

Respiratory carriage of hypervirulent *Klebsiella pneumoniae* by indigenous populations of Malaysia

Das, Souradeep; Pandey, Anish K.; Morris, Denise E; Anderson, Rebecca; Lim, Victor; Wie, Chong Chun; Yap, Ivan Kok Seng; Alattraqchi, Ahmed Ghazi; Simin, Hafis; Abdullah, Ramle; Yeo, Chew Chieng; Clarke, Stuart C.; Cleary, David W.

DOI:

[10.1186/s12864-024-10276-4](https://doi.org/10.1186/s12864-024-10276-4)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Das, S, Pandey, AK, Morris, DE, Anderson, R, Lim, V, Wie, CC, Yap, IKS, Alattraqchi, AG, Simin, H, Abdullah, R, Yeo, CC, Clarke, SC & Cleary, DW 2024, 'Respiratory carriage of hypervirulent *Klebsiella pneumoniae* by indigenous populations of Malaysia', *BMC Genomics*, vol. 25, no. 1, 381. <https://doi.org/10.1186/s12864-024-10276-4>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.




If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

RESEARCH

Open Access



Respiratory carriage of hypervirulent *Klebsiella pneumoniae* by indigenous populations of Malaysia

Souradeep Das¹, Anish K. Pandey^{1,2}, Denise E Morris¹, Rebecca Anderson¹, Victor Lim³, Chong Chun Wie^{4,5}, Ivan Kok Seng Yap⁵, Ahmed Ghazi Alattraqchi⁶, Hafis Simin⁷, Ramle Abdullah⁸, Chew Chieng Yeo⁶ , Stuart C. Clarke^{1,2,4,9}  and David W. Cleary^{10*} 

Abstract

Klebsiella pneumoniae is a Gram-negative Enterobacteriaceae that is classified by the World Health Organisation (WHO) as a Priority One ESKAPE pathogen. South and Southeast Asian countries are regions where both healthcare associated infections (HAI) and community acquired infections (CAI) due to extended-spectrum β -lactamase (ESBL)-producing and carbapenem-resistant *K. pneumoniae* (CRKp) are of concern. As *K. pneumoniae* can also exist as a harmless commensal, the spread of resistance genotypes requires epidemiological vigilance. However there has been no significant study of carriage isolates from healthy individuals, particularly in Southeast Asia, and specially Malaysia. Here we describe the genomic analysis of respiratory isolates of *K. pneumoniae* obtained from Orang Ulu and Orang Asli communities in Malaysian Borneo and Peninsular Malaysia respectively. The majority of isolates were *K. pneumoniae* species complex (KpSC) 1 *K. pneumoniae* ($n = 53$, 89.8%). Four *Klebsiella variicola* subsp. *variicola* (KpSC3) and two *Klebsiella quasipneumoniae* subsp. *similipneumoniae* (KpSC4) were also found. It was discovered that 30.2% ($n = 16$) of the KpSC1 isolates were ST23, 11.3% ($n = 6$) were of ST65, 7.5% ($n = 4$) were ST13, and 13.2% ($n = 7$) were ST86. Only eight of the KpSC1 isolates encoded ESBL, but importantly not carbapenemase. Thirteen of the KpSC1 isolates carried yersiniabactin, colibactin and aerobactin, all of which harboured the *rmpADC* locus and are therefore characterised as hypervirulent. Co-carriage of multiple strains was minimal. In conclusion, most isolates were KpSC1, ST23, one of the most common sequence types and previously found in cases of *K. pneumoniae* infection. A proportion were hypervirulent (hvKp) however antibiotic resistance was low.

Keywords *Klebsiella pneumoniae*, Malaysia, Antimicrobial resistance, ESBL, Hypervirulent

*Correspondence:

David W. Cleary

d.w.cleary@bham.ac.uk

¹Faculty of Medicine, Institute for Life Sciences, University of Southampton, Southampton, UK

²NIHR Southampton Biomedical Research Centre, University Hospital Southampton Foundation NHS Trust, Southampton, UK

³School of Medicine, International Medical University, Kuala Lumpur, Malaysia

⁴Institute for Research, Development and Innovation, International Medical University, Kuala Lumpur, Malaysia

⁵School of Pharmacy, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia

⁶Centre for Research in Infectious Diseases and Biotechnology (CeRIDB), Faculty of Medicine, Universiti Sultan Zainal Abidin, Medical Campus, 20400 Kuala Terengganu, Terengganu, Malaysia

⁷Faculty of Applied Social Sciences, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Nerus, Terengganu, Malaysia

⁸Centre of Excellence in National Indigenous Pedagogy, Institute of Teacher Education Tengku, Ampuan Afzan Campus, Pahang, Malaysia

⁹Global Health Research Institute, University of Southampton, Southampton, UK

¹⁰Institute of Microbiology and Infection, College of Medical and Dental Sciences, University of Birmingham, University of Birmingham, UK



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

For indigenous communities across the globe there is no doubt that there continues to exist disparities in susceptibility to, and burden from various diseases [1]. Marginalisation, resulting in inadequate access to health and social care infrastructure, often results in these communities being at risk of higher infant mortality and reduced life expectancy [1]. Whilst the term ‘indigenous’ is not prescribed as a term to encompass all such peoples, examples of ongoing burden in Canadian First Nations, Métis, and Inuit communities [2] and Australian Aboriginals [3] are all too easy to find. The Orang Ulu and Orang Asli are considered the indigenous peoples of Malaysia, living traditional lifestyles with similar socio-economic and health burdens [4, 5]. Typically, these communities live in poverty which translates into a life expectancy difference of ~20 years compared to the rest of the Malaysian population [6]. Given the distinct challenges that indigenous communities encounter, it is therefore critical that specific and targeted attention is given to them with respect to the infectious disease burden they may face [7]. To that end we draw our attention to *Klebsiella pneumoniae*.

K. pneumoniae is a Gram-negative Enterobacteriaceae that is classified by the World Health Organisation (WHO) as a Priority One ESKAPE pathogen [8]. As a major threat to public health, it remains a common nosocomial pathogen causing serious healthcare associated infections (HAI) including bacteraemia, pneumonia, and sepsis. Similarly to many infectious diseases the burden is shouldered most by those at the extremes of age, as well as the immunocompromised [9].

The *K. pneumoniae* species complex (KpSC) comprises seven phylogroups, based on genomic relatedness: *K. pneumoniae* (KpSC1), *K. quasipneumoniae* subsp. *quasipneumoniae* (KpSC2), *K. variicola* subsp. *variicola* (KpSC3), *K. quasipneumoniae* subsp. *similipneumoniae* (KpSC4), *K. variicola* subsp. *tropica* (KpSC5), *K. quasivariicola* (KpSC6), and *K. africana* (KpSC7) [10, 11]. Whilst KpSC1 is the predominant phylogroup responsible for global disease [12], this diversity in the species complex has hampered efforts to generate effective vaccines and therapeutics. Moreover, the emergence of multidrug resistance to carbapenems, broad-spectrum β -lactams, fluoroquinolones, and aminoglycosides pose a significant challenge. Consequently, genomic surveillance, to monitor for example the spread of problematic clonal complexes such as CC258 [13], is an absolute requirement for this important human pathogen.

As *K. pneumoniae* can also exist as a harmless commensal, the prevalence of resistant genotypes within these reservoirs also requires epidemiological vigilance. To date there has been no significant study of respiratory carriage isolates from healthy individuals in South-east Asia, and especially Malaysia. However, broadly,

South and Southeast Asian countries are regions where both HAI and community acquired infections (CAI) due to extended-spectrum β -lactam (ESBL)-producing and carbapenem-resistant *K. pneumoniae* (CRKp) are of particular concern [14–16]. A recent study of blood stream infections, involving seven major hospitals in South and Southeast Asian countries, found 17% of *K. pneumoniae* strains were carbapenemase producers and 47% of the strains carried markers for ESBL. Additional concerns were raised due to the high prevalence of aerobactin synthesis locus (*iuc*) in association with ESBL and/or carbapenemases [14].

Here we report the genomic epidemiology of *K. pneumoniae* that were taken during an all-age, upper respiratory tract carriage study in partnership with Orang Ulu communities in Sarawak, Malaysian Borneo and Orang Asli communities in Peninsular Malaysia.

Methods

Study population

Isolation of *K. pneumoniae* was done during an all-age, Orang Ulu community carriage study in April 2016 in Sarawak, Malaysian Borneo (isolates collected between 09/04/2016 and 01/05/2016) and a similar study in Orang Asli communities in Peninsular Malaysia in August 2017 (isolates collected between 01/08/2017 and 05/08/2017). These studies have both been published separately without the data presented herein [17, 18]. Briefly, in Sarawak, the largest state in Malaysian Borneo, four rural long-house communities and one village were visited. These varied in both isolation and affluence, from Rumah Bana the most affluent and only 30 km from a nearby town, to Ba Marong which was located in dense forest 157 km from Long Lama. The reported ethnicities of the total population recruited for the initial study ($n=140$) encompassed Iban, Kelabit, Kenyah and Penan. In Terengganu, situated in the north-east of peninsular Malaysia two villages were visited from which $n=130$ participants were recruited. All participants were reported as Orang asli. There were no exclusion criteria for recruitment. Participants were asked to complete a questionnaire requesting demographic data and medical history such as recent respiratory tract infections, and any history of use of antimicrobials and immunisation status.

Klebsiella sp. isolates

Participants received whole mouth (WMS), oropharyngeal (OP), nasopharyngeal and/or nasal swabbing (N). Swabs were either rayon tipped Transwab® Perna-sal Amies with charcoal (Medical Wire and Equipment, Corsham, UK) for paediatric NP, or viscose tipped sterile Amies swabs with charcoal (Deltalab, Chalgrove, UK) for WMS, OP and N. Initial culture, performed at the University of Southampton (UK) following transportation,

was done as outlined previously [17, 18], but briefly swabs were plated onto multiple media for the purposes of isolating common respiratory pathogens. These included: CBA (Columbia blood agar with horse blood), CHOC (Columbia blood agar with chocolate horse blood), CNA (Columbia Blood Agar with Colistin and Naladixic Acid), BACH (Columbia Agar with Chocolate Horse Blood and Bacitracin), GC (Lysed GC Selective Agar) and *Pseudomonas* CFC Selective agar (all Oxoid, UK). Suspected *K. pneumoniae*, those being mucoid coliforms on CBA plates, were sub-cultured onto CLED agar (Oxoid, UK) for confirmation using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF) using a Microflex[®] LRF (Bruker, UK). Where multiple isolates were taken from one individual these were included regardless of the niche from which they were taken i.e., if a *Klebsiella* sp. was isolated from the nose and mouth of a participant, both were stored. These were used to determine frequency of multi-strain carriage based on genomic analyses.

Antibiotic susceptibility

K. pneumoniae was spread over Mueller-Hinton agar plates (MH, Oxoid, UK). Antibiotic discs (Oxoid, UK) were added before plates were incubated at 35±1°C for 18 h (±2 h). *K. pneumoniae* were tested with cefotaxime (5 µg), ciprofloxacin (5 µg), meropenem (10 µg) and ceftazidime (10 µg) antibiotic discs. Susceptibility was determined against EUCAST Clinical Breakpoint guidelines v6.0.

DNA extraction

DNA was extracted from a sweep of growth using a QIAmp DNA Mini kit (Qiagen, UK), as per the manufacturer's instructions. The concentration of genomic DNA was determined using Qubit 2.0 fluorometric quantification (Thermo-Fisher, UK).

Whole Genome Sequencing: Library preparation was done using the Nextera XT DNA kit (Illumina, UK) following the manufacturer's instructions. Briefly DNA was quantified with Qubit fluorometric quantification (ThermoFisher, UK) and diluted to 0.2 ng/µl. Following dilution, 1 ng DNA was tagged with the Nextera XT transposome. Tagmented libraries were amplified with 12 cycles of PCR and dual-indexed primers. Libraries were cleaned and size selected using 0.5× volume AMPure XP beads (Beckman Coulter™, Fisher Scientific). Library concentrations were normalised using the bead-based normalisation protocol implemented in the Nextera XT kit, then normalised libraries were pooled in equal volumes. Sequencing to generate 2×250 bp paired end reads was done on a MiSeq (Illumina, UK) using the 500 cycle v2 reagent kit.

Sequence quality control and assembly

FASTQC v0.11.3 (<https://github.com/s-andrews/FastQC>) with MULTIQC v1.9 [19] was used to perform quality assessment on paired-end reads. Kraken v7.3.0 [20], a kmer-based taxonomy classification tool, was used to check for contamination and confirm the isolates were *K. pneumoniae*. Paired-end reads were trimmed using trimmomatic v0.39 [21] and *de novo* assembled using SPAdes v3.14.1 [22]. Assemblies were iteratively improved using pilon v1.23 [23] and then Quast v5.0.2 was used to evaluate the assemblies [24].

Phylogenomics

Snippy v4.4.1 (<https://github.com/tseemann/snippy>), using snippy-multi, was used to generate a core SNP-based alignment using *Klebsiella pneumoniae* subsp. *pneumoniae* Ecl8 (accession: GCA_000315385.1) as a reference. The subsequent alignment was character-corrected using snippy-clean before recombination regions were identified and removed using the accompanying run_gubbins.py and snp-sites scripts, before FastTree was used to generate a phylogenetic tree using the GTR model of nucleotide substitution.

MLST, virulence and antibiotic gene identification/detection

Kleborate v1.0.0 [25, 26] was used to define *K. pneumoniae* species complex (KpSC), MLST, serotype predictions of K (capsule) and O antigen as well as a range of ICEKp and plasmid associated virulence loci in addition to antimicrobial resistance determinants.

Data analysis

All analysis was done in R version 3.6.0 (2019-04-26) using RStudio version 1.2.1335 [27, 28] with graphics built using the grammar of graphics package, ggplot [28]. The phylogenetic tree was visualised using the r package treedataverse, specifically ggtree [29].

Results

K. pneumoniae colonised participants

A total of sixty-seven isolates (from forty-eight individuals) were isolated from the two previously described carriage studies. Participant demographics and location of those from whom *K. pneumoniae* were isolated are shown in Table 1. Children represented only 8.4% ($n=4/48$) of this cohort, with most isolates recovered from those between the ages of 18 and 64 (35.4%; $n=17$). There were more female participants (54.2%) than male. Forty-two participants came from the locations in Sarawak, with the majority (37.5%; $n=18$) taken during the visit to Rumah Numpang, an isolated longhouse community located in Sebauh, Bintulu.

Table 1 Demographics of participants from whom *K. pneumoniae* were isolated

	Group	N (%)	
Age (Years)	< 5	1 (2.1)	
	5–17	3 (6.3)	
	18–49	17 (35.4)	
	50–64	16 (33.3)	
	65+	10 (20.8)	
	NA	1 (2.1)	
Gender	M	16 (33.3)	
	F	26 (54.2)	
	NA	2 (4.2)	
Location	Peninsular Malaysia	Kampung Sungai Pergam	4 (8.4)
		Kampung Berua	2 (4.2)
	Sarawak	Rumah Numpang	18 (37.5)
		Rumah Bana	12 (25.0)
		Long Nen	5 (10.4)
		Long Kerangan	3 (6.3)
		Ba Marong	2 (4.2)
		Kampung Sebir	2 (4.2)

Eight isolates were excluded from genome analysis after failing assembly quality checks. Four had N50s<1Mbp, three had N50s<1Mbp and genome sizes>6Mbp and a final isolate had a genome size>6Mbp and the assembly was highly fragmented (>1000 contigs). Of the remaining fifty-nine isolates the majority (89.8%; n=53) were KpSC1 *K. pneumoniae*, with four *Klebsiella variicola* subsp. *variicola* (KpSC3) and two *Klebsiella quasipneumoniae* subsp. *similipneumoniae* (KpSC4).

Twenty-six unique MLSTs were identified (Fig. 1). The majority of the STs were only observed once, with most isolates (n=29/59, 49.2%) belonging to only three STs: ST23 (n=16, 27.1%), ST86 (n=7, 11.9%) and ST65 (n=6, 10.2%). ST5584 and ST5585 were novel to this study.

Fourteen capsule (K) types were identified. Of these KL1 accounted for most at 27.1% (n=16/59), followed by KL2 (18.6%, n=11/59) and KL3 (8.5%, n=5/59). For twelve (20.3%) isolates no accurate capsule locus could be identified. It is likely these ‘unknown’ loci were due to lack of assembly contiguity or low/absence of coverage given that on average the locus evidence was present on between six and seven contigs and was typically missing between five to six genes.

Genomics of *K. pneumoniae* isolates

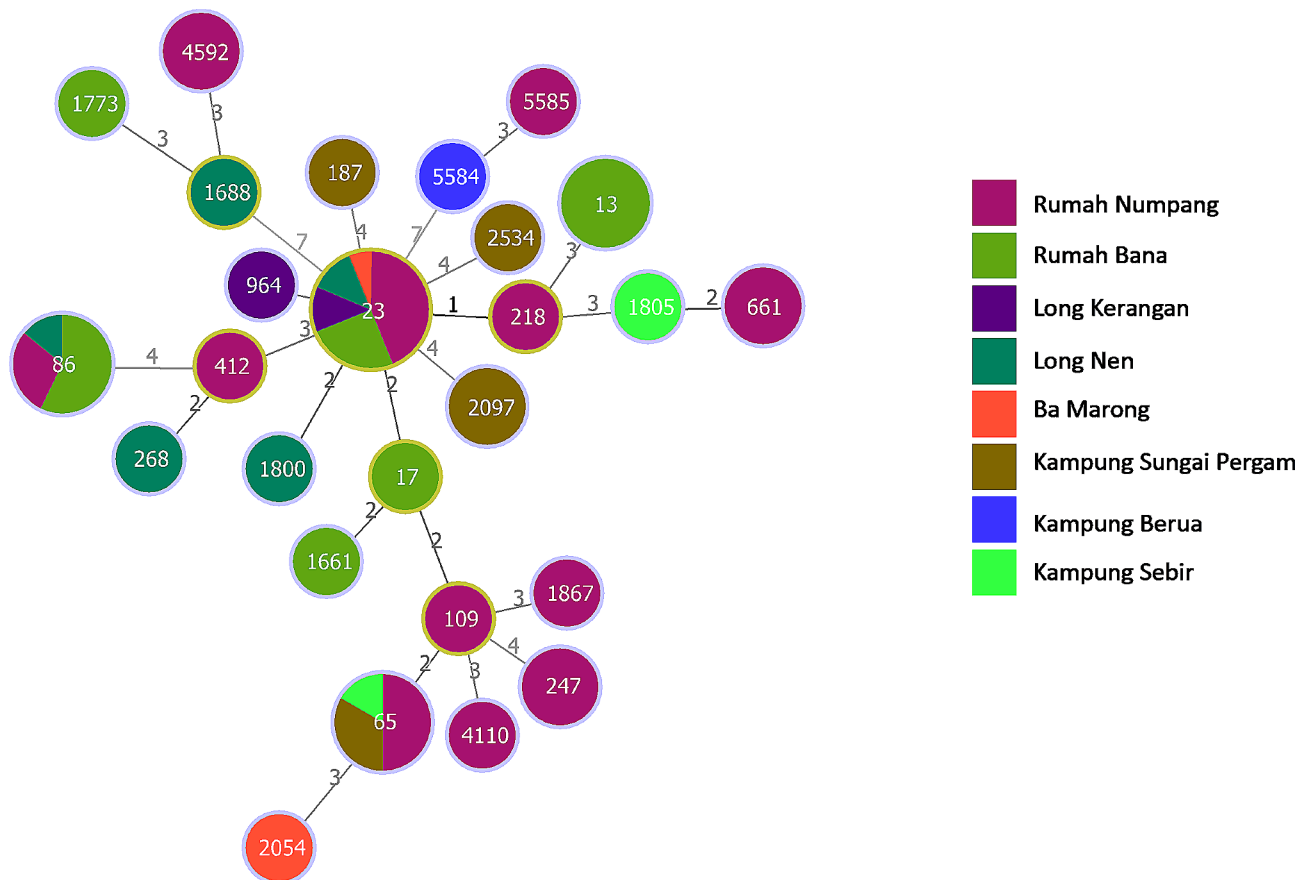


Fig. 1 Minimum-spanning tree generated using goeBurst and MLST. Node numbers are STs and edge numbers represent allelic differences. Node size is proportional to number of isolates belonging to that ST and are coloured by location of isolation

Virulence and antibiotic resistance

Acquired virulence gene and resistance scores are shown in Fig. 2. Only eight isolates, three of which were ST13 KL3, had a resistance score (Fig. 2A), and these were all ‘1’ indicating they were ESBL producers but without carbapenemases. This is in keeping with the phenotypic antibiotic resistance testing where all isolates were susceptible to all antibiotics tested. Resistance scores (Fig. 2B) were more varied. Those with the highest scores (5; indicating the presence of yersiniabactin, colibactin and aerobactin) were KL1 ST23 (*n*=10: *ybt* 1 *ICEKp*10, *clb* 2, *iuc* 1) and KL2 ST65 (*n*=2: *ybt* 17 *ICEKp*10, *clb* 3, *iuc* 1).

The distribution of O serotypes is shown in Fig. 3. O1/02 serotypes accounted for 66% of isolates (*n*=39) and were found associated with six capsule types (K1, K2, K3, K24, and K109).

Phylogeny of *K. pneumoniae* isolates

The three KpSCs, KpSC1, 3 and 4, are clearly visible on the isolate phylogeny with KpSC1 being the largest cluster (Fig. 4). No distinction between geographic location or body site sampling can be seen. All isolates which had a virulence score of 5 (indicated by black squares for yersiniabactin, colibactin and aerobactin) also harboured the *rmpADC* locus (*n*=13) conferring a hypermuroid phenotype and would therefore be classified as hypervirulent *K. pneumoniae* (*hvKp*). These isolates in particular only harboured the chromosomally encoded β-lactamase (B.broad column, Fig. 4) and are therefore not considered a problem with respect to antimicrobial resistance. Moreover, the presence of resistance genes overall was very limited. Only eight isolates had ESBL of which only two (both KL24 ST661) had any of the virulence genes, which in this case was *ybt*. We also note the high proportion of isolates generally encoding *rmpA/A2* (*n*=29/60). All were KpSC1, with KL1 (*n*=16), KL2 (*n*=11) and KL57

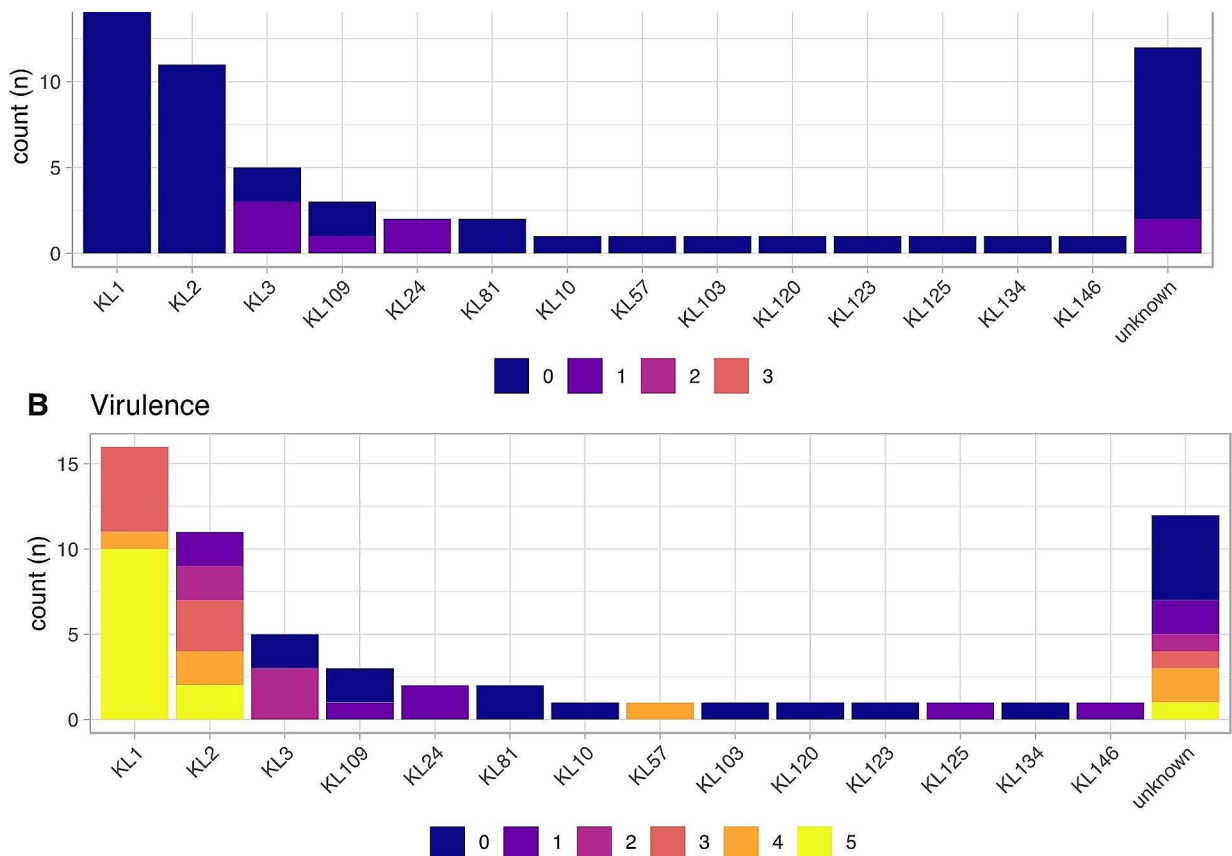


Fig. 2 Antibiotic resistance (A) and acquired virulence (B) scores for isolates with known capsular types. Resistance scores are on a scale of 0 (low) to 3 (high), and virulence 0 (low) to 5 (high). Resistance scores of 0 indicates no ESBL and no carbapenemase, 1 indicates the presence of an ESBL but still without carbapenemase, 2 indicates the presence of a carbapenemase without colistin resistance and 3 a carbapenemase with colistin resistance. Scores of 0/1 are irrespective of colistin resistance with scores of 2/3 irrespective of the presence of an ESBL. Virulence scores are based on the presence of yersiniabactin (*ybt*), colibactin (*clb*) and/or aerobactin (*iuc*). A score of 0 indicates none of these genes were found with a score of 5 showing all three were present. Scores of 1–4 indicate the following: 1 - yersiniabactin only, 2 - yersiniabactin and colibactin (or colibactin only), 3 aerobactin only, 4- aerobactin and yersiniabactin. Bars are ordered by frequency of capsule type

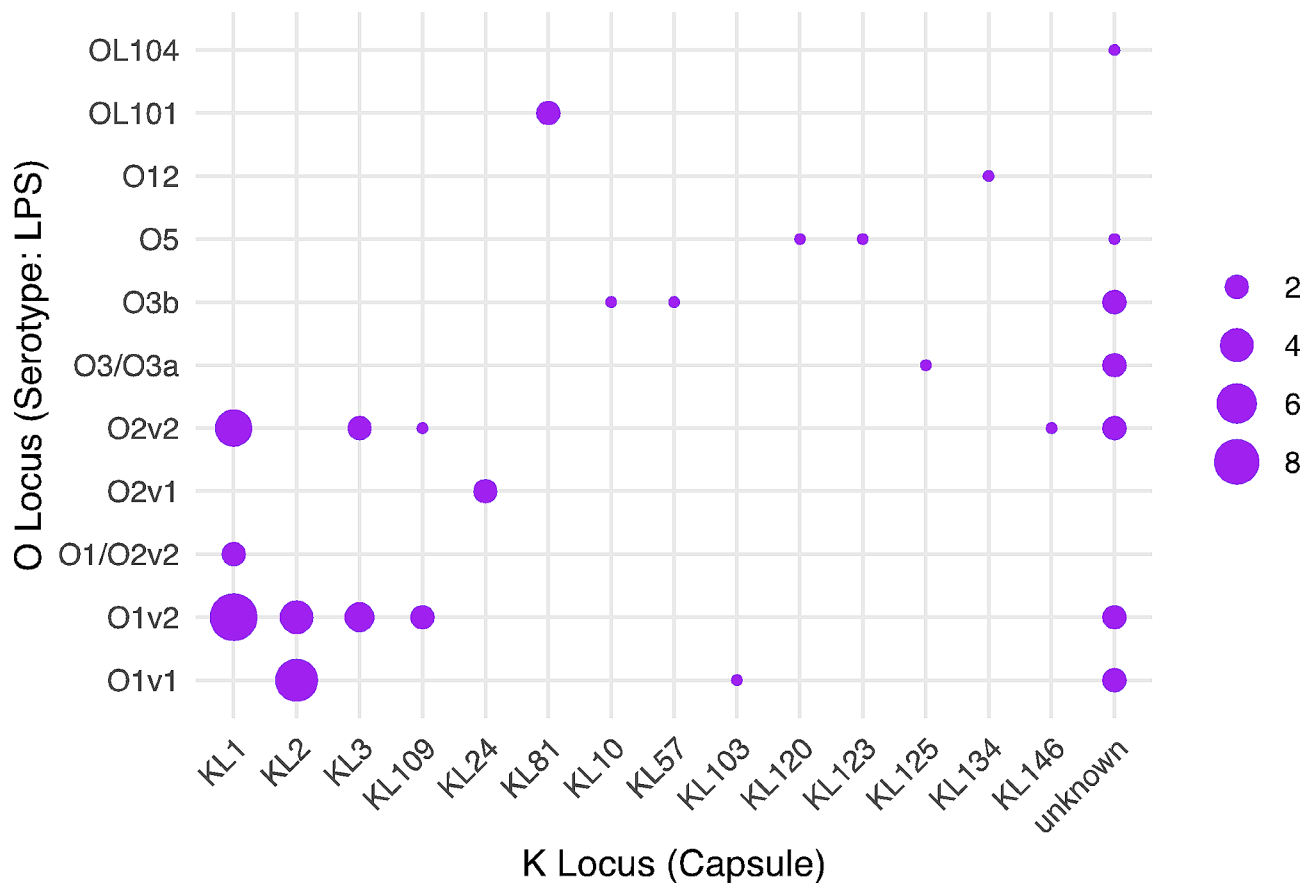


Fig. 3 Correlation between capsule (K) and serotype (O) loci. Each point displays the number of isolates within each K/O group. Most isolates were characterized by O1 or O2 polysaccharide which were found in KL1, KL2, KL3, KL109 and KL124 (bottom left quadrant)

($n=2$) being the only capsule locus types associated. All KL1 were ST23, whereas KL2 was split reasonably evenly between ST86 ($n=7$) and ST65 ($n=4$). All isolates were from adults, with the majority ($n=22/29$, 76%) coming from those over the age of 50-years-old. There was no gender separation with those coming from females and males in the same proportion (1:1.6) mirroring study recruitment. All but two isolates were from oropharyngeal/whole mouth swabs.

Carriage of multiple *K. pneumoniae* strains

We were particularly interested in the question of multiple strain carriage. Of the 18 participants from whom multiple isolations of *K. pneumoniae* were made, nine had isolates from both WMS and OP swabs, three from both OP and N and two from a N and NP sample. This excludes four participants for whom the genome assemblies of one or both isolates were not analysed as outlined above. Of those remaining, only two participants exhibited a multiple-strain carriage phenotype: an 80-year-old female from Rumah Bana had an ST86 KL2 in her OP sample and an ST17 with unknown capsule type in her nasal sample, and a 48-year-old female from Long Nen

harboured a ST1800 KL109 and a ST268 unknown capsule type in OP and WMS samples respectively.

Discussion

Countering the spectre of rising antimicrobial resistance in *K. pneumoniae* requires continued genomic epidemiological vigilance. Whilst there are efforts to fill the gap in knowledge regarding strains of this important human pathogen from Malaysia, specifically carbapenemase producers isolated in clinical settings [30, 31], there is a paucity of information related to carriage. Therefore, to add to this burgeoning field, we present the first genomic study of *K. pneumoniae* isolates from respiratory tract carriage, with a focus on an understudied and marginalised community of Malaysia. We highlight the high frequency of hypervirulent strains with, at the present time, minimal repertoires for antimicrobial resistance.

With limited regional data it is difficult to compare our isolates either with those from elsewhere in Malaysia, or within the Southeast Asian region, particularly with respect to respiratory carriage. Nevertheless, there are useful comparisons which may be made. For example, our finding that 19% of our isolates were hvKp is

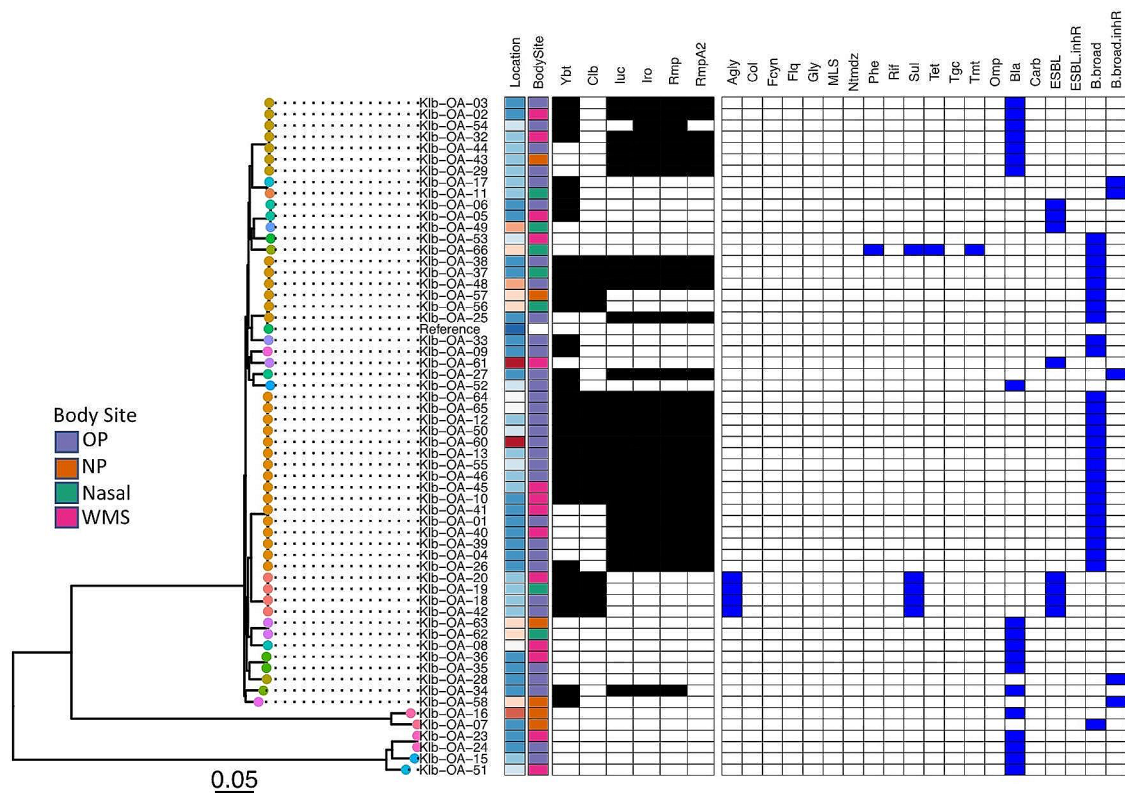


Fig. 4 Phylogeny of *K. pneumoniae*. Phylogenetic tree based on core-genome SNPs and constructed using FastTree (GTR+GAMMA). Leaves are colored by ST. The presence of virulence genes are shown in black, with antibiotic resistance markers in blue (white indicates absence). Virulence: Ybt yersiniabactin, Clb colibactin, luc aerobactin, Iro salmochelin, Rmp and RmpA2 hypermucoidy. Antibiotics: Agly aminoglycosides, Col colistin, Fcyn Fosfomycin, Flq fluoroquinolones, Gly glycopeptides, MLS macrolides, Ntmdz nitroimidazoles, Phe phenicols, Rif rifampin, Sul sulfonamides, Tet tetracyclines, Tgc tigecycline, Tmt trimethoprim, Omp osmoporin mutations, Bla beta-lactamases, Carb carbapenemase, ESBL extended spectrum beta-lactamases, ESBL(inhR) extended spectrum beta-lactamases with resistance to beta-lactamase inhibitors, B.broad beta-lactamases, B.broad.inhR beta-lactamases with resistance to beta-lactamase inhibitors

reasonably similar to the ~10% identified in studies of gut colonisation in a Taiwanese hospital setting and a study of healthy Chinese adults [32, 33]. In the latter study, KL1 and KL2 were also the most frequent capsule types observed, and that included in those Chinese adults residing in Malaysia [34]. This dominance of KL1/2 has also been noted in Singapore [35]. In keeping with the Singaporean epidemiology, the most common sequence type from our study was also ST23, which has been found to be associated with hypervirulence and antimicrobial resistance of *K. pneumoniae* in South East Asia [36] and previously identified as a common cause of disease in a collection of isolates from a Malaysian teaching hospital [37]. There are however important contrasts to note from the latter clinical study. There, *hvKp* accounted for a significantly greater proportion of isolates (38%), and MDR/ESBL was also more prevalent (31.9 and 27.8% respectively) [37]. Indeed, whilst ST23 was shared, ST22 and ST412, the next two most common STs were not found in the present study [37]. The increased proportion of

ESBL-producers in clinical samples was also observed in a two year study of hospital blood stream infections [38]. Here, 53% of isolates were classed as ESBL-producers ($n=45/303$) [38]. Our study provides a further stark contrast to the 87% prevalence of carbapenem-resistant *K. pneumoniae* clinical isolates that were observed over the course of one year at the University Malaya Medical Centre, Kuala Lumpur [31]. Our study is however in keeping with the global picture of associations between *hvKp* and antimicrobial resistance, in that those isolates which would be classed as *hvKp* only had limited resistance, genotypically, to β -lactamas without ESBL and carbapenemases [25]. This reflects perhaps an important difference between carriage and clinical strains, as according to the most recent National Antibiotic Resistance Surveillance Report from 2020, antibiotic resistance has been rising in clinical strains in Malaysia [39]. Regardless of these distinctions, a recent surveillance study of MDR and *hvKp* within low- and lower-middle income countries raised the importance of strain convergence

[40]. This phenomenon whereby a hypervirulent and carbapenem-resistant strain arises was noted in 2015 in China [41]. As such, and for other examples since, Asia was flagged as a region where convergence was of particular concern [40]. Whilst Malaysia was not one of the three countries highlighted, the circulation of both *hvKp*, as shown here, with the many examples above of clinical carbapenemase/ESBL-producing strains is cause for concern.

There are several limitations to this study. Firstly, the use of Illumina short-read sequencing and the resulting inability to accurately reconstruct K loci for all isolates would have been negated using long-read approaches. Further, no phenotypic analysis of capsule production was undertaken which would have clarified these discrepancies. The lack of accurate medical records makes extrapolation between the low AMR observed in these isolates and antibiotic use in these populations difficult. Our primary focus was respiratory carriage and therefore we do not know the carriage prevalence in the gut. This would be an important additional epidemiological reference point for future studies. Finally, our demographic is skewed towards an older population in only two locations. Despite these issues, perhaps the most important strength of our study is the focus on a marginalised, indigenous community. It is well documented that similar communities across the world are at increased risk of respiratory disease. Such examples include from the epidemiology of severe community acquired pneumonia in Australian Aboriginals [42] to increased incidence of invasive Group A Streptococci in First Nations populations in Alberta Canada [43] and *Staphylococcus aureus* disease in Native American individuals, where burden is many times the national average [44]. Whilst we have not determined *K. pneumoniae* disease burden in our Orang Ulu and Asli populations, this study is an important step in future attempts to do so.

In conclusion, we present the first study of carried respiratory *K. pneumoniae* from Malaysia. Currently, whilst there is clearly a reservoir of strains capable of causing disease, they at present do not harbour the genetics for resistance to therapeutic interventions.

Acknowledgements

Sequencing was performed by the Environmental Sequencing Facility at the University of Southampton to whom we are grateful. We would like to thank the Department of Orang Asli Affairs and Development (JAKOA). Additionally, the authors would especially like to thank the communities who partook for their support and participation in this study.

Author contributions

CCY, SCC and DWC conceived the study. CCY, SCC, DWC and VL secured funding. CCW, IKS, AGA, HS and RA were responsible for study planning and visits in Malaysia. RA, DEM, IKS and HS conducted visits for sampling. RA and DEM undertook microbiological culture and sample preparation for sequencing. SD, AKP and DWC undertook data analysis. SD and DWC wrote the manuscript. All authors reviewed the manuscript.

Funding

This work was funded by a Newton Fund Institutional Links award to Prof Stuart C. Clarke and Prof Mohd Nor Norazmi [grant number 172686537], an International Medical University grant awarded to Dr Chong Chun Wie and Dr Ivan Kok Seng Yap (BP I-01/12 (09) 2015), and two University of Southampton HEFCE Newton Fund Official Development Assistance (ODA) awards, one each to Prof Stuart C. Clarke and Dr David W. Cleary. Dr Cleary was supported by the National Institute for Health Research through the NIHR Southampton Biomedical Research Centre.

Data availability

The datasets generated and/or analysed during the current study are available in the European Nucleotide Archive (ENA) under project PRJEB51966 with accession numbers ERX9079753 to ERX9079811.

Declarations

Ethics approval and consent to participate

Ethical approval for isolates taken in Peninsular Malaysia was provided by Universiti Sultan Zainal Abidin (UniSZA) Ethics Committee: approval no. UniSZA/C/1/UHREC/628–1(85) dated 27 June 2016, the Department of Orang Asli Affairs and Development (JAKOA): approval no. JAKOA/PP.30.052.Jld11 [42], and by the University of Southampton Faculty of Medicine Ethics Committee (Submission ID: 20831). For those from Sarawak, approval was provided by the International Medical University Malaysia (IMU) Joint-Committee on Research and Ethics (Project ID No: IMU R 175/2015) and the University of Southampton Faculty of Medicine Ethics Committee (Submission ID: 14478). In all cases written informed consent was taken with parents/guardians providing consent for those < 18 years old.

Consent for publication

All authors have provided consent for publication.

Competing interests

DWC was a post-doctoral researcher on projects funded by Pfizer and GSK between April 2014 and October 2017. SCC acts as principal investigator on studies conducted on behalf of University Hospital Southampton NHS Foundation Trust/University of Southampton that are sponsored by vaccine manufacturers but receives no personal payments from them. SCC has participated in advisory boards for vaccine manufacturers but receives no personal payments for this work. SCC has received financial assistance from vaccine manufacturers to attend conferences. All grants and honoraria are paid into accounts within the respective NHS Trusts or Universities, or to independent charities. All other authors have no conflicts of interest.

Received: 3 October 2023 / Accepted: 1 April 2024

Published online: 17 April 2024

References

- Anderson I, Robson B, Connolly M, Al-Yaman F, Bjertness E, King A, et al. Indigenous and tribal peoples' health (the Lancet-Lowitja Institute Global Collaboration): a population study. *Lancet*. 2016;388(10040):131–57.
- Lee NR, King A, Vigil D, Mullaney D, Sanderson PR, Ametepée T, et al. Infectious diseases in indigenous populations in North America: learning from the past to create a more equitable future. *Lancet Infect Dis*. 2023;23(10):e431–44.
- Carville KS, Lehmann D, Hall G, Moore H, Richmond P, de Klerk N et al. Infection is the major component of the Disease Burden in Aboriginal and non-aboriginal Australian children: a Population-based study. *Pediatr Infect Dis J*. 2007;26(3).
- SyedHussain T, Krishnasamy DS, Hassan AAG. Distribution and demography of the Orang Asli in Malaysia. *Int J Humanit Social Sci Invention*. 2017;6(1):40–5.
- Yeo L-F, Aghakhaniyan FF, Tan JSY, Gan HM, Phipps ME. Health and saliva microbiomes of a semi-urbanized indigenous tribe in Peninsular Malaysia. *F1000Research*. 2019;8:175.
- Mohd Asri MN. Advancing the Orang Asli through Malaysia clusters of Excellence Policy. *J Int Comp Educ*. 2012;1(2).

7. Butler JC, Crengle S, Cheek JE, Leach AJ, Lennon D, O'Brien KL, et al. Emerging infectious diseases among indigenous peoples. *Emerg Infect Dis*. 2001;7(3 Suppl):554–5.
8. Shrivastava SR, Shrivastava PS, Ramasamy J. World health organization releases global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. *J Med Soc*. 2018;32(1):76.
9. Podschun R, Ullmann U. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev*. 1998;11(4):589–603.
10. Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. *Nat Rev Microbiol*. 2020;18(6):344–59.
11. Rodríguez-Medina N, Barrios-Camacho H, Duran-Bedolla J, Garza-Ramos U. *Klebsiella variicola*: an emerging pathogen in humans. *Emerg Microbes Infect*. 2019;8(1):973–88.
12. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D 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.
13. Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D, et al. Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS ONE*. 2015;10(7):e0133727.
14. Wyres KL, Nguyen TNT, Lam MMC, Judd LM, van Vinh Chau N, Dance DAB, et al. Genomic surveillance for hypervirulence and multi-drug resistance in invasive *Klebsiella pneumoniae* from South and Southeast Asia. *Genome Med*. 2020;12(1):11.
15. Shein AMS, Wannigama DL, Higgins PG, Hurst C, Abe S, Hongsing P, et al. High prevalence of mgrb-mediated colistin resistance among carbapenem-resistant *Klebsiella pneumoniae* is associated with biofilm formation, and can be overcome by colistin-EDTA combination therapy. *Sci Rep*. 2022;12(1):12939.
16. Wannigama DL, Sithu Shein AM, Hurst C, Monk PN, Hongsing P, Phattharapornjaroen P, et al. Ca-EDTA restores the activity of ceftazidime-avibactam or aztreonam against carbapenemase-producing *Klebsiella pneumoniae* infections. *iScience*. 2023;26(7):107215.
17. Cleary DW, Morris DE, Anderson RA, Jones J, Alattraqchi AG, Rahman A. The upper respiratory tract microbiome of indigenous Orang Asli in north-eastern Peninsular Malaysia. *Npj Biofilms Microbiomes*. 2021;7(1):1.
18. Morris DE, McNeil H, Hocknell RE, Anderson R, Tuck AC, Tricarico S, et al. Carriage of upper respiratory tract pathogens in rural communities of Sarawak, Malaysian Borneo. *Pneumonia*. 2021;13(1):6.
19. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047–8.
20. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*. 2014;15.
21. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
22. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a New Genome Assembly Algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77.
23. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE*. 2014;9(11):e112963.
24. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29(8):1072–5.
25. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat Commun*. 2021;12(1):4188.
26. Wyres KL, Wick RR, Gorrie C, Jenney A, Follador R, Thomson NR et al. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb Genomics*. 2016;2(12).
27. R. A language and environment for statistical computing. Computing, Vienna, Austria. 2017.
28. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2016.
29. Yu G. Using ggtree to visualize data on Tree-Like structures. *Curr Protocols Bioinf*. 2020;69(1):e96.
30. Gan HM, Eng WWH, Dhanoa A. First genomic insights into carbapenem-resistant *Klebsiella pneumoniae* from Malaysia. *J Global Antimicrob Resist*. 2020;20:153–9.
31. Lau MY, Teng FE, Chua KH, Ponnampalavanar S, Chong CW, Abdul Jabar K, et al. Molecular characterization of Carbapenem resistant *Klebsiella pneumoniae* in Malaysia Hospital. *Pathogens*. 2021;10(3):279.
32. Juan C-H, Chou S-H, Chen IR, Yang C-I, Lin Y-T, Chen L. Intestinal colonisation with hypervirulent or third-generation cephalosporin-resistant *Klebsiella pneumoniae* strains upon hospital admission in a general ward in Taiwan. *Int J Antimicrob Agents*. 2022;60(2):106624.
33. Lin Y-T, Siu LK, Lin J-C, Chen T-L, Tseng C-P, Yeh K-M, et al. Seroepidemiology of *Klebsiella pneumoniae* colonizing the intestinal tract of healthy Chinese and overseas Chinese adults in Asian countries. *BMC Microbiol*. 2012;12(1):13.
34. Cao RG, Suarez NM, Obermoser G, Lopez SM, Flano E, Mertz SE, et al. Differences in antibody responses between trivalent inactivated influenza vaccine and live attenuated influenza vaccine correlate with the kinetics and magnitude of interferon signaling in children. *J Infect Dis*. 2014;210(2):224–33.
35. Chew KL, Lin RTP, Teo JWP. *Klebsiella pneumoniae* in Singapore: Hypervirulent infections and the carbapenemase threat. *Front Cell Infect Microbiol*. 2017;7:515.
36. Gorrie CL, Mirceta M, Wick RR, Edwards DJ, Thomson NR, Strugnelli RA, et al. Gastrointestinal carriage is a Major Reservoir of *Klebsiella pneumoniae* infection in Intensive Care patients. *Clin Infect Dis*. 2017;65(2):208–15.
37. Mobasser G, Thong KL, Rajasekaram G, Teh CSJ. Molecular characterization of extended-spectrum β -lactamase-producing *Klebsiella pneumoniae* from a Malaysian hospital. *Brazilian J Microbiol*. 2020;51(1):189–95.
38. Subramaniam K, Khaithir TMN, Ding CH, Hussin NSC. Epidemiology of bloodstream infections in the paediatric population in a Malaysian general hospital over a 2-year period. *Malays J Pathol*. 2021;43(2):291–301.
39. Hashim R. National Antibiotic Resistance Surveillance Report 2020. Antibiotic Resistance Surveillance Reference Laboratory, Institute for Medical Research 2020 2018.
40. Silvester R, Madhavan A, Kokkat A, Parolla A, B M A MH, et al. Global surveillance of antimicrobial resistance and hypervirulence in *Klebsiella pneumoniae* from LMICs: an in-silico approach. *Sci Total Environ*. 2022;802:149859.
41. Zhang Y, Zeng J, Liu W, Zhao F, Hu Z, Zhao C, et al. Emergence of a hypervirulent carbapenem-resistant *Klebsiella pneumoniae* isolate from clinical infections in China. *J Infect*. 2015;71(5):553–60.
42. Tsai D, Chiong F, Secombe P, Hnin KM, Stewart P, Goud R, et al. Epidemiology and microbiology of severe community-acquired pneumonia in Central Australia: a retrospective study. *Intern Med J*. 2022;52(6):1048–56.
43. Tyrrell GJ, Bell C, Bill L, Fathima S. Increasing Incidence of Invasive Group A *Streptococcus* Disease in First Nations Population, Alberta, Canada, 2003–2017. *Emerg Infect Dis*. 2021;27(2):443–51.
44. Cella E, Sutcliffe CG, Tso C, Paul E, Ritchie N, Coleley J et al. Carriage prevalence and genomic epidemiology of *Staphylococcus aureus* among native American children and adults in the Southwestern USA. *Microb Genomics*. 2022;8(5).

Publisher's Note

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