S27 guidelines. *Escherichia coli* ATCC 25922 was used as the quality control strain for antibiotic susceptibility testing.

### **Whole genome sequencing**

Five to ten bacterial colonies were scrapped from the 2-day old agar plate culture with a sterile P200 pipette tip and used for genomic DNA extraction using a conventional SDS-based extraction method. The purified gDNA was quantified with Qubit 2 (Invitrogen, Santa Clara, CA), normalized to 0.2 ng/μl and processed using the Nextera XT library preparation kit (Illumina, San Diego, CA). Sequencing was performed on the Miseq (run configuration of 2 × 250 bp) located at the Monash University Malaysia Genomics Facility.

### **Genome assembly and analysis**

Raw paired-end reads were adapter-trimmed using Trimmomatic v0.36<sup>10</sup>. Error correction followed by *de novo* assembly of the trimmed reads used Unicycler v0.3.0<sup>11</sup>. Genome-based species verification was performed with Jspecies v1.2.1 using the “mummer” algorithm<sup>12</sup>. . The average nucleotide identity matrix produced by Jspecies v.1.2.1 was subsequently inputted into R<sup>13</sup> to generate a clustered heatmap with the pheatmap package (default setting)<sup>14</sup>. Upon species confirmation, *in-silico* multi-locus sequence typing were performed on the assembled genomes using the software mlst v2.16 (<https://github.com/tseemann/mlst>) that scans the assembled contigs against the *Klebsiella pneumoniae* PubMLST typing schemes<sup>15</sup>. Screening of contigs for antimicrobial resistance genes against the Resfinder database (accessed on 28-Jul-2018)<sup>16</sup> used ABRicate v0.8.11 (<https://github.com/tseemann/abricate>).

### **Identification and visualization of plasmid-derived contigs**

Mlplasmids was used to accurately predict plasmid- and chromosomal-derived sequences in the genome assemblies based on pentamer frequency and machine learning approach<sup>17</sup>. More specifically, mlplasmids used the support-vector machine models that have been trained on various complete *Klebsiella pneumoniae* genomes to classify the origin of contigs, enabling accurate *in-silico* localization of antibiotic resistance genes to the plasmid or chromosome of *Klebsiella pneumoniae* isolates. Visualization and alignment of the plasmid-derived contigs used BRIG v0.95 and EasyFig v2.1<sup>18,19</sup>.

## RESULTS

### Genome Assembly and statistics:

The assembled genome size ranges from 5.23 Mb to 5.64 Mb (GC content of 57.14 – 57.57%) (See Supplemental Table 1 for NCBI accession codes). Each strain exhibited more than 98% pairwise average nucleotide identity (ANI) to the currently described type strains of *K. pneumoniae* subspecies with consistently less than 95% pairwise ANI to *K. quasipneumoniae* subsp. *similipneumoniae* 07A044<sup>T</sup> and *K. quasipneumoniae* subsp. *quasipneumoniae* 01A030<sup>T</sup><sup>20</sup> (Figure 1). Within the ANI heatmap, minor clustering could be observed for isolates that exhibit strikingly high pairwise ANI (>99.8%) for example MGF001 and MGF018 as well as MGF002 and MGF009. Although a relatively high pairwise ANI (>99%) was observed among *K. pneumoniae* subsp. *pneumoniae* DSM 30104<sup>T</sup>, *K. pneumoniae* subsp. *ozaenae* DSM 16358<sup>T</sup> and *K. pneumoniae* subsp. *rhinoscleromatis* DSM 16231<sup>T</sup>, this is not the case for *K. quasipneumoniae* subsp. *similipneumoniae* 07A044<sup>T</sup> and *K. quasipneumoniae* subsp. *quasipneumoniae* 01A030<sup>T</sup> (pairwise ANI <96%) (Figure 1).

### Multiple genetic origins of *Klebsiella pneumoniae* isolates as revealed by *in-silico* MLST



All *Klebsiella pneumoniae*-MLST housekeeping genes (*rpoB*, *gapA*, *mdh*, *pgi*, *phoE*, *infB* and *tonB*)<sup>21</sup> are complete and present as a single-copy in the assembled *Klebsiella pneumoniae* genomes. Although we did not uncover any novel sequence type which is somewhat unusual given the paucity of Malaysian *K. pneumoniae* strain representation in the MLST database, we found a high diversity of sequence types among the 8 carbapenem-resistant strains. A total of 5 sequence types (ST) e.g. ST11 (MGF001 and MGF018), ST258 (MGF011), ST530 (MGF004), ST584 (MGF020), and ST3157 (MGF002 and MGF009) were identified and strains forming minor clusters in the ANI heatmap were classified to the same ST consistent with their high genomic relatedness (Table 1). It is also worth noting that strains sharing the same ST also share the same isolation source. For example, all strains from ST11, ST530, ST3157 and ST3414 were isolated from urine samples while strains belonging to ST258 and ST584 were isolated from body fluid samples.

### **Evidence of genomic potential for carbapenamase production in the sequenced isolates**

VITEKS2 assessment indicates that all isolates were highly resistant to the 3 carbapenems tested namely ertapenem, imipenem and meropenem (Table 1). In addition to carbapenems, the isolates were also highly resistant to cephalosporins (cefazolin, ceftazidime, ceftriaxone, cefepime) and beta-lactam/beta-lactamase inhibitor combination (ampicillin-sulbactam, piperacillin-tazobactam). Genome mining for antibiotic resistance genes revealed that each strain harbours one of the two main classes of carbapenem resistance gene namely *bla*<sub>NDM</sub> and *bla*<sub>KPC</sub> in addition to some ESBL genes (Table 1). Two known genetic variants were found for each of the carbapenemase genes. Based on currently sampling, the *bla*<sub>NDM-1</sub> gene is the most commonly found followed by *bla*<sub>NDM-5</sub>. The distribution of *bla*<sub>NDM</sub> and *bla*<sub>KPC</sub> variants is not random and appears to be sequence type-specific. It is also worth noting that, despite sharing

the identical *bla*<sub>NDM-1</sub> gene sequence with MGF004 and MGF019 belonging to ST530 and ST3414, respectively, MGF001 and MGF018 (both assigned to ST11) exhibited a substantially lower resistance (<16 µg/ml) to the carbapenem drug, meropenem (Table 1).

### Gene synteny in the neighbourhood of carbapenemase genes

Plasmid-derived contigs were identified in all 8 strains with cumulative length ranging from 226kb to 580kb (Supplemental Table 2). The contigs containing *bla*<sub>NDM</sub> or *bla*<sub>KPC</sub> were all classified as “plasmid” by mlplasmids. Five out of eight of the *bla*<sub>NDM</sub>/*bla*<sub>KPC</sub>-containing contigs are more than 15 kb with the longest one being the *bla*<sub>KPC-6</sub>-containing contig (83kb) from strain MGF020 (Figures 2 and 3). The three remaining plasmid contigs are only less than 3,200 bp in length and consist of a *bla*<sub>NDM</sub> gene and three upstream genes (*nagA*, *trpF* and *ble*) (Figure 2). The gene cluster *nagA-trpF-ble* that encodes for alpha-N-acetylgalactosaminidase, anthranilate isomerase and bleomycin-binding protein is consistently found upstream of *bla*<sub>NDM</sub> in the longer *bla*<sub>NDM</sub>-containing contigs. In two of the *bla*<sub>NDM</sub>-containing contigs e.g. contig36<sub>MGF004</sub> and contig37<sub>MGF019</sub>, two genes both coding for integrase were located downstream of *bla*<sub>NDM</sub> and arranged in tandem. On the contrary, *bla*<sub>KPC</sub> was immediately flanked by transposase-coding genes transcribed in opposite orientation.

### The first *K. pneumoniae* plasmid-derived contig harbouring the *bla*<sub>KPC-6</sub> variant

Using the entire contig25 from the KPC-6-producing MGF020 isolate as the query to search against the NCBI non-redundant nucleotide database (NCBI nt database, accessed on 27<sup>th</sup> February 2019), significant matches with high query coverage (>90% of query length) were found for a few *bla*<sub>KPC-2</sub>-containing complete InA/C plasmid sequences. Interestingly, BLASTN search using only the annotated *bla*<sub>KPC-6</sub> gene in contig25<sub>MGF020</sub> as the query against the NCBI

non-redundant nucleotide database (accessed on 27<sup>th</sup> Feb 2019) returned only the reference CDS of *bla*<sub>KPC-6</sub> (GenBank Accession Number: EU555534.1) with exact match (100% identity) followed by several 99.887% identity hits representing a single mismatch to the 882 bp *bla*<sub>KPC-6</sub> gene to some *bla*<sub>KPC-2</sub>-containing plasmids. Within the Enterobacteriaceae whole genome shotgun sequence database, we only found exact match of the *bla*<sub>KPC-6</sub> gene to two contigs, JRTV01000008 and JRTV01000009, both from the genome assembly of *K. variicola* strain 223/14 also isolated from Malaysia.

Using the plasmid-derived contigs from MGF020 as the query (Supplemental Table 2), we observed a significant nucleotide alignment (>90% nucleotide identity) across the entire plasmid KP1766\_p1 of *K. pneumoniae* strain KP1766 (GenBank Accession : CP025147.1) except for the region ranging from 15 kb to 40 kb that consists of phage-protein coding genes (purple outer arrows in Figure 3). This region was similarly absent from the complete pKPC\_CAV1344 plasmid of *K. pneumoniae* strain CAV1344 (GenBank Accession: CP011622.1). Contig25<sub>MGF020</sub> covered approximately 30% length fraction of plasmid KP1766\_p1 and contained the typical plasmid conjugation *tra* genes (red outer arrows in Figure 3). Interestingly, despite harbouring the *bla*<sub>KPC-6</sub> gene, very little significant sequence match to the reference IncA/C plasmid was observed when the whole genome assembly of Malaysian *bla*<sub>KPC6</sub>-producing *K. variicola* was used as the query (Figure 3).

## DISCUSSION

We assembled the genomes of 8 *K. pneumoniae* isolates from a Malaysian hospital. The genomes represent a 100% increase in the number of publicly available Malaysian *K. pneumoniae* genomes (NCBI Assembly database as of 27<sup>th</sup> February 2019), underscoring the paucity of genomic representation of this clinically important bacterial species from Malaysia. In addition, these genomes are first carbapenem-resistant *K. pneumoniae* genomes

reported from Malaysia although it is worth mentioning that the genomes of carbapenem-resistant *K. variicola* and *K. quasipneumoniae* isolated from Malaysia have been also recently published <sup>5,6</sup>. The high genetic diversity of *K. pneumoniae* in this current genomic sampling as indicated by the number of different STs recovered from this sampling is unexpected given their identical sampling site, suggesting multiple introduction/origin of carbapenem-producing strains in the hospital during the strain isolation period. While some of the STs identified were common clinically important STs such as ST11 and ST258 <sup>22-24</sup>, some STs are rare with only a few representative strains reported worldwide. For example, ST3157 and ST3414 was only reported once in Australia (id 6751; isolate DMG1800058) and China (id 7803; isolate 42182), respectively <sup>15</sup>. It is also worth noting that the genomic relatedness among the subspecies of *K. pneumoniae* type strains is substantially higher than that among *K. quasipneumoniae* subspecies, indicating the lack of standardized subspecies delineation criteria among *Klebsiella* spp.

Despite being categorized as carbapenem-resistant, the *K. pneumoniae* strains belonging to ST11 in this study exhibited a lower resistance to the meropenem drug compared to other *bla*<sub>NDM</sub>-harboring strains from different STs. It is possible the strains from ST11 have a genetic makeup that increases their baseline sensitivity towards meropenem. For example, variations in the gene coding for porin channel have been associated with decreased susceptibility to meropenem drug given the direct involvement of such protein in the permeation of carbapenem drug <sup>25</sup>. Transposon mutagenesis of these strains followed by selection on higher concentration of meropenem will be instructive to identify genes that are associated with increased sensitivity to the carbapenem drug <sup>26,27</sup>. Alternatively, it is also possible that the novel *bla*<sub>SHV</sub> variants that are present in all other *bla*<sub>NDM</sub>-containing strains may be contributing to the increased resistance towards carbapenem <sup>28</sup>.

Most reported carbapenem-resistant Malaysian *K. pneumoniae* strains harbour the *bla*<sub>NDM1</sub>, *bla*<sub>oxa-48</sub> or *bla*<sub>IMP4</sub> gene variant <sup>29,30</sup>. Isolate MGF020 represents the first *bla*<sub>KPC-6</sub> harbouring *K. pneumoniae* strain to be isolated from Malaysia. More intriguingly, despite the abundance of *K. pneumoniae* genome assemblies in the NCBI database, contig25 from strain MGF020 is the first *K. pneumoniae* plasmid contig containing the *bla*<sub>KPC-6</sub> gene. In addition, its significant coverage to the 205 Kb *incA/C* plasmid KP1766\_p1 suggests that the *bla*<sub>KPC-6</sub> gene in MGF020 is harboured on a large *incA/C* plasmid. The putative plasmid length is relatively large compared to most plasmids found in clinical Enterobacteriaceae that are usually less than 100 Kb <sup>31</sup>. Although a large plasmid can carry more virulence and resistance genes, it also represents a significant burden to the host's metabolism and need to be maintained in low copy number. The persistence of large plasmid among *Klebsiella* strains as well as other gram-negative bacterial genera <sup>32-34</sup> despite the lack of constant selection in the environment is largely attributed to the presence of elaborate plasmid-encoded maintenance system <sup>35,36</sup>.

The high prevalence of repetitive elements observed in the vicinity of carbapenem resistance genes is the most plausible explanation for the recovery of mostly short *bla* gene-containing contigs from our short-read only assemblies <sup>37</sup>. Such elements are known to break short-read assembly graphs thus preventing complete assembly of circular plasmid <sup>38</sup>. The complete assembly of clinically important *Klebsiella* plasmids is highly valuable as it can provide novel insights into plasmid dynamics and facilitate the tracking of plasmid transmission during outbreaks <sup>39</sup>. Future sequencing work incorporating long reads that can span long repeats such as those generated by PacBio and Nanopore technology will be instructive. Nanopore sequencing in particular is now commonly used among researchers to close microbial genome assembly gaps <sup>40,41</sup> due to its ease of use, low capital cost and wide community support.

## CONCLUSIONS

We provide a significant genomic resource for clinical carbapenem-resistant *K. pneumoniae* isolated from Malaysia. In addition to identifying multiple *K. pneumoniae* sequence types, some of which are rarely reported worldwide, we also uncovered three carbapenem resistance gene variants that are possibly encoded on different plasmid backbones. Despite the active global genomic sampling of *K. pneumoniae*, our study is the first to report a *K. pneumoniae* plasmid sequence harbouring the *bla*<sub>KPC-6</sub> gene. Our findings highlight the need for increasing the genomic surveillance of clinical *K. pneumoniae* at the national level in view of the emergence of carbapenem-resistant *K. pneumoniae* in this region.

## DECLARATIONS

**Funding:** This study was funded by the Monash University Malaysia Tropical Medicine Biology Small Grant Scheme awarded to AD.

**Ethical Approval:** Not required

**Competing Interests:** None

## ACKNOWLEDGEMENTS

We thank Monash University Malaysia Genomics Facility for the provision of computational resources and sequencing support.

## REFERENCES

1. Gupta N, Limbago BM, Patel JB, Kallen AJ. Carbapenem-resistant Enterobacteriaceae: epidemiology and prevention. *Clinical infectious diseases* 2011;53:60-7.
2. van Loon K, Voor AF, Vos MC. A systematic review and meta-analyses of the clinical epidemiology of carbapenem-resistant Enterobacteriaceae. *Antimicrobial agents and chemotherapy* 2018;62:e01730-17.
3. Ripabelli G, Sammarco M, Flocco R, Scutellà M, Recchia L, Grasso G. *Klebsiella pneumoniae* isolated from intensive care unit patients with respiratory tract infections: characterization by pulsed-field gel electrophoresis, antimicrobial resistance and pcrs for carbapenemase genes detection. *J Respir Med Lung Dis* 2017; 2 (1) 2017;1008.
4. Cella E, Ciccozzi M, Presti AL, et al. Multi-drug resistant *Klebsiella pneumoniae* strains circulating in hospital setting: whole-genome sequencing and Bayesian phylogenetic analysis for outbreak investigations. *Scientific reports* 2017;7:3534.
5. Ahmad N, Chong TM, Hashim R, Shukor S, Yin W-F, Chan K-G. Draft genome of multidrug-resistant *Klebsiella pneumoniae* 223/14 carrying KPC-6, isolated from a general hospital in Malaysia. *Journal of genomics* 2015;3:97.
6. Gan HM, Rajasekaram G, Eng WWH, Kaniappan P, Dhanoa A. Whole-genome sequences of two carbapenem-resistant *Klebsiella quasipneumoniae* strains isolated from a tertiary hospital in Johor, Malaysia. *Genome Announc* 2017;5:e00768-17.
7. Ho W-S, Gan H-M, Yap K-P, Balan G, Yeo CC, Thong K-L. Genome sequence of multidrug-resistant *Escherichia coli* EC302/04, isolated from a human tracheal aspirate. *Am Soc Microbiol*; 2012.
8. Yap K-P, Gan HM, Teh CSJ, et al. Genome sequence and comparative pathogenomics analysis of a *Salmonella enterica* serovar Typhi strain associated with a typhoid carrier in Malaysia. *Am Soc Microbiol*; 2012.
9. National antibiotic resistance surveillance report. 2017. at [https://www.imr.gov.my/images/uploads/NSAR/NSAR\\_2017/NSAR\\_report\\_2017-edited-31.1.2019.pdf](https://www.imr.gov.my/images/uploads/NSAR/NSAR_2017/NSAR_report_2017-edited-31.1.2019.pdf).)
10. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114-20.
11. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS computational biology* 2017;13:e1005595.

12. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 2015;32:929-31.
13. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010.
14. Kolde R. Pheatmap: pretty heatmaps. R package version 2012;61:617.
15. Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010;11:595.
16. Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *Journal of antimicrobial chemotherapy* 2012;67:2640-4.
17. Arredondo-Alonso S, Rogers MRC, Braat JC, et al. mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. *Microbial Genomics* 2018;4:-.
18. Alikhan N-F, Petty NK, Zakour NLB, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics* 2011;12:1.
19. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011;27:1009-10.
20. Brisse S, Passet V, Grimont PAD. Description of *Klebsiella quasipneumoniae* sp. nov., isolated from human infections, with two subspecies, *Klebsiella quasipneumoniae* subsp. *quasipneumoniae* subsp. nov. and *Klebsiella quasipneumoniae* subsp. *similipneumoniae* subsp. nov., and demonstration that *Klebsiella singaporensis* is a junior heterotypic synonym of *Klebsiella variicola*. *International Journal of Systematic and Evolutionary Microbiology* 2014;64:3146-52.
21. Diancourt L, Passet V, Verhoef J, Grimont PA, Brisse S. Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *Journal of clinical microbiology* 2005;43:4178-82.
22. Holt KE, Wertheim H, Zadoks RN, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proceedings of the National Academy of Sciences* 2015;112:E3574-E81.
23. Chen L, Mathema B, Pitout JD, DeLeo FR, Kreiswirth BN. Epidemic *Klebsiella pneumoniae* ST258 is a hybrid strain. *MBio* 2014;5:e01355-14.
24. Qi Y, Wei Z, Ji S, Du X, Shen P, Yu Y. ST11, the dominant clone of KPC-producing *Klebsiella pneumoniae* in China. *The Journal of antimicrobial chemotherapy* 2011;66:307-12.
25. Gröbner S, Linke D, Schütz W, et al. Emergence of carbapenem-non-susceptible extended-spectrum  $\beta$ -lactamase-producing *Klebsiella pneumoniae*



isolates at the university hospital of Tübingen, Germany. Journal of medical microbiology 2009;58:912-22.

26. Goryshin IY, Jendrisak J, Hoffman LM, Meis R, Reznikoff WS. Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. Nature biotechnology 2000;18:97.

27. Gan HM, Ibrahim Z, Shahir S, Yahya A. Identification of genes involved in the 4-aminobenzenesulfonate degradation pathway of *Hydrogenophaga* sp. PBC via transposon mutagenesis. FEMS microbiology letters 2011;318:108-14.

28. Ma P, Laibinis HH, Ernst CM, Hung DT. Carbapenem resistance caused by high-level expression of OXA-663  $\beta$ -lactamase in an OmpK36-Deficient *Klebsiella pneumoniae* clinical isolate. Antimicrobial Agents and Chemotherapy 2018;62:e01281-18.

29. Hamzan NI, Yean CY, Rahman RA, Hasan H, Rahman ZA. Detection of blaIMP4 and blaNDM1 harboring *Klebsiella pneumoniae* isolates in a university hospital in Malaysia. Emerging health threats journal 2015;8:26011-.

30. Low Y-M, Yap PS-X, Abdul Jabar K, et al. The emergence of carbapenem resistant *Klebsiella pneumoniae* in Malaysia: correlation between microbiological trends with host characteristics and clinical factors. Antimicrobial resistance and infection control 2017;6:5-.

31. Sengupta M, Austin S. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. Infection and immunity 2011;79:2502-9.

32. Platt TG, Morton ER, Barton IS, Bever JD, Fuqua C. Ecological dynamics and complex interactions of *Agrobacterium* megaplasmids. Frontiers in plant science 2014;5:635.

33. Lowe N, Gan HM, Chakravartty V, et al. Quorum-sensing signal production by *Agrobacterium vitis* strains and their tumor-inducing and tartrate-catabolic plasmids. FEMS microbiology letters 2009;296:102-9.

34. Hegedűs B, Kós PB, Bálint B, et al. Complete genome sequence of *Novosphingobium resinovorum* SA1, a versatile xenobiotic-degrading bacterium capable of utilizing sulfanilic acid. Journal of Biotechnology 2017;241:76-80.

35. Williams JJ, Hergenrother PJ. Exposing plasmids as the Achilles' heel of drug-resistant bacteria. Current opinion in chemical biology 2008;12:389-99.

36. Vivian A, Murillo J, Jackson RW. The roles of plasmids in phytopathogenic bacteria: mobile arsenals? Microbiology 2001;147:763-80.

37. Utturkar SM, Klingeman DM, Hurt Jr RA, Brown SD. A case study into microbial genome assembly gap sequences and finishing strategies. Frontiers in microbiology 2017;8:1272.

38. Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics* 2012;13:36.
39. Conlan S, Park M, Deming C, et al. Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. *MBio* 2016;7:e00742-16.
40. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microbial Genomics* 2017;3:-.
41. Gan HM, Lee YP, Austin CM. Nanopore long-read guided complete genome assembly of *Hydrogenophaga intermedia*, and genomic insights into 4-aminobenzenesulfonate, p-aminobenzoic acid and hydrogen metabolism in the genus *Hydrogenophaga*. *Frontiers in microbiology* 2017;8:1880.

**Figure legends:**

Figure1. A heatmap showing the hierarchical clustering of *Klebsiella pneumoniae* and *Klebsiella quasipneumoniae* strains based on genomic distance. The sequence types of each isolate reported in this study were indicated either in brackets or next to the red vertical lines. Values in boxes indicate pairwise average nucleotide identity (See Supplemental Table 1 for NCBI accession codes). *Kpr*, *K. pneumoniae* subsp. *rhinoscleromatis* ATCC 13884<sup>T</sup>; *Kpp*(D), *K. pneumoniae* subsp. *pneumoniae* DSM 30104<sup>T</sup>; *Kpp*(A), *K. pneumoniae* subsp. *pneumoniae* ATCC 13883<sup>T</sup>; *Kpo*, *K. pneumoniae* subsp. *ozaene* ATCC 11296<sup>T</sup>; *Kqs*, *K. quasipneumoniae* subsp. *similipneumoniae* 07A044<sup>T</sup>; *Kqq*, *K. quasipneumoniae* subsp. *quasipneumoniae* 01A030<sup>T</sup>.

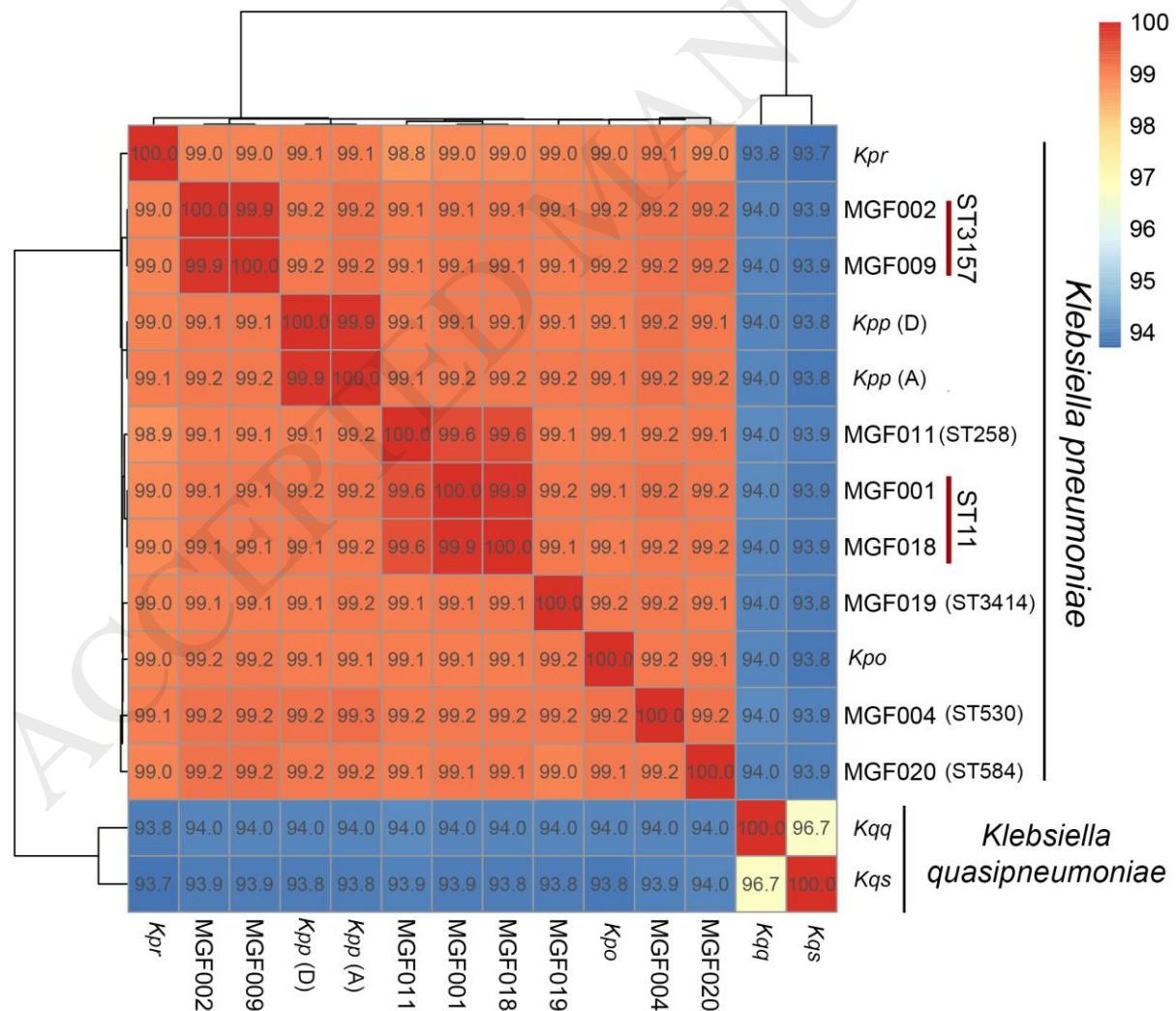


Figure 2. The gene neighbourhood of carbapenamase genes. Arrows of similar colour represent genes predicted to have similar functions. Direction of arrows indicates transcription orientation and scale bar represents contig length.

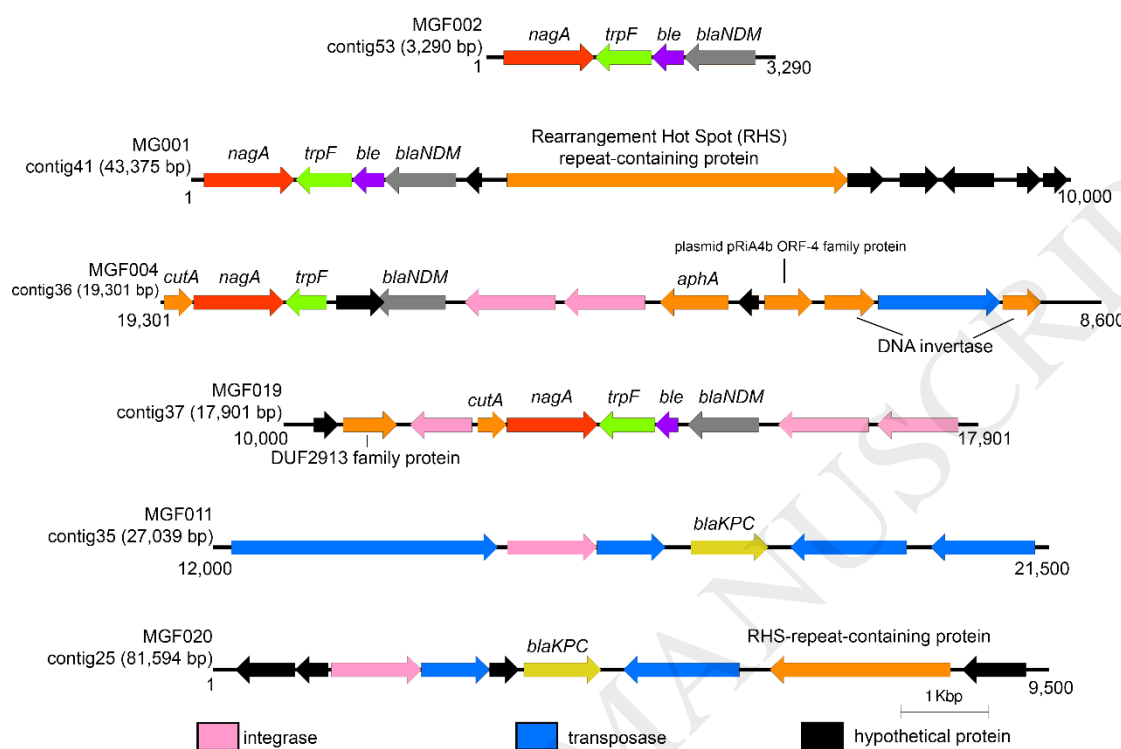


Figure 3. Circular comparison of *bla*<sub>KPC</sub> plasmids. Each plasmid is represented by a colored ring shaded based on nucleotide similarity to the reference plasmid KP1766\_p1 (min. 90%; max. 100%). The outermost ring highlights the gene regions involved in plasmid conjugation (*tra*), antibiotic resistance (*bla*<sub>TEM</sub> and *bla*<sub>KPC</sub>), phage assembly (red arrows) and gene mobility (pink arrows).

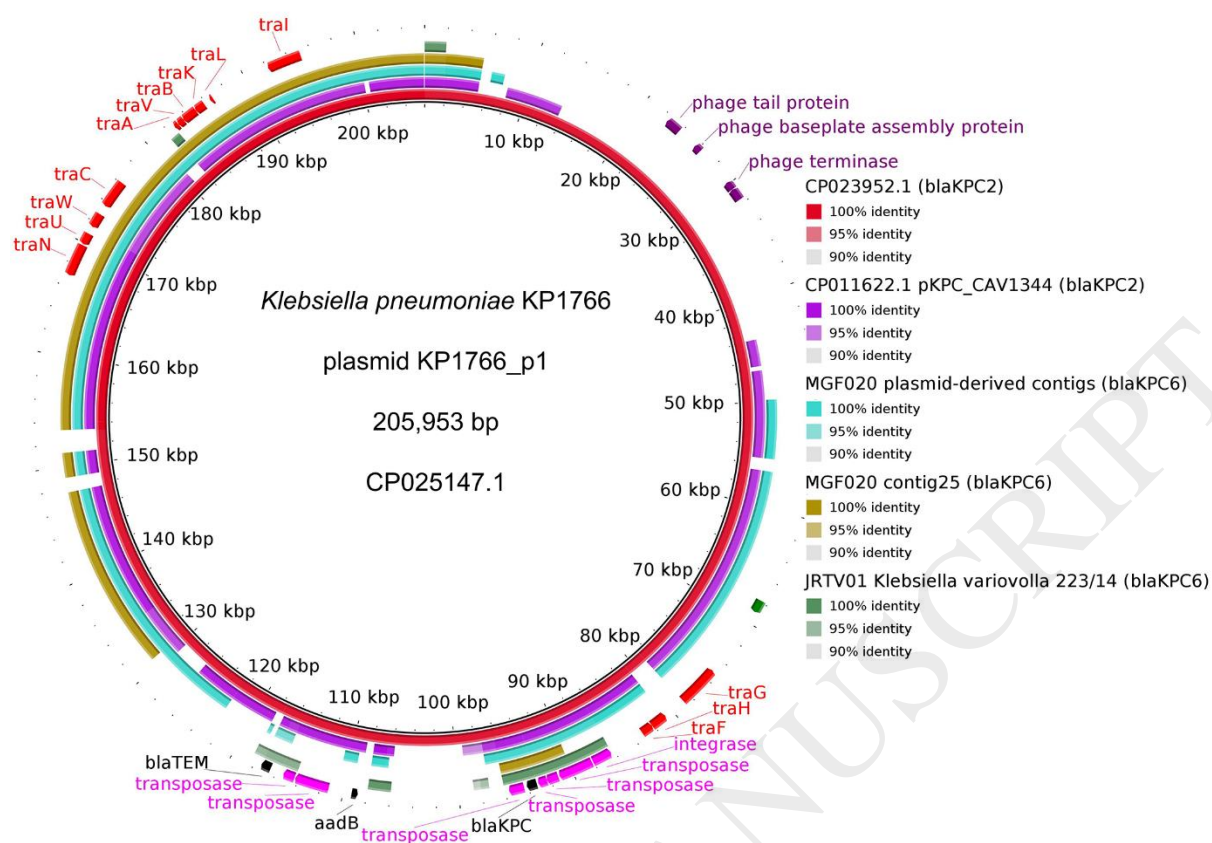


Table 1: Characteristics and susceptibility profiles of carbapenemase-producing isolates

Isolate	Sequence Type	Isolation Source	Isolation Date (mm/dd/yy)	$\beta$ -Lactamase						Beta-lactam		2 <sup>nd</sup> generatio n cephalos porin	3 <sup>rd</sup> generation cephalosporin		4 <sup>th</sup> generation cephalospo rin	Carbapenem		
				NDM	KPC	CTX-M	OXA	SHV	TEM	Ampi cillin/ sulbac tam	Piperac illin/ tazoba ctam	CFZ	CAZ	CTX	CFM	ETP	IMP	ME M
MGF001	11	Urine	03/15/12	1		15				$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	8	4
MGF002	3157	Urine	04/10/12	5		15		187*	1C	$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	$\geq 16$
MGF004	530	Urine	05/02/12	1		15		187*		$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	$\geq 16$
MGF009	3157	Urine	04/07/15	5		15		187*	1C	$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	$\geq 16$
MGF011	258	Body fluid	03/05/15		2			182*		$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	8	$\geq 16$
MGF018	11	Urine	08/18/16	1		15				$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	8
MGF019	3414	Urine	08/15/16	1			9*	187*	1B	$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	$\geq 16$
MGF020	584	Body Fluid	09/13/16		6	15			1C	$\geq 32$	$\geq 128$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 64$	$\geq 8$	$\geq 16$	$\geq 16$

\*less than 98% nucleotide identity to the reference gene