Extended genomic analyses of the broad-host-range phages vB_KmiM-2Di and vB_KmiM-4Dii reveal slopekviruses have highly conserved genomes

Smith-Zaitlik, Thomas; Shibu, Preetha; McCartney, Anne L.; Foster, Geoffrey; Hoyles, Lesley; Negus, David

DOI:
10.1101/2022.04.06.486684

First published: 06/04/2022

Document Version
Early version, also known as preprint

Extended genomic analyses of the broad-host-range phages vB_KmiM-2Di and vB_KmiM-4Dii reveal slopekviruses have highly conserved genomes

Thomas Smith-Zaitlik[1], Preetha Shibu[2]†, Anne L. McCartney[3], Geoffrey Foster[4], Lesley Hoyles[1], David Negus[1]

ORCiD information

Preetha Shibu: 0000-0001-7279-2236

Geoffrey Foster: 0000-0002-5527-758X

Lesley Hoyles: 0000-0002-6418-342X

David Negus: 0000-0001-9047-4565

Affiliation

[1] Department of Biosciences, Nottingham Trent University, UK

[2] Life Sciences, University of Westminster, UK

[3] Department of Food and Nutritional Sciences, University of Reading, UK

[4] SRUC Veterinary Services, Inverness, UK

† **Present address:** Berkshire and Surrey Pathology Services, Frimley Health NHS Trust, Wexham Park Hospital, Slough, UK

Corresponding authors

Lesley Hoyles, lesley.hoyles@ntu.ac.uk; David Negus, david.negus@ntu.ac.uk

Keywords

Slopekvirus, phage diversity, lytic phage, myovirus, pangenome, *Klebsiella oxytoca* complex, homing endonuclease

Repositories: The sequences for the two phage genomes described herein have been deposited in DDBJ/ENA/GenBank under accession numbers MZ707156 (vB_KmiM-2Di) and MZ707157 (vB_KmiM-4Dii) (BioProject PRJNA750911). Supplementary data and material associated with this article are available from figshare at https://figshare.com/projects/KP15-like_bacteriophages/127976.

Abbreviations: AAHC, antibiotic-associated haemorrhagic colitis; AMR, antimicrobial resistance; ANI, average nucleotide identity; HEN, homing endonuclease; KoC, *Klebsiella oxytoca* complex; MAG, metagenome-assembled genome; MALDI-TOF MS, matrix-assisted laser desorption/ionisation-time of flight mass spectrometry; MLST, multi-locus sequence typing; MSA, multiple-sequence alignment; NA, nutrient agar; NB, nutrient broth; TEM, transmission electron microscopy; UTI, urinary tract infection.

ABSTRACT

High levels of antimicrobial resistance among members of the *Klebsiella oxytoca* complex (KoC) have led to renewed interest in the use of bacteriophage (phage) therapy to tackle infections caused by these bacteria. In this study we characterized two lytic phages, vB_KmiM-2Di and vB_KmiM-4Dii, that were isolated from sewage water against two GES-5-positive *Klebsiella michiganensis* strains (PS_Koxy2 and PS_Koxy4, respectively). ViPTree analysis showed both phages belonged to the genus *Slopekvirus*. *rpoB* gene-based sequence analysis of 108 presumptive *K. oxytoca* isolates (*n*=59 clinical, *n*=49 veterinary) found *K. michiganensis* to be more prevalent (46 % clinical and 43 % veterinary, respectively) than *K. oxytoca* (40 % clinical and 6 % veterinary, respectively). Host range analysis against these 108 isolates found both vB_KmiM-2Di and vB_KmiM-4Dii showed broad lytic activity against KoC species. Several putative homing endonuclease genes were encoded within the genomes of both phages, which may contribute to their broad host range. Pangenome analysis of 24 slopekviruses found that genomes within this genus are highly conserved, with more than 50 % of all predicted coding sequences representing core genes at ≥95 % identity and ≥70 % coverage. Given their broad host ranges, our results suggest vB_KmiM-2Di and vB_KmiM-4Dii represent attractive potential therapeutics. In addition, current recommendations for phage-based pangenome analyses may require revision.

INTRODUCTION

Members of the *Klebsiella oxytoca* complex (KoC) are divided into phylogroups based on the sequence of their chromosomally encoded β-lactamase (bla_{OXY}) gene. The current phylogroups are *Klebsiella michiganensis* (KoI, with KoV sub-lineage), *K. oxytoca* (KoII), *K. spallanzanii* (KoIII), *K. pasteurii* (KoIV), *K. grimontii* (KoVI) and *K. huaxiensis* (KoVIII). KoVII has been described based on a single isolate (1–3).

Several members of the KoC can cause a variety of infections in humans including urinary tract infections (UTIs), septicaemia and *Clostridioides*-negative antibiotic-associated haemorrhagic colitis (AAHC) (4–6). The rapid development of antimicrobial resistance (AMR) and the lack of novel antibiotics is a serious public health concern. Of 41 strains of the KoC isolated from bloodstream infections in the UK and Ireland, 100 % were phenotypically resistant to amoxicillin and cefuroxime, 75.6 % to piperacillin-tazobactam, 73.2 % to amoxicillin-clavulanate and 48.8 % to ciprofloxacin (7). In a survey of 5,724 clinical isolates of *K. oxytoca*, the SENTRY Antimicrobial Surveillance Program identified rates of non-susceptibility of *K. oxytoca* to various antibiotics: 1.8 % carbapenems, 12.5 % ceftriaxone, 7.1 % ciprofloxacin, 0.8 % colistin and 0.1 % tigecycline (8). GES-positive clinical strains of the KoC have also been identified recently (9,10).

KoC bacteria can also cause disease in animals. Of 336 samples collected from companion animals, 11 (3.3 %) isolates were identified as *K. oxytoca*. These were typically recovered from the urogenital system and 81.8 % were resistant to ampicillin (11). *Klebsiella* spp. were detected in 51/1541 (3.3 %) equine samples. Two *K. michiganensis* and one *K. oxytoca* were identified as non-repetitive cefotaxime-resistant isolates. These three isolates were phenotypically resistant to gentamicin, tobramycin, tetracycline, doxycycline, chloramphenicol and trimethoprim/sulfamethoxazole (12).

Accurate identification at the species level is important for recognising the epidemiological and clinical significance of each member of the KoC in both humans and animals. Current diagnostics are unable to consistently differentiate between members of the KoC leading to historic misidentification of non-*K. oxytoca* strains as *K. oxytoca* (10). The phenotypic similarity of the

complex members prevents accurate identification using biochemical tests such as API 20E tests (1). While matrix-assisted laser desorption/ionisation-time of flight mass spectrometry (MALDI-TOF MS) is a rapid and cost-effective diagnostic tool, many databases in clinical and veterinary use have not been updated and lack biomarkers required to consistently differentiate members of the KoC (13,14). The use of 16S rRNA gene sequencing is considered unsuitable for identifying *Klebsiella* at the species level as the gene sequence is so highly conserved in this taxon (15). *gyrA* and *rpoB* gene sequences have been used to distinguish *K. grimontii* and *K. huaxiensis* from other members of the KoC, supported by results from average nucleotide identity (ANI) analysis of genome sequence data (16,17). More recently, *rpoB* gene sequence analysis has been used to identify KoC, *Klebsiella pneumoniae* complex and *Raoultella* spp. isolates recovered from the faeces of healthy women and breast-fed infants (18). Multi-locus sequence typing (MLST) can also be used to assign strains to species of the KoC (10).

Bacteriophages (phages) are promising alternatives or adjuncts to current antibiotic therapies. We previously published an extensive review on *Klebsiella* phage and their potential as therapeutics (19). The majority of *Klebsiella* phage publications focus on *K. pneumoniae*. Phage vB_Kox_ZX8 was isolated from human faeces; this phage was shown to clear bacteraemia caused by a clinical strain of *K. oxytoca* in BALB/c mice (20). Another recent study described 30 novel phages that were active against *Klebsiella* species including KoC members *K. oxytoca* and *K. michiganensis* (21). Of the phages isolated, 15 were active against the single *K. michiganensis* strain tested, whereas 16 showed activity against at least one of the five *K. oxytoca* strains tested. More recently, the lytic *Drexlerviridae* phage KMI8 was isolated against *K. michiganensis* (22). KMI8 was lytic against 3/5 *K. michiganensis* strains but not *K. pneumoniae* (0/5) or *K. oxytoca* (0/1). To the best of our knowledge, ISF3 and ISF6 (23) and RP180 (24) are the only three *Raoultella* phages recorded in the literature. They were isolated against *Raoultella ornithinolytica* and *Raoultella* spp., respectively.

This study aimed to characterise the morphology, genomes and host ranges of two lytic phages, vB_KmiM-2Di and vB_KmiM-4Dii, isolated against two strains of GES-5-encoding *K. michiganensis* (PS_Koxy2 and PS_Koxy4) (25). We used *rpoB* gene sequence analysis to accurately identify 108 clinical and veterinary isolates identified as *K. oxytoca* using MALDI-TOF MS and/or API 20E. These isolates were used in our host range analysis, to determine the therapeutic potential of vB_KmiM-

2Di and vB_KmiM-4Dii against *Klebsiella* spp. A pangenome analysis was undertaken to compare our new phage genomes with those of their closest relatives.

METHODS

Strain information

Details of all strains included in this study can be found in **Table 1**.

Isolation of lytic phage

Filter-sterilised sewage samples (0.45 µm cellulose acetate filter; Millipore) collected from mixed-liquor tanks at Mogden Sewage Treatment Works (March 2017) were screened against *K. michiganensis* strains PS_Koxy2 and PS_Koxy4 (10,25). Firstly, 9 mL of filter-sterilised sewage were added to 1 mL of 10× concentrated sterile nutrient broth (NB) (Oxoid Ltd) containing 50 mM $CaCl_2$ and 50 mM $MgCl_2$. This was then inoculated with 200 µL of overnight culture from each strain and incubated for 6 h at 37 °C. The samples were centrifuged at 10,000 rpm for 5 min; the supernatants were aliquoted (200 µL) and used in spot assays to identify lytic phage. Plaques were propagated to purity to create phage stocks.

Growth media and culture conditions

Bacterial cultures were initially streaked onto MacConkey agar (Sigma Aldrich) to ensure purity before being grown on nutrient agar (NA) (Sigma Aldrich). NB (Sigma Aldrich) was used for overnight cultures, incubated aerobically at 37 °C. All media used for phage assays were supplemented with $CaCl_2$ and $MgCl_2$ (final concentration 0.5 mM) unless otherwise specified.

143 **Colony PCR and sequencing of *rpoB* gene products**

144 Forward (5'-GTTTTCCCAGTCACGACGTTGTAGGCGAAATGGCGGAAAACCA-3') and reverse (5'-
145 TTGTGAGCGGATAACAATTTCGAGTCTTCGAAGTTGTAACC-3') *rpoB*-specific primers (Macrogen) (23)
146 were diluted to 10 µM in DNase-free H_2O. A single colony for each isolate was touched using a sterile
147 loop and dipped into the PCR master mix: 25 µL MangoMix™ (Meridian Bioscience); forward and
148 reverse primers (0.5 µM final concentration); 20 µL DNase-free H_2O. DNA loading dye was included
149 in the MangoMix™. The positive control tube contained 2 µL of concentrated *K. michiganensis*
150 PS_Koxy1 DNA (10) and the negative control contained master mix alone (no template DNA). The
151 cycle conditions were: initial denaturation, 95 °C for 10 min; 35 cycles 95 °C for 30 s, 54 °C for 30 s,
152 72 °C for 1 min; final extension, 72 °C for 5 min. PCR products were checked for single bands of
153 expected size (1076 nt) using agarose gel electrophoresis (1 % agarose gel in 1x TAE buffer; 100 V, 40
154 min) against a GeneRuler 1 kb ladder (ThermoFisher Scientific).

155

156 The Thermo Scientific™ Gene JET PCR Purification Kit was used to clean PCR products. Purified
157 samples were checked for DNA concentration and purity using a NanoDrop™ 2000
158 spectrophotometer (24). Samples were then adjusted to 10 ng/µL and sent for sequencing (Source
159 BioScience) using the MLST forward primer (5'-GTTTTCCCAGTCACGACGTTGTA-3') (25).

160

161 **Phylogenetic analysis of *rpoB* gene sequences**

162 Returned *rpoB* gene sequences were trimmed to 501 nt using Geneious Prime (v2020.0.5) by
163 extracting the sequence between 276 and 776 nt, and with reference to the 45 *rpoB* allele
164 sequences (released 19 June 2020) available for download from the PubMLST *Klebsiella*
165 *oxytoca/michiganensis/grimontii* typing database (26). *rpoB* gene sequences were extracted from
166 the following genomes and used in analyses: *K. oxytoca* (GCA_900977765), *K. spallanzanii*
167 (GCA_901563875), *K. pasteurii* (GCA_901563825), *K. grimontii* (GCA_900200035), *K. michiganensis*
168 (GCA_901556995), *K. pneumoniae* subsp. *ozaenae* (GCA_000826585), *K. pneumoniae* subsp.
169 *rhinoscleromatis* (GCA_000163455), *K. pneumoniae* subsp. *pneumoniae* (GCA_000742135), *K.*
170 *quasipneumoniae* subsp. *similipneumoniae* (GCA_900978135), *K. quasipneumoniae* subsp.
171 *quasipneumoniae* (GCA_000751755), *K. africana* (GCA_900978845), '*K. quasivariicola*'
172 (GCA_000523395), *K. variicola* subsp. *tropica* (GCA_900978675), *K. variicola* subsp. *variicola*
173 (GCA_900977835), *K. aerogenes* (GCA_003417445), *K. indica* (GCA_005860775) and *K. huaxiensis*
174 (GCA_003261575). *Raoultella electrica* (GCA_006711645), *R. terrigena* (GCA_006539725), *R.*

planticola (GCA_000735435) and *R. ornithinolytica* (GCA_001598295) *rpoB* gene sequences were also included in analyses, as these taxa should be classified as *Klebsiella* spp. (27). A multiple-sequence alignment (MSA; available as **Supplementary Material**) was created using Clustal Omega (v1.2.2). The Jukes–Cantor genetic distance model was used to generate a neighbour-joining tree using the *rpoB* gene sequence of *K. aerogenes* ATCC 13048^T as an outgroup. The resulting newick file (available as **Supplementary Material**) was exported to iTOL (https://itol.embl.de/) (v6.1.1) (28)) for visualisation and annotation of the phylogenetic tree.

Host range analysis

Sterile molten top NA (3 mL; 0.2 % SeaPlaque Agarose, Lonza) supplemented with $CaCl_2$ and $MgCl_2$ (both at 5 mM) was aliquoted into sterile test tubes held at 45 °C. Each tube was then inoculated with 250 µL of an overnight culture of the prospective host strain, and gently swirled to mix the contents before being poured onto an NA plate. The plate was gently swirled to ensure even distribution of top agar. Once set, 5 µL aliquots of both phages were spotted onto the plate. Plates were incubated overnight at 37 °C. Next day, plates were inspected for lysis, with results recorded according to a modification of Haines *et al.* (32): ++, complete lysis; +, hazy lysis; 0, no visible plaques. We also noted whether depolymerase activity (d) was evident (i.e., formation of haloes around plaques).

Phage concentration

The Vivaspin 20 50 kDa centrifugal concentrator (Cytiva) was used to concentrate 20 mL of filter-sterilised propagated phage. Samples were spun at 3000 ***g*** until only 200 µL of sample remained. This concentrated phage stock was stored at 4 °C.

TEM

Formvar/carbon-coated 200 mesh copper grids (Agar Scientific) were prepared via glow discharge (10 mA, 10 s) using a Q150R ES sputter coater (Quorum Technologies Ltd). Phage suspensions (15 µL) were pipetted onto the grid surface for 30 s before removal using filter paper. Samples were stained using 15 µL of 2 % phosphotungstic acid. Excess stain was removed using filter paper and grids were

204 air-dried. Samples were visualized using a JEOL JEM-2100Plus (JEOL Ltd) TEM and an accelerating
205 voltage of 200 kV. Images were analysed and annotated using ImageJ (https://imagej.net/Fiji).

206

207 **Phage DNA extraction**

208 Nuclease-free H_2O was added to 200 µL of concentrated phage for a total final volume of 450 µL.
209 This was then incubated at 37 °C for 1.5 h with 50 µL of 10× CutSmart Buffer (New England BioLabs)
210 supplemented with 5 mM $CaCl_2$, 10 µL of DNase I (1 U/µL) (Thermo Scientific) and RNase A (10
211 mg/mL) (Thermo Scientific). Next, 20 µL of EDTA (final concentration 20 mM) and 1.3 µL of
212 Proteinase K (20 mg/mL) (Qiagen) were added and incubated at 56 °C for 1.5 h. The Qiagen DNeasy
213 Blood & Tissue Kit (Qiagen) was used to extract and purify the phage DNA. DNA was eluted in 20 µL
214 of the kit's AE buffer. Phage DNA integrity was checked using agarose gel electrophoresis (1 %
215 agarose gel in 1× TAE buffer; 70 V, 90 min) against a GeneRuler 1 kb ladder (ThermoFisher Scientific).

216

217 **Phage DNA sequencing, genome assembly and characterization**

218 Sequence data were generated on our in-house Illumina MiSeq platform. Extracted DNA was
219 adjusted to a concentration of 0.2 ng/µL and treated using the Nextera XT DNA library preparation
220 kit (Illumina) to produce fragments of approximately 500 bp. Fragmented and indexed samples were
221 run on the sequencer using a Micro flow cell with the MiSeq Reagent Kit v2 (Illumina; 150-bp paired-
222 end reads) following Illumina's recommended denaturation and loading procedures. Quality of raw
223 sequence data was assessed using FastQC v0.11.9. Reads had a mean phred score above 30 and
224 contained no adapter sequences, so data were not trimmed. Genomes were assembled using
225 SPAdes v3.13.0 (33), and visualized using Bandage v0.8.1 (34). Contamination and completeness of
226 genomes were determined using CheckV v0.8.1 (35). Genomes were screened for antimicrobial
227 resistance genes using the Resistance Gene Identifier (v5.2.0) of the Comprehensive Antibiotic
228 Resistance Database (v3.1.4) (36).

229

230 ViPTree v1.9 (37) was used to determine whether the phage genomes were closely related to
231 previously described double-stranded DNA viruses. Initial analyses showed them to be closely
232 related to *Klebsiella virus KP15* (genus *Slopekvirus*) (38). Other *Slopekvirus* genomes were identified
233 in GenBank and from the literature (**Table 2**) and included in a second ViPTree analysis. All publicly

234 available genome sequences were also compared against the genomes of *Klebsiella* phage
235 vB_KmiM-2Di and *Klebsiella* phage vB_KmiM-4Dii using pyani v0.2.11 (ANIm) (39).

236

237 **Analysis of homing endonucleases (HENs) encoded within phage genomes**

238 Genes in all genomes included in the initial ViPTree analysis (**Table 2**) were predicted and annotated
239 using Prokka 1.14.6 (34) using the PHROG (41) dataset. Data were imported into R using Biostrings
240 v2.58.0 (42) and predicted protein names searched for 'HNH|homing' to identify HENs encoded with
241 the phage genomes. The HEN sequences were exported in fasta format and imported into Geneious
242 Prime. An MSA was created using Clustal Omega v1.2.3 (options selected: group sequences by
243 similarity, evaluate full distance matrix, 5 refinement iterations). From the MSA, a neighbour-joining
244 tree (Jukes–Cantor) was generated. The tree was uploaded to iToL v6.4.2 (43) for annotation. All
245 data associated with the HEN analysis are available as **Supplementary Material**.

246

247 **Comparative genome analyses**

248 PhageClouds uses a gene-network approach to allow rapid searching of ~640,000 phage genomes via
249 a web interface, and includes metagenome-assembled genomes (MAGs) derived from several large-
250 scale metagenome and virome studies (44). Sequences of our initial set of GenBank genomes (**Table**
251 **2**) were searched against the PhageClouds database to identify slopekvirus genomes not included in
252 our original analysis, and to determine whether such genomes had been detected in
253 metagenome/virome studies. Results from PhageClouds searches were manually checked to identify
254 a non-redundant set of genomes potentially representing slopekviruses. Quality and completeness
255 of the genomes were determined using CheckV. ViPTree was used to confirm the newly identified
256 genomes fell within the genus *Slopekvirus*.

257

258 For all genomes found to be of high quality or complete (*n*=24; **Table 2**), VIRIDIC was used to
259 calculate the intergenomic similarities of the virus sequences (45). The 24 genomes were Prokka-
260 annotated as described above and included in a pangenome analysis (Roary v3.12.0, 95 % identity;
261 (46). Using treeio v1.18.1 (47) and ggtree v3.2.1 (48) a phylogenetic tree was generated from the
262 Roary-generated newick file (accessory_binary_genes.fa.newick), while the binary gene
263 presence/absence file (gene_presence_absence.Rtab) was used to visualize the core and accessory
264 genes identified in the slopekvirus pangenome. For each of the 24 genome sequences, amino acid

sequences of core genes with ≥95 % nucleotide identity and ≥70 % coverage (based on comparisons of minimum, average and maximum gene group sizes determined from Roary outputs) were concatenated. These sequences were then used to generate an MSA with MUSCLE v3.8.1551 (49). A maximum likelihood tree was generated from the MSA using PHYML v3.3.20180214 (BLOSUM62, 100 bootstraps; (50), and visualized using iToL v6.4.2. The Prokka-annotated genomes and outputs from Roary are available as **Supplementary Material**.

RESULTS

Morphological and genomic characterization of phages isolated on *K. michiganensis*

Two phages (vB_KmiM-2Di and vB_KmiM-4Dii) had been isolated and purified on *K. michiganensis* strains PS_Koxy2 and PS_Koxy4, respectively, during ongoing studies focussed on finding alternatives to antibiotics for treating *Klebsiella*-associated infections (25). In this study, genome sequence data and TEM images were generated for both phages. The genomes were found to be of high quality and free of contamination using CheckV (**Table 2**): the linear genome of vB_KmiM-2Di was 99.34 % complete, comprising 177,200 nt and encoding 275 genes; the linear genome of vB_KmiM-4Dii was 98.02 % complete, comprising 174,857 nt and encoding 271 genes. Neither phage encoded antimicrobial resistance genes. Initial BLAST-based and ViPTree analyses suggested the genomes represented members of the genus *Slopekvirus*. A ViPTree analysis, incorporating all slopekvirus genomes known to us at the time of analysis, confirmed this association (**Fig. 1a**). vB_KmiM-2Di clustered most closely with *Klebsiella virus PMBT1* and *Enterobacter virus Eap3*, whereas vB_KmiM-4Dii was most closely related to *Klebsiella* phage vB_KoM-MeTiny and *Klebsiella* phage vB_KoM-Pickle. ANIm analysis (**Fig. 1b**) showed vB_KmiM-2Di shared ≥99 % ANI with *Klebsiella virus PMBT1*. vB_KmiM-4Dii shared ≥99 % ANI with *Escherichia* phage phT4A, *Klebsiella* phage vB_KoM-MeTiny and *Klebsiella* phage vB_KoM-Pickle.

TEM images (**Fig. 1c, d**) showed both phages were myoviruses due to their contractile tails. Phage vB_KmiM-2Di (**Fig. 1c**) had a clearly visible base plate and long tail fibres. The tail and baseplate were recorded at 122.7 nm in length and the capsid diameter was 119.4 nm. The total length of the phage

was 242.1 nm. Phage vB_KmiM-4Dii (**Fig. 1d**) had a capsid diameter of 119.7 nm. The length of its tail and baseplate were recorded at 132.1 nm and the total length was 251.8 nm. Three short tail fibres could be seen attached at the bottom of the base plate and short whiskers protruded from the collar under the capsid.

Identification of clinical and veterinary KoC isolates

rpoB gene sequence data were generated for clinical (*n*=59), and veterinary (*n*=49) isolates previously identified as *K. oxytoca* using MALDI-TOF MS and API 20E tests (**Table 1**). These data were compared against the 45 *rpoB* reference allele sequences (released 19 June 2020) available for download from the PubMLST *Klebsiella oxytoca/michiganensis/grimontii* typing database and *rpoB* gene sequences of the type strains of *Klebsiella* (including *Raoultella*) spp. (**Supplementary Fig. 1**). This was done to confirm identity of isolates as we (and others) have previously shown that phenotypic tests and MALDI-TOF are often inadequate for characterization of KoC isolates (10).

Analysis of the sequence data revealed *K. michiganensis* was the most prevalent species represented in both clinical and veterinary isolates (46 % and 43 %, respectively; 27/59 clinical, 21/49 veterinary; **Fig. 2a**), with the *rpoB* gene sequences of all isolates clustering in the clade with the *rpoB* gene sequence of *K. michiganensis* W14^{T} (**Supplementary Fig. 1**). *K. oxytoca* was the second most prevalent species represented in the clinical isolates (40 %; 24/59), but only represented 6 % of veterinary isolates (3/49). *K. grimontii* was the second most-common bacterium among the veterinary isolates (24 %; 12/49) and third most-prevalent species in the clinical isolates (12 %; 7/59). *R. ornithinolytica* (12 %; 6/49), *K. huaxiensis* (4 %; 2/49) and *R. terrigena* (2 %; 1/49) were only represented in veterinary isolates. *K. pneumoniae* represented 2 % (1/59) of clinical isolates and 8 % (4/49) of veterinary isolates.

The majority (73 %; 36/49) of veterinary isolates were of bovine origin, isolated predominantly from milk-related samples (61 %; 30/49) (**Fig. 2b**). While isolates of *K. grimontii* (8/36), *K. huaxiensis* (2/36), *K. oxytoca* (1/36), *K. pneumoniae* (2/36), *R. ornithinolytica* (4/36) and *R. terrigena* (1/36) had been recovered from milk samples, the majority (44 %; 16/36) of the bovine isolates were *K. michiganensis* (**Fig. 2c**).

Host range analysis

vB_KmiM-2Di and vB_KmiM-4Dii were screened against clinical (*n*=59) and veterinary (*n*=49) isolates using an agar overlay method (**Table 3**). Phage activity against each strain was recorded as ++, complete lysis; +, hazy lysis; 0, no visible plaques. Depolymerase activity (d) was also recorded.

Both phages displayed a broad host range against clinical and veterinary isolates tested. vB_KmiM-2Di showed lytic activity 66 % (71/108) of tested strains (78 %, 46/59 clinical and 51 %, 25/49 veterinary strains, respectively). Whereas vB_KmiM-4Dii showed lytic activity against 84 % (91/108) strains (92 %, 54/59 clinical and 76 %, 37/49 veterinary strains, respectively). The two phages showed lytic activity against one or more strains of *K. michiganensis*, *K. oxytoca*, *K. grimontii*, *R. ornithinolytica*, *K. huaxiensis* and *R. terrigena*. Additionally, vB_KmiM-4Dii showed activity against *K. pneumoniae*.

The formation of haloes around plaques, indicating depolymerase activity, was observed for both phages against strains Ko13 and GFKo7. vB_KmiM-4Dii also displayed depolymerase activity against strain GFKo41. Neither vB_KmiM-2Di nor vB_KmiM-4Dii was able to infect 17 % (18/108) of the strains tested, the majority (80 %, 12/15) of which were of veterinary origin.

Analysis of HENs encoded within slopekvirus genomes

It has previously been suggested that differences in host range for members of the genus *Slopekvirus* may be due to HENs encoded within their genomes. These HENs may act as regulators of DNA modification and provide resistance to host restriction enzymes (51). Fifty-four putative HENs were encoded within 13/16 *Slopekvirus* genomes (**Fig. 3**). Several genomes encoded multiple putative HENs, with *Escherichia* phage phT4A (NC_055712) encoding nine. vB_KmiM-2Di encoded five HENs whereas vB_KmiM-4Dii encoded four. Of these, two had possible homologues in both genomes (**Supplementary Fig. 2**). MZ707156_00002 shared homology with MZ707157_00002; 79.7 % pairwise amino acid identity. MZ707156_00164 shared homologywith MZ707157_00163; 100 % pairwise amino acid identity.

Diversity of slopekviruses in GenBank and metagenomic data

PhageClouds was used to identify slopekviruses not included in our initial scoping of viruses related to phages vB_KmiM-2Di and vB_KmiM-4Dii. Each PhageClouds search generated 33 hits. **Supplementary Table 1** provides example output for a search done using the genome sequence of phage vB_KmiM-2Di. Seventeen of the 33 hits represented sequences included in our initial phylogenetic analysis (**Fig. 1a**), six represented isolated phages not included in our initial list of viruses and 11 represented MAGs recovered from virome studies (**Table 2**). ViPTree analysis showed all 17 viruses were related to the genus *Slopekvirus* (data not shown). CheckV analysis showed four of the GenBank and four of the MAG sequences represented complete or high-quality phage genomes (**Table 2**).

Our initial analyses (**Fig. 1**, **Supplementary Fig. 3**) showed that slopekvirus genomes shared high sequence similarity, with the HENs contributing to their diversity. Consequently, we used all 24 high-quality/complete slopekvirus genomes identified in this study (**Table 2**) in a pangenome analysis to determine whether core genes could be identified within the genus *Slopekvirus* (**Fig. 4**). It has been recommended that sequence and similarity coverage of proteins are set to >30 % identity and 50 % coverage, respectively, for genus-level phage-based pangenome studies (52). However, for the genus *Slopekvirus* we found these criteria were too lax (not shown). Our Roary-based pangenome analysis run at 95 % identity identified 155 core genes in a total pangenome of 425 genes (**Fig. 4a, b**; **Supplementary Table 2**). The pangenome was open, as the plot showing the total number of genes is not asymptotic (**Fig. 4c**). Filtering the core genes based on comparisons of minimum, maximum and average group nucleotide coverage (**Supplementary Table 2**) identified 148 core genes had ≥95 % identity and ≥70 % coverage. Given that the 24 genomes encoded a mean of 272 genes (+/- 5 genes) each, these core genes represented 54 % (148/272) of the total genome content of the slopekviruses. The majority (61/148, 41 %) of these genes were predicted to encode hypothetical proteins, with baseplate wedge subunit proteins (6/148, 4 %), proteins of unknown function (4/148, 3 %), tail tube (3/148, 2%) and RIIB lysis inhibitor, major head, head scaffolding, clamp loader of DNA polymerase, baseplate hub and 5'−3' deoxyribonucleotidase proteins (2/148, 1 % each) making the greatest contribution to the core genes (**Supplementary Fig. 4**). As expected, the HENs contributed to the accessory genes.

If two phage genome sequences, tested reciprocally, are more than 95 % identical at the nucleotide level over their full genome length, they are assigned to the same species (52). We, therefore,

sought to determine the species diversity within the genus *Slopekvirus*. VIRIDIC analysis of the 24 genome sequences included in our pangenome analysis suggested there were eight species represented within the genus (**Supplementary Table 3**), though bidirectional hierarchical clustering of these data showed no obvious clustering of the species (**Fig. 5a**). An MSA of the concatenated protein sequences for the 148 core genes (available as **Supplementary Material**) showed the total alignment of 32,904 aa shared a minimum identity of 97.16 % across all genomes (**Supplementary Table 4**). Maximum-likelihood phylogenetic analysis (**Fig. 5b**) of these concatenated sequences did not support species separation, nor did clustering of the accessory gene (binary) data in the pangenome analysis (**Fig. 4a**).

DISCUSSION

In this study we characterized the morphology and genomes of two phages (vB_KmiM-2Di and vB_KmiM-4Dii) we had isolated on two strains (PS_Koxy2 and PS_Koxy4, respectively) of multidrug-resistant *K. michiganensis* (10,25). Both phages were found to belong to the genus *Slopekvirus* and had the myoviridae-type morphology consistent with other known slopekviruses (**Fig. 1**). The host ranges of these phages were determined on a collection of clinical (*n*=59) and veterinary (*n*=49) bacteria that had been originally identified as belonging to the KoC (**Table 1**). Both phages were found to have broad host ranges (**Table 3**). This feature of the slopekviruses is thought to be due to the number of HENs encoded within their genomes (**Fig. 3**). Searches among ~640,000 phage genomes, using PhageClouds, allowed us to identify slopekviruses within metagenome and virome datasets. We also characterized the genomes of our phages and those of their closest relatives, undertaking a pangenome analysis to better understand the genomic diversity of slopekviruses.

Identification of members of the KoC from clinical and veterinary sources

Despite our clinical and veterinary isolates being presumptively identified as *K. oxytoca* by MALDI-TOF MS and API 20E profiling, *K. oxytoca* only represented 25 % (27/108) of the isolates identified by *rpoB* gene sequence analysis. While we have previously highlighted that phenotypic assays and MALDI-TOF MS frequently do not allow accurate identification of clinical isolates of the KoC (10,53), this is the first time we have encountered issues identifying veterinary KoC isolates. Veterinary

MALDI-TOF MS databases should be updated to include relevant reference spectra to allow differentiation of KoC species (1)

The *rpoB* gene-based sequence analysis identified *K. michiganensis* as the most common isolate from both clinical and veterinary samples (27/59 and 21/49 respectively). This supports our previous findings with respect to *K. michiganensis* being more clinically relevant than *K. oxytoca* (53). Gómez *et al*. also used *rpoB* gene sequence analysis to identify commensal, community-acquired and neonatal intensive care unit *Klebsiella* spp. isolates. Routine biochemical testing identified 21 *K. oxytoca*; however, *rpoB* gene analysis identified them as *K. michiganensis* (16/21), *K. grimontii* (5/21) and *K. pneumoniae* (1/21) (54). Taken together with our work, these findings reinforce the clinical relevance of *K. michiganensis* and its historic underappreciation as a human pathogen due to misidentification.

In recent years, *K. michiganensis* has been isolated from a diverse range of animals: farm animals (cows, poultry and pigs) (13,55), companion animals (cat, dog and horses) (12,55) and other animals including hedgehogs, guineapigs, mice, fruit bats, turtles and invertebrates (13,55–58). This is from using tools capable of discriminating between members of the KoC, but further work is needed to assess the true clinical and epidemiological significance of *K. michiganensis* in animals. Furthermore, many of these studies are retrospective, using accurate but time- and resource-intensive methods that would not be practical for rapid clinical or veterinary diagnosis. It is notable that 16/21 of our *K. michiganensis* veterinary isolates had been recovered from bovine milk samples (**Table 1**; **Fig. 2c**). Whether these isolates encode genetic determinants specific to mastitis, as observed for an acquired *lac* operon in bovine-associated *K. pneumoniae* (59), will be determined when we analyse genome sequence data for them in the future, with results to be reported elsewhere.

K. grimontii has previously been isolated from animals in Germany (cattle/milk, rabbit, pig, sheep, dog, pig, tortoise, hedgehog, roe deer) (13), and during a longitudinal study undertaken in Pavia, Italy (sheep, horse, fly, cattle, pig, cat, duck, turtle, dog, cockroach, wasp, chicken) (55). We found *K. grimontii* in seal (*n*=2), bovine (*n*=8), gecko (*n*=1) and sparrow (*n*=1) samples collected in the UK. *K. grimontii* appears to be the second most-common KoC species of veterinary relevance.

K. huaxiensis was originally described based on one isolate recovered from human urine in China (16). Since then, the bacterium has been isolated from cow and human faeces in Italy (1), and from cows (*n*=4), water (*n*=2), a horse (*n*=1) and hospital carriage (*n*=1) (55). In this study we identified two isolates (GFKo11, GFKo50) from bovine milk collected in Scotland. As for other members of the KoC, more work is needed to determine the wider relevance of *K. huaxiensis* to veterinary infections.

Host range determination for phages vB_KmiM-2Di and vB_KmiM-4Dii

We found that phages vB_KmiM-2Di and vB_KmiM-4Dii both exhibited broad host ranges against several different species of bacteria (**Table 3**). This includes isolates of *K. pneumoniae*, *K. michiganensis*, *K. oxytoca*, *K. grimontii*, *R. ornithinolytica*, *K. huaxiensis* and *R. terrigena*. Ordinarily, phage host range is narrow, sometimes down to the strain level. However broad-host-range phages are reported in the literature (60–62) and phages with extended host ranges have been identified within the genus *Slopekvirus* (21,38). For example, phage vB_KoM-MeTiny, which is genetically similar to vB_KmiM-2Di and vB_KmiM-4Dii (95.2 % and 96.2 % respectively; VIRIDIC), is reported to form plaques on *K. michiganensis*, *K. oxytoca*, *K. pneumoniae*, *K. variicola* and *K. quasipneumoniae*. These findings suggest bacteriophage belonging to the genus *Slopekvirus* are useful as potential therapeutics owing to their broad host ranges and lack of antimicrobial resistance genes (21,38), in agreement with others who have worked with slopekviruses (21,32).

Despite a high level of sequence identity between vB_KmiM-2Di and vB_KmiM-4Dii (95.25 %; VIRIDIC) differences in host range were observed between the two phages. Maciejewska *et al*. previously characterised two slopekviruses (vB_KpnM_KP15 and vB_KpnM_KP27) that exhibit broad lysis against *Klebsiella* spp. (63). They also noted differences in the host ranges of the two phages, despite a high-level of DNA identity (94.2 %; our VIRIDIC analysis). The authors suggested the discord in host range may be due to the presence of two HENs encoded within the genome of vB_KpnM_KP27, both of which are absent from vB_KpnM_KP15. HENs are site-specific DNA endonucleases that function as mobile genetic elements by catalysing a double-strand break at specific DNA target sites in a recipient genome that lacks the endonuclease. The double-strand break is then repaired by homologous recombination using the allele containing the HEN as the template. The result is the incorporation of the HEN into the cleavage site. Maciejewska *et al*. hypothesised that HENs encoded within vB_KpnM_KP27 (YP_007348875.1 and YP_007348891.1) may act as regulators of DNA modification by splicing events near DNA modification genes located in their close

vicinity. Such events may result in protection against host restriction enzymes and subsequently modulate host range. We identified a homologue of YP_007348875.1 (86 % pairwise amino acid identity with MZ707157_00254) and its associated flanking region in the genome of vB_KmiM-4Dii that is absent from vB_KmiM-2Di (**Supplementary Fig. 5**). The second HEN identified in KP27 (YP_007348891.1) was absent from both vB_KmiM-2Di and vB_KmiM-4Dii. The presence of YP_007348875.1 in the genome of vB_KmiM-4Dii may help explain its slightly broader host range compared to vB_KmiM-2Di.

We also conducted a search for other HENs encoded in other slopekvirus genomes. Despite a high degree of genetic conservation across the genus (**Fig. 4**, **Fig. 5**), there was a high level of divergence with respect to the number and type of HENs encoded across the 16 genomes and between vB_KmiM-2Di and vB_KmiM-4Dii. We found that vB_KmiM-2Di encodes five putative HENs, whereas vB_KmiM-4Dii encodes four. Of note was a HEN identified in the genomes of both vB_KmiM-2Di and vB_KmiM-4Dii (MZ707156_00164 and MZ707157_00163, respectively). This HEN was found immediately upstream of a CDS encoding a putative SbcC-like subunit of a predicted palindrome-specific endonuclease (**Supplementary Fig. 6**). This may again be a case of a HEN splicing event near a gene involved in DNA modification, with the result influencing host range. These findings lend support to the theory that HENs may influence host range and account for the differences in host lysis observed between vB_KmiM-2Di and vB_KmiM-4Dii.

It has been noted by others (21,64) that phages with highly similar sequences may have altered lytic spectra due to selection pressures applied by different hosts used for propagation. In the current study, *K. michiganensis* strains PS_Koxy2 and PS_Koxy4 were used for propagation of vB_KmiM-2Di and vB_KmiM-4Dii, respectively. These two strains have been extensively characterised both genomically and phenotypically and have been shown to share >99.9 % ANI and the same multi-locus sequence type (10). Therefore, we conclude that any differences observed with respect to host range are unlikely to be a consequence of the host used for propagation.

Genetic diversity of slopekviruses: implications for phage taxonomy

Our Roary-based pangenome analysis of 24 high-quality/complete slopekvirus genomes showed the genus *Slopekvirus* comprises phages with highly conserved genomes, with the core genome

(determined using ≥95 % identity and ≥70 % coverage) representing over half of the total genome content of the genus (**Fig. 4**). The recommended cut-off criteria (>30 % identity, 50 % coverage (52)) for genus-level phage pangenome analysis was found to be inappropriate for use with the genus *Slopekvirus*. Results from our analysis (along with other unpublished work from our laboratory) suggest wider-ranging studies are required to determine appropriate recommended identity and coverage cut-off values for use in phage-based pangenome studies at any taxonomic level.

The use of VIRIDIC (and other ANI tools) to assign phage that share ≥95 % reciprocal sequence identity at the nucleotide level over their full genome length to the same species is also questioned (52). Our VIRIDIC analysis of 24 high-quality/complete slopekvirus genomes suggested our dataset represented eight different species (**Fig. 5**). However, phylogenetic analysis based on the protein sequences of 148 core genes (encoding 32,904 aa) did not support separation of the genus into eight species. ANI analyses must be supplemented with phylogenetic analyses to provide robust evidence to support 'species' designations within phage genera, as recommended for bacterial and archaeal taxonomy (65).

AUTHOR STATEMENTS

Authors and contributors

Conceptualisation: PS, LH, DN. Data curation: all authors. Formal analysis: TSZ, PS, ALM, GF, LH, DN. Funding acquisition: PS, ALM, LH. Investigation: TSZ, PS, ALM, LH, DN. Methodology: ALM, LH, DN. Resources: GF, ALM, LH, DN. Supervision: ALM, LH, DN. Visualisation: TSZ, LH, DN. Writing – original draft: TSZ, LH, DN. Writing – reviewing and editing: all authors.

Conflicts of interest

The author(s) declare that there are no conflicts of interest.

Funding information

PS was in receipt of an IBMS Research Grant (project title "Isolation of lytic bacteriophages active against antibiotic-resistant *Klebsiella pneumoniae*"). Imperial Health Charity is thanked for contributing to registration fees for the Professional Doctorate studies of PS. LH was funded by UK Med-Bio (Medical Research Council grant number MR/L01632X/1). SRUC Veterinary Services receive funding from the Scottish Government as part of its Veterinary Advisory Services programme.

Ethical approval

The study of anonymised clinical isolates provided by the Nottingham University Hospitals NHS Trust (NUH) Pathogen Bank was approved by NUH Research and Innovation (19MI001).

Acknowledgements

This work used computing resources funded by the Research Contingency Fund of the Department of Biosciences, Nottingham Trent University (NTU). TSZ completed this work as part of an MRes degree at NTU. We thank Andrew Millard (University of Leicester) for providing us with the Prokka-formatted PHROG dataset for annotation of phage genomes.

555

REFERENCES

1. Merla C, Rodrigues C, Passet V, Corbella M, Thorpe HA, Kallonen TVS, et al. Description of *Klebsiella spallanzanii* sp. Nov. and of *Klebsiella pasteurii* sp. nov. Frontiers in Microbiology [Internet]. 2019 Oct 25 [cited 2021 Jun 13];10(OCT):2360. Available from: http://ggdc.dsmz.de

2. Granier SA, Leflon-Guibout V, Goldstein FW, Nicolas-Chanoine MH. New *Klebsiella oxytoca* β-Lactamase Genes *bla*OXY-3 and *bla*OXY-4 and a Third Genetic Group of *K oxytoca* Based on *bla*OXY-3. Antimicrobial Agents and Chemotherapy [Internet]. 2003 Sep 1 [cited 2021 Jun 21];47(9):2922–8. Available from: https://journals.asm.org/journal/aac

3. Fevre C, Jbel M, Passet V, Weill FX, Grimont PAD, Brisse S. Six Groups of the OXY β-Lactamase Evolved Over Millions of Years in *Klebsiella oxytoca*. Antimicrobial Agents and Chemotherapy [Internet]. 2005 Aug [cited 2021 Jun 21];49(8):3453–62. Available from: https://journals.asm.org/journal/aac

4. Youn Y, Lee SW, Cho H-H, Park S, Chung H-S, Seo JW. Antibiotics-Associated Hemorrhagic Colitis Caused by *Klebsiella oxytoca*: Two Case Reports. Pediatric Gastroenterology, Hepatology \& Nutrition [Internet]. 2018 Apr [cited 2021 Jan 26];21(2):141–6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5915692/

5. Abadi RM. Emerging Carbapenemase *Klebsiella oxytoca* with Multidrug Resistance Implicated in Urinary Tract Infection. Biomedical and Biotechnology Research Journal [Internet]. 2020 Apr 1 [cited 2021 Jun 13];4(2):148–51. Available from: https://www.bmbtrj.org/article.asp?issn=2588-9834;year=2020;volume=4;issue=2;spage=148;epage=151;aulast=AL-Khikani

6. England PH. MRSA, MSSA and Gram-Negative Bacteraemia and CDI: Annual Report [Internet]. 2020 Dec [cited 2021 Jan 18]. Available from: https://www.gov.uk/government/statistics/mrsa-mssa-and-e-coli-bacteraemia-and-c-difficile-infection-annual-epidemiological-commentary

7. Moradigaravand D, Martin V, Peacock SJ, Parkhill J. Population Structure of Multidrug Resistant *Klebsiella oxytoca* Within Hospitals Across the UK and Ireland Identifies Sharing of Virulence and Resistance Genes with *K pneumoniae*. Genome biology and evolution [Internet]. 2017 Mar 1 [cited 2021 Jun 15];9(3):574–87. Available from: https://academic.oup.com/gbe/article/9/3/574/2977333

8. Yang J, Long H, Hu Y, Feng Y, McNally A, Zong Z. *Klebsiella oxytoca Complex: Update on Taxonomy, Antimicrobial Resistance, and Virulence*. Clinical microbiology reviews [Internet]. 2022 Jan 19 [cited 2022 Feb 28];35(1). Available from: https://pubmed.ncbi.nlm.nih.gov/34851134/

9. Ellington MJ, Davies F, Jauneikaite E, Hopkins KL, Turton JF, Adams G, et al. A Multispecies Cluster of GES-5 Carbapenemase–Producing Enterobacterales Linked by a Geographically Disseminated Plasmid. Clinical Infectious Diseases [Internet]. 2020 Nov [cited 2021 Jan 27];71(10):2553–60. Available from: https://doi.org/10.1093/cid/ciz1130

10. Shibu P, McCuaig F, McCartney AL, Kujawska M, Hall LJ, Hoyles L. Improved Molecular Characterization of the *Klebsiella oxytoca* Complex Reveals the Prevalence of the Kleboxymycin Biosynthetic Gene Cluster. Microbial Genomics [Internet]. 2021 Jun 18 [cited 2021 Nov 15];7(6):000592. Available from: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000592

11. Lee D, Oh JY, Sum S, Park HM. Prevalence and Antimicrobial Resistance of *Klebsiella* Species Isolated from Clinically Ill Companion Animals. Journal of Veterinary Science [Internet]. 2021 Feb 9 [cited 2021 Jun 13];22(2):1–13. Available from: https://vetsci.org

12. Loncaric I, Rosel AC, Szostak MP, Licka T, Allerberger F, Ruppitsch W, et al. Broad-Spectrum Cephalosporin-Resistant *Klebsiella* spp. Isolated from Diseased Horses in Austria. Animals [Internet]. 2020 Feb 1 [cited 2021 Jun 13];10(2):332. Available from: www.mdpi.com/journal/animals

13. Klaper K, Hammerl JA, Rau J, Pfeifer Y, Werner G. Genome-Based Analysis of *Klebsiella* spp. Isolates from Animals and Food Products in Germany, 2013–2017. Pathogens [Internet]. 2021 May 8 [cited 2021 Jun 16];10(5):573. Available from: https://doi.org/10.3390/pathogens10050573

14. Bridel S, Watts SC, Judd LM, Harshegyi T, Passet V, Rodrigues C, et al. Klebsiella MALDI TypeR: A Web-Based Tool for *Klebsiella* Identification Based on MALDI-TOF Mass Spectrometry. Research in Microbiology. 2021 May 15;103835.

15. Boye K, Hansen DS. Sequencing of 16S rDNA of *Klebsiella*: Taxonomic Relations Within the Genus and to Other Enterobacteriaceae. International Journal of Medical Microbiology. 2003 Jan 1;292(7–8):495–503.

16. Hu Y, Wei L, Feng Y, Xie Y, Zong Z. *Klebsiella huaxiensis* sp. nov., Recovered from Human Urine. International Journal of Systematic and Evolutionary Microbiology [Internet]. 2019 Feb 1 [cited 2021 Jun 16];69(2):333–6. Available from: https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.003102

17. Passet V, Brisse S. Description of *Klebsiella grimontii* sp. nov. International Journal of Systematic and Evolutionary Microbiology [Internet]. 2018 Jan 1 [cited 2021 Jun 16];68(1):377–81. Available from: https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.002517

18. Gómez M, Valverde A, del Campo R, Rodríguez JM, Maldonado-Barragán A. Phenotypic and Molecular Characterization of Commensal, Community-Acquired and Nosocomial *Klebsiella* spp. Microorganisms [Internet]. 2021 Nov 12 [cited 2021 Dec 1];9(11):2344. Available from: https://pubmed.ncbi.nlm.nih.gov/34835469/

19. Herridge WP, Shibu P, O'Shea J, Brook TC, Hoyles L. Bacteriophages of *Klebsiella* spp., Their Diversity and Potential Therapeutic Uses. Journal of Medical Microbiology [Internet]. 2020 Feb [cited 2021 Jan 22];69(2):176–94. Available from: /pmc/articles/PMC7431098/

20. Li P, Zhang Y, Yan F, Zhou X. Characteristics of a Bacteriophage, vB_Kox_ZX8, Isolated From Clinical *Klebsiella oxytoca* and Its Therapeutic Effect on Mice Bacteremia. Frontiers in microbiology [Internet]. 2021 Dec 3 [cited 2022 Feb 28];12. Available from: https://pubmed.ncbi.nlm.nih.gov/34925270/

21. Townsend EM, Kelly L, Gannon L, Muscatt G, Dunstan R, Michniewski S, et al. Isolation and Characterization of *Klebsiella* Phages for Phage Therapy. https://home.liebertpub.com/phage [Internet]. 2021 Mar 17 [cited 2021 Jul 18];2(1):26–42. Available from: https://www.liebertpub.com/doi/abs/10.1089/phage.2020.0046

22. Ku H, Kabwe M, Chan HT, Stanton C, Petrovski S, Batinovic S, et al. Novel *Drexlerviridae* Bacteriophage KMI8 with Specific Lytic Activity Against *Klebsiella michiganensis* and its Biofilms. PLoS ONE [Internet]. 2021 Sep 1 [cited 2021 Nov 10];16(9). Available from: /pmc/articles/PMC8423285/

23. Zamani I, Bouzari M, Emtiazi G, Ghasemi SM, Chang H-I. Molecular Investigation of Two Novel Bacteriophages of a Facultative Methylotroph, *Raoultella ornithinolytica*: First Report of *Raoultella* Phages. Archives of Virology 2019 164:8 [Internet]. 2019 May 20 [cited 2021 Aug 19];164(8):2015–22. Available from: https://link.springer.com/article/10.1007/s00705-019-04282-1

24. Fofanov M v., Morozova V v., Kozlova YN, Tikunov AY, Babkin I v., Poletaeva YE, et al. *Raoultella* Bacteriophage RP180, A New Member of the Genus *Kagunavirus*, Subfamily *Guernseyvirinae*. Archives of virology [Internet]. 2019 Oct 1 [cited 2021 Nov 10];164(10):2637–40. Available from: https://pubmed.ncbi.nlm.nih.gov/31372754/

25. Shibu P. Investigations of Carbapenem-Resistant *Klebsiella* Species and Associated Clinical Considerations. 2019.

26. Herzog KAT, Schneditz G, Leitner E, Feierl G, Hoffmann KM, Zollner-Schwetz I, et al. Genotypes of *Klebsiella oxytoca* Isolates from Patients with Nosocomial Pneumonia are Distinct from Those of Isolates from Patients with Antibiotic-Associated Hemorrhagic Colitis. Journal of Clinical Microbiology [Internet]. 2014 [cited 2021 Apr 2];52(5):1607–16. Available from: /pmc/articles/PMC3993621/

27. NanoDrop™ 2000/2000c Spectrophotometers [Internet]. [cited 2021 Apr 2]. Available from: https://www.thermofisher.com/order/catalog/product/ND-2000#/ND-2000

28. Herzog KAT, Schneditz G, Leitner E, Feierl G, Hoffmann KM, Zollner-Schwetz I, et al. Genotypes of *Klebsiella oxytoca* Isolates From Patients With Nosocomial Pneumonia are Distinct From Those of Isolates From Patients With Antibiotic-Associated Hemorrhagic Colitis. Journal of Clinical Microbiology [Internet]. 2014 [cited 2021 Apr 2];52(5):1607–16. Available from: /pmc/articles/PMC3993621/

29. Jolley KA, Bray JE, Maiden MCJ. Open-Access Bacterial Population Genomics: BIGSdb Software, the PubMLST.org Website and Their Applications. Wellcome Open Research [Internet]. 2018 [cited 2021 Nov 15];3. Available from: /pmc/articles/PMC6192448/

30. Ma Y, Wu X, Li S, Tang L, Chen M, An Q. Proposal for Reunification of the Genus *Raoultella* with the Genus *Klebsiella* and Reclassification of *Raoultella electrica* as *Klebsiella electrica* comb. nov. Research in Microbiology. 2021 Sep 1;172(6):103851.

31. Letunic I, Bork P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. Nucleic Acids Research [Internet]. 2019 Jul 1 [cited 2021 Apr 15];47(W1):W256–9. Available from: https://academic.oup.com/nar/article/47/W1/W256/5424068

32. Haines MEK, Hodges FE, Nale JY, Mahony J, van Sinderen D, Kaczorowska J, et al. Analysis of Selection Methods to Develop Novel Phage Therapy Cocktails Against Antimicrobial Resistant Clinical Isolates of Bacteria. Frontiers in Microbiology. 2021 Mar 29;12:564.

33. 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. Journal of Computational Biology [Internet]. 2012 May 1 [cited 2021 Aug 19];19(5):455. Available from: /pmc/articles/PMC3342519/

34. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: Interactive Visualization of *de novo* Genome Assemblies. Bioinformatics [Internet]. 2015 Oct 15 [cited 2021 Aug 19];31(20):3350–2. Available from: https://academic.oup.com/bioinformatics/article/31/20/3350/196114

35. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV Assesses the Quality and Completeness of Metagenome-Assembled Viral Genomes. Nature biotechnology [Internet]. 2021 May 1 [cited 2022 Feb 28];39(5):578–85. Available from: https://pubmed.ncbi.nlm.nih.gov/33349699/

36. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: Antibiotic Resistome Surveillance with the Comprehensive Antibiotic Resistance Database. Nucleic acids research [Internet]. 2020 Jan 1 [cited 2022 Feb 28];48(D1):D517–25. Available from: https://pubmed.ncbi.nlm.nih.gov/31665441/

37. Y N, T Y, M K, H U, H O, S G. ViPTree: The Viral Proteomic Tree Server. Bioinformatics (Oxford, England) [Internet]. 2017 Aug 1 [cited 2021 Aug 19];33(15):2379–80. Available from: https://pubmed.ncbi.nlm.nih.gov/28379287/

38. Kęsik-Szeloch A, Drulis-Kawa Z, Weber-Dąbrowska B, Kassner J, Majkowska-Skrobek G, Augustyniak D, et al. Characterising the Biology of Novel Lytic Bacteriophages Infecting Multidrug Resistant *Klebsiella pneumoniae*. Virology Journal 2013 10:1 [Internet]. 2013 Mar 28 [cited 2021 Aug 8];10(1):1–12. Available from: https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-10-100

39. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Analytical Methods [Internet]. 2015 Dec 17 [cited 2021 Nov 6];8(1):12–24. Available from: https://pubs.rsc.org/en/content/articlehtml/2016/ay/c5ay02550h

40. Seemann T. Prokka: Rapid Prokaryotic Genome Annotation. Bioinformatics (Oxford, England) [Internet]. 2014 Jul 15 [cited 2021 Dec 9];30(14):2068–9. Available from: https://pubmed.ncbi.nlm.nih.gov/24642063/

41. Terzian P, Olo Ndela E, Galiez C, Lossouarn J, Pérez Bucio RE, Mom R, et al. PHROG: Families of Prokaryotic Virus Proteins Clustered Using Remote Homology. NAR genomics and bioinformatics [Internet]. 2021 Jun 23 [cited 2021 Dec 9];3(3). Available from: https://pubmed.ncbi.nlm.nih.gov/34377978/

42. Pagès H, Aboyoun P, Gentleman R, DebRoy S. Biostrings: Efficient Manipulation of Biological Strings. R package version 2620.

43. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. Nucleic acids research [Internet]. 2021 Jul 2 [cited 2021 Dec 9];49(W1):W293–6. Available from: https://pubmed.ncbi.nlm.nih.gov/33885785/

44. Rangel-Pineros G, Millard A, Michniewski S, Scanlan D, Sirén K, Reyes A, et al. From Trees to Clouds: PhageClouds for Fast Comparison of ~640,000 Phage Genomic Sequences and Host-Centric Visualization Using Genomic Network Graphs. https://home.liebertpub.com/phage [Internet]. 2021 Dec 16 [cited 2022 Feb 28];2(4):194–203. Available from: https://www.liebertpub.com/doi/abs/10.1089/phage.2021.0008

45. Moraru C, Varsani A, Kropinski AM. VIRIDIC—A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. Viruses 2020, Vol 12, Page 1268 [Internet]. 2020 Nov 6 [cited 2022 Feb 28];12(11):1268. Available from: https://www.mdpi.com/1999-4915/12/11/1268/htm

46. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: Rapid Large-Scale Prokaryote Pan Genome Analysis. Bioinformatics [Internet]. 2015 Nov 15 [cited 2022 Feb 28];31(22):3691–3. Available from: https://academic.oup.com/bioinformatics/article/31/22/3691/240757

47. Wang LG, Lam TTY, Xu S, Dai Z, Zhou L, Feng T, et al. Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. Molecular Biology and Evolution [Internet]. 2020 Feb 1 [cited 2022 Feb 28];37(2):599–603. Available from: https://academic.oup.com/mbe/article/37/2/599/5601621

48. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees with their Covariates and Other Associated Data. Methods in Ecology and Evolution [Internet]. 2017 Jan 1 [cited 2022 Feb 28];8(1):28–36. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12628

49. Edgar RC. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. Nucleic Acids Research [Internet]. 2004 Mar 1 [cited 2022 Feb 28];32(5):1792–7. Available from: https://academic.oup.com/nar/article/32/5/1792/2380623

50. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology [Internet]. 2010 May 1 [cited 2022 Feb 28];59(3):307–21. Available from: https://academic.oup.com/sysbio/article/59/3/307/1702850

51. Maciejewska B, Roszniowski B, Espaillat A, Kęsik-Szeloch A, Majkowska-Skrobek G, Kropinski AM, et al. Klebsiella Phages Representing a Novel Clade of Viruses With an Unknown DNA Modification and Biotechnologically Interesting Enzymes. Applied Microbiology and Biotechnology [Internet]. 2017 Jan 1 [cited 2021 Jan 28];101(2):673–84. Available from: http://www.ncbi.nlm.nih.gov/blast/

52. Turner D, Kropinski AM, Adriaenssens EM. A Roadmap for Genome-Based Phage Taxonomy. Viruses [Internet]. 2021 Mar 18 [cited 2021 Apr 14];13(3):506. Available from: https://www.mdpi.com/1999-4915/13/3/506

53. Chen Y, Brook TC, Soe CZ, O'Neill I, Alcon-Giner C, Leelastwattanagul O, et al. Preterm Infants Harbour Diverse *Klebsiella* Populations, Including Atypical Species That Encode and Produce an Array of Antimicrobial Resistance- and Virulence-Associated Factors. Microbial Genomics. 2020 Jun;6(6).

54. Gómez M, Valverde A, del Campo R, Rodríguez JM, Maldonado-Barragán A. Phenotypic and Molecular Characterization of Commensal, Community-Acquired and Nosocomial *Klebsiella* spp. Microorganisms [Internet]. 2021 Nov 1 [cited 2021 Dec 27];9(11). Available from: /pmc/articles/PMC8625991/

55. Thorpe H, Booton R, Kallonen T, Gibbon MJ, Couto N, Passet V, et al. One Health or Three? Transmission Modelling of *Klebsiella* Isolates Reveals Ecological Barriers to Transmission Between humans, Animals and the Environment. bioRxiv [Internet]. 2021 Aug 11 [cited 2021 Dec 28];2021.08.05.455249. Available from: https://www.biorxiv.org/content/10.1101/2021.08.05.455249v2

56. Cai Z, Guo Q, Yao Z, Zheng W, Xie J, Bai S, et al. Comparative Genomics of *Klebsiella michiganensis* BD177 and Related Members of *Klebsiella* sp. Reveal the Symbiotic Relationship With *Bactrocera dorsalis*. BMC Genetics [Internet]. 2020 Dec 1 [cited 2021 Jul 7];21(Suppl 2). Available from: /pmc/articles/PMC7747454/

57. McDougall FK, Wyres KL, Judd LM, Boardman WSJ, Holt KE, Power ML. Novel Strains of *Klebsiella africana* and *Klebsiella pneumoniae* in Australian Fruit Bats (*Pteropus poliocephalus*). Research in Microbiology. 2021 Nov 1;172(7–8):103879.

58. Oliveira RA, Ng KM, Correia MB, Cabral V, Shi H, Sonnenburg JL, et al. *Klebsiella michiganensis* Transmission Enhances Resistance to Enterobacteriaceae Gut Invasion by Nutrition competition. Nature microbiology. 2020 Apr;5(4):630–41.

59. 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 of the United States of America. 2015 Jul;112(27):E3574--3581.

60. Pan Y-J, Lin T-L, Chen C-C, Tsai Y-T, Cheng Y-H, Chen Y-Y, et al. *Klebsiella* Phage ΦK64-1 Encodes Multiple Depolymerases for Multiple Host Capsular Types. Journal of Virology [Internet]. 2017 Feb 15 [cited 2021 Jan 28];91(6). Available from: /pmc/articles/PMC5331798/

61. Verma V, Kusum AE, Ae H, Chhibber S. Characterization of a T7-Like Lytic Bacteriophage of *Klebsiella pneumoniae* B5055: A Potential Therapeutic Agent.

62. Wu L-T, Chang S-Y, Yen M-R, Yang T-C, Tseng Y-H. Characterization of Extended-Host-Range Pseudo-T-Even Bacteriophage Kpp95 Isolated on *Klebsiella pneumoniae*. Applied and Environmental Microbiology [Internet]. 2007 Apr [cited 2021 Jul 17];73(8):2532. Available from: /pmc/articles/PMC1855606/

63. Maciejewska B, Roszniowski B, Espaillat A, Kęsik-Szeloch A, Majkowska-Skrobek G, Kropinski AM, et al. *Klebsiella* Phages Representing A Novel Clade of Viruses with an Unknown DNA Modification and Biotechnologically Interesting Enzymes. Applied microbiology and biotechnology [Internet]. 2017 Jan 1 [cited 2021 Nov 16];101(2):673–84. Available from: https://pubmed.ncbi.nlm.nih.gov/27766357/

64. Jensen EC, Schrader HS, Rieland B, Thompson TL, Lee KW, Nickerson KW, et al. Prevalence of Broad-Host-Range Lytic Bacteriophages of *Sphaerotilus natans*, *Escherichia coli*, and *Pseudomonas aeruginosa*. Applied and Environmental Microbiology [Internet]. 1998 [cited 2022 Feb 28];64(2):575–80. Available from: https://journals.asm.org/doi/abs/10.1128/AEM.64.2.575-580.1998

65. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, et al. Proposed Minimal Standards for the Use of Genome Data for the Taxonomy of Prokaryotes. International journal of systematic and evolutionary microbiology [Internet]. 2018 Jan 1 [cited 2022 Feb 28];68(1):461–6. Available from: https://pubmed.ncbi.nlm.nih.gov/29292687/

66. Provasek VE, Lessor LE, Cahill JL, Rasche ES, Kuty Everett GF. Complete Genome Sequence of Carbapenemase-Producing *Klebsiella pneumoniae* Myophage Matisse. Genome announcements [Internet]. 2015 [cited 2021 Nov 16];3(5). Available from: https://pubmed.ncbi.nlm.nih.gov/26430049/

67. Pereira C, Moreirinha C, Teles L, Rocha RJM, Calado R, Romalde JL, et al. Application of Phage Therapy During Bivalve Depuration Improves *Escherichia coli* Decontamination. Food microbiology [Internet]. 2017 Feb 1 [cited 2021 Nov 16];61:102–12. Available from: https://pubmed.ncbi.nlm.nih.gov/27697159/

68. Zhao J, Zhang Z, Tian C, Chen X, Hu L, Wei X, et al. Characterizing the Biology of Lytic Bacteriophage vB_EaeM_ϕEap-3 Infecting Multidrug-Resistant *Enterobacter aerogenes*. Frontiers in microbiology [Internet]. 2019 [cited 2021 Nov 16];10(MAR). Available from: https://pubmed.ncbi.nlm.nih.gov/30891025/

69. Mijalis EM, Lessor LE, Cahill JL, Rasche ES, Everett GFK. Complete Genome Sequence of *Klebsiella pneumoniae* Carbapenemase-Producing K. pneumoniae Myophage Miro. Genome announcements [Internet]. 2015 [cited 2021 Nov 16];3(5). Available from: https://pubmed.ncbi.nlm.nih.gov/26430050/

70. Koberg S, Brinks E, Fiedler G, Hüsing C, Cho GS, Hoeppner MP, et al. Genome Sequence of *Klebsiella pneumoniae* Bacteriophage PMBT1 Isolated from Raw Sewage. Genome announcements [Internet]. 2017 [cited 2021 Nov 16];5(8). Available from: https://pubmed.ncbi.nlm.nih.gov/28232430/

71. Wang Z, Cai R, Wang G, Guo Z, Liu X, Guan Y, et al. Combination Therapy of Phage vB_KpnM_P-KP2 and Gentamicin Combats Acute Pneumonia Caused by K47 Serotype *Klebsiella pneumoniae*. Frontiers in Microbiology [Internet]. 2021 Apr 22 [cited 2021 Aug 12];12. Available from: /pmc/articles/PMC8100603/

72. Tisza MJ, Buck CB. A Catalog of Tens of Thousands of Viruses from Human Metagenomes Reveals Hidden Associations with Chronic Diseases. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 2021 Jun 8 [cited 2022 Mar 7];118(23). Available from: /pmc/articles/PMC8201803/

73. Camarillo-Guerrero LF, Almeida A, Rangel-Pineros G, Finn RD, Lawley TD. Massive Expansion of Human Gut Bacteriophage Diversity. Cell [Internet]. 2021 Feb 18 [cited 2022 Mar 7];184(4):1098. Available from: /pmc/articles/PMC7895897/

74. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut. Cell Host & Microbe [Internet]. 2020 Nov 11 [cited 2022 Mar 7];28(5):724. Available from: /pmc/articles/PMC7443397/

772 **Table 1.** Clinical and veterinary isolates included in this study.

773

774 With the exceptions of Ko8, Ko31 and GFKo17 (*K. oxytoca/R. ornithinolytica*), all strains were initially
775 identified as *K. oxytoca*.

776

Lab ID*	Source	Method of initial identification†	*rpoB* gene-based identity (sequence similarity, %)
Ko1	Blood culture (hand)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko2	Blood culture	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko3	Sputum	M	*K. michiganensis* W14^T (99.60)
Ko4	Blood culture	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko5	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.60)
Ko6	Urine	M	*K. oxytoca* ATCC 13182^T (99.60)
Ko7	Blood culture (arterial line)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko8	Blood culture	M	*K. grimontii* 06D021^T (99.40)
Ko9	Blood culture (peripheral blood)	M	*K. grimontii* 06D021^T (99.40)
Ko10	Blood culture	M	*K. michiganensis* W14^T (99.40)
Ko11	Blood culture (white waste)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko12	Blood culture	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko13	Blood culture	M	*K. michiganensis* W14^T (99.60)
Ko14	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.40)
Ko15	Tissue (toe)	M	*K. oxytoca* ATCC 13182^T (99.60)
Ko16	Tissue (hip)	M	*K. grimontii* 06D021^T (99.60)
Ko17	Peritoneal fluid	M	*K. oxytoca* ATCC 13182^T (99.40)
Ko18	Sputum	M	*K. michiganensis* W14^T (99.40)
Ko19	Blood culture (peripheral blood)	M	*K. oxytoca* ATCC 13182^T (100.00)
Ko20	Blood culture	M	*K. oxytoca* ATCC 13182^T (99.60)
Ko21	Perfusion fluid	M	*K. michiganensis* W14^T (99.60)

Lab ID*	Source	Method of initial identification†	*rpoB* gene-based identity (sequence similarity, %)
Ko22	Perfusion fluid	M	*K. michiganensis* W14^T (99.60)
Ko23	Blood culture (venous blood)	M	*K. michiganensis* W14^T (99.60)
Ko24	Blood culture (systemic)	M	*K. michiganensis* W14^T (99.57)
Ko25	Clean catch urine	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko26	Urine	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko27	Brocho-alveolar lavage	M	*K. grimontii* 06D021^T (99.60)
Ko28	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.60)
Ko29	Sputum	M	*K. michiganensis* W14^T (99.40)
Ko30	Peritoneal fluid	M	*K. grimontii* 06D021^T (99.35)
Ko31	Sputum	M	*K. michiganensis* W14^T (99.40)
Ko32	Blood culture (venous blood)	M	*K. michiganensis* W14^T (99.60)
Ko33	Sputum	M	*K. michiganensis* W14^T (99.35)
Ko34	Blood culture (venous blood)	M	*K. pneumoniae* ATCC 13883^T (99.80)
Ko35	Sputum	M	*K. michiganensis* F107 CP024643.1 (97.67)‡
Ko36	Blood culture (Hickman line)	M	*K. michiganensis* W14^T (99.60)
Ko37	Blood culture (venous blood)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko38	Blood culture (venous blood)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko39	Blood culture (CVC line)	M	*K. michiganensis* W14^T (99.60)
Ko40	Blood culture (arm)	M	*K. oxytoca* ATCC 13182^T (99.60)
Ko41	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.20)
Ko42	Peritoneal fluid	M	*K. grimontii* 06D021^T (99.60)
Ko43	Bile fluid	M	*K. michiganensis* W14^T (99.80)
Ko44	CAPD fluid	M	*K. oxytoca* ATCC 13182^T (100.00)
Ko45	Aspirate (liver)	M	*K. oxytoca* ATCC 13182^T (99.35)
Ko46	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (100.00)
Ko47	Blood culture (central line)	M	*K. oxytoca* ATCC 13182^T (99.80)

Lab ID*	Source	Method of initial identification†	*rpoB* gene-based identity (sequence similarity, %)
Ko48	Blood culture (ACF blood)	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko49	Tissue	M	*K. michiganensis* W14^T (100.00)
Ko50	Bone biopsy (heel)	M	*K. michiganensis* W14^T (99.20)
Ko51	Fluid (pancreas)	M	*K. grimontii* 06D021^T (99.40)
Ko52	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.60)
Ko53	Blood culture	M	*K. oxytoca* ATCC 13182^T (99.60)
Ko54	Blood culture (peripheral blood)	M	*K. oxytoca* ATCC 13182^T (99.78)
Ko55	Mid stream urine	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko56	Blood culture (peripheral blood)	M	*K. michiganensis* W14^T (99.60)
Ko57	Sputum	M	*K. oxytoca* ATCC 13182^T (99.80)
Ko58	Perfusion fluid	M	*K. michiganensis* W14^T (100.00)
Ko59	Blood culture (venous blood)	M	*K. michiganensis* W14^T (99.40)
GFKo2	Seal, foetal stomach; Rossh-shire	A	*K. grimontii* 06D021^T (99.40)
GFKo3	Common seal, peritoneal fluid; Angus	A	*K. grimontii* 06D021^T (99.40)
GFKo4	Common seal, lung; Inverness-shire	A	*R. ornithinolytica* NBRC 105727^T (99.60)
GFKo5	Sparrow, eye; England	A	*K. grimontii* 06D021^T (99.60)
GFKo6	African elephant, lung; Perthshire	A	*K. oxytoca* ATCC 13182^T (99.80)
GFKo7	Porpoise, SI; Berwickshire	A	*K. michiganensis* W14^T (99.60)
GFKo8	Porpoise, lung; Berwickshire	A	*K. michiganensis* W14^T (99.60)
GFKo9	Porpoise, lung; Perthshire	A	*R. ornithinolytica* NBRC 105727^T (99.80)
GFKo10	Bottlenose dolphin, intestine; Ross-shire	A	*K. michiganensis* W14^T (99.00)
GFKo11	Bovine, Milk, Perthshire	M	*K. huaxiensis* WCHKl090001^T (100.00)
GFKo12	Bovine, Milk, Not known	A	*K. grimontii* 06D021^T (99.40)
GFKo13	Bovine, Milk, Dumfries	M	*K. michiganensis* W14^T (99.60)
GFKo14	Bovine, Milk, Dumfries	M	*K. michiganensis* W14^T (99.40)
GFKo15	Bovine, Milk, England	M	*K. michiganensis* W14^T (99.40)

Lab ID*	Source	Method of initial identification†	*rpoB* gene-based identity (sequence similarity, %)
GFKo16	Bovine, Milk, England	M	*K. oxytoca* ATCC 13182^T (99.60)
GFKo17	Bovine, Milk sock, Fife	M	*R. ornithinolytica* NBRC 105727^T (99.40)
GFKo18	Bovine, Milk, England	A	*K. michiganensis* W14^T (99.40)
GFKo19	Bovine, Foetal stomach, England	M	*K. michiganensis* W14^T (99.40)
GFKo20	Bovine, Foetal stomach, England	A	*K. grimontii* 06D021^T (99.60)
GFKo21	Bovine, Milk, Midlothian	A	*K. grimontii* 06D021^T (99.40)
GFKo22	Bovine, Milk, England	M	*K. grimontii* 06D021^T (99.60)
GFKo24	Bovine, Milk, England	M	*R. ornithinolytica* NBRC 105727^T (99.80)
GFKo25	Bovine, Milk, England	A	*R. ornithinolytica* NBRC 105727^T (99.80)
GFKo26	Bovine, Lung, Caithness	A	*K. grimontii* 06D021^T (99.60)
GFKo27	Bovine, Milk, unknown	M	*K. grimontii* 06D021^T (99.60)
GFKo28	Bovine, Milk, Wales?	M	*K. michiganensis* W14^T (99.40)
GFKo29	Bovine, Milk, Lanark	M	*K. michiganensis* W14^T (99.40)
GFKo30	Bovine, Milk, unknown	M	*K. michiganensis* W14^T (100.00)
GFKo31	Bovine, Milk sock, Moray	A	*K. michiganensis* W14^T (99.20)
GFKo32	Canine, Urine, Aberdeen	A	*K. michiganensis* W14^T (99.40)
GFKo33	Bovine, Lung, Orkney	A	*K. grimontii* 06D021^T (99.40)
GFKo34	Gecko, Skin, Caithness	A	*K. grimontii* 06D021^T (99.60)
GFKo35	Bovine, Spleen, Orkney	A	*K. michiganensis* W14^T (99.40)
GFKo36	Red deer, Lung, Orkney	A	*K. michiganensis* W14^T (99.40)
GFKo37	Bovine, Milk, Orkney	A	*K. michiganensis* W14^T (99.59)
GFKo38	Bovine, Milk, Dumfries	A	*K. michiganensis* W14^T (99.35)
GFKo39	Bovine, Milk, Dumfries	A	*K. grimontii* 06D021^T (99.40)
GFKo40	Bovine, Milk, Dumfries	A	*K. michiganensis* W14^T (99.60)
GFKo41	Bovine, Milk, Dumfries	A	*K. michiganensis* W14^T (99.60)
GFKo42	Bovine, Milk, Wigtown	A	*K. pneumoniae* ATCC 13883^T (99.60)

Lab ID*	Source	Method of initial identification†	*rpoB* gene-based identity (sequence similarity, %)
GFKo43	Bovine, Milk, Dumfries	A	*R. ornithinolytica* NBRC 105727^T (99.80)
GFKo44	Bovine, Milk, Dumfries	A	*K. michiganensis* W14^T (99.60)
GFKo45	Bovine, Milk, Dumfries	A	*K. michiganensis* W14^T (99.56)
GFKo46	Poultry layer, Liver, Lanark	Unknown	*K. pneumoniae* ATCC 13883^T (99.60)
GFKo47	Canine, Nose, Peebles	Unknown	*K. oxytoca* ATCC 13182^T (99.80)
GFKo48	Bovine, Foetal stomach, Ayr	Unknown	*K. pneumoniae* ATCC 13883^T (99.80)
GFKo49	Partridge, Liver, Wigtown	Unknown	*K. pneumoniae* ATCC 13883^T (99.60)
GFKo50	Bovine, Milk, Lanark	Unknown	*K. huaxiensis* WCHKl090001T (99.40)
GFKo52	Bovine, Milk, Ayr	Unknown	*R. terrigena* NBRC 14941^T (100.00)

777 * Ko prefix, clinical isolate (all obtained from the Pathogen Bank of Queen's Medical Centre

778 Nottingham); GFKo prefix, veterinary isolate.

779 † A, API 20E; M, MALDI-TOF MS.

780 ‡ Identified through BLASTn analysis.

781 **Table 2.** *Slopekvirus* genomes included in this study

Phage	Genome accession	Genome size (nt)	CDS*	CheckV			Reference
				Quality†	Completeness (%)	Contamination (%)	
Klebsiella phage vB_KmiM-2Di	MZ707156	177200	275	High-quality	99.34	0	This study
Klebsiella phage vB_KmiM-4Dii	MZ707157	174857	272	High-quality	98.02	0	This study
Klebsiella virus KP15	NC_014036	174436	269	High-quality	97.79	0	(63)
Klebsiella virus KP27	NC_020080	174413	271	High-quality	97.78	0	(63)
Klebsiella virus Matisse	NC_028750	176081	273	High-quality	98.73	0	(66)
Escherichia phage phT4A	NC_055712	171598	259	High-quality	96.20	0	(67)
Enterobacter virus Eap3	NC_041980	175814	275	High-quality	98.56	0	(68)
Klebsiella virus Miro	NC_041981	176055	274	High-quality	98.70	0	(69)
Klebsiella virus PMBT1	NC_042138	175206	274	High-quality	98.22	0	(70)
Klebsiella phage vB_KpM-KalD	LR881140	174351	268	Complete	100	0	(21)
Klebsiella phage vB_KpM-SoFaint	LR881141	175933	271	Complete	100	0	(21)
Klebsiella phage vB_KpM-Milk	LR881142	176734	271	Complete	100	0	(21)
Klebsiella phage vB_KoM-Liquor	LR881143	176734	270	Complete	100	0	(21)
Klebsiella phage vB_KoM-Pickle	LR881145	175221	273	High-quality	98.23	0	(21)

Phage	Genome accession	Genome size (nt)	CDS*	CheckV			Reference
				Quality†	Completeness (%)	Contamination (%)	
Klebsiella phage vB_KpM-Mild	LR881147	176856	271	Complete	100	0	(21)
Klebsiella phage vB_KoM-MeTiny	LR883651	175419	272	Complete	100	0	(21)
Klebsiella phage vB_KpnM_05F‡	LR746310	174231	272	Complete	100	0	(32)
Klebsiella phage KOX8‡	MN101221	131200	190	Medium-quality	73.55	0	Unpublished
Klebsiella phage KOX10‡	MN101223	168074	252	High-quality	94.22	0	Unpublished
Klebsiella phage KOX11‡	MN101224	59583	88	Low-quality	33.42	0	Unpublished
Klebsiella_phage vB_KpnM_P-KP2‡	MT157285	172138	262	High-quality	96.51	0	(71)
Klebsiella phage vB_KpnM_M1‡	MW448170	176329	275	High-quality	98.85	0	Unpublished
SAMN00791912_a1_ct12238‡	–	174367	272	High-quality	97.76	0	(72)
uvig_221808‡	–	18523	17	Genome-fragment	10.78	0	(73)
uvig_279022‡	–	16907	33	Genome-fragment	9.48	0	(73)
uvig_279024‡	–	15914	37	Genome-fragment	8.93	0	(73)
uvig_279033‡	–	12820	25	Genome-fragment	7.19	0	(73)
uvig_298784‡	–	175996	274	High-quality	98.67	0	(73)
uvig_335830‡	–	174400	272	High-quality	97.77	0	(73)

Phage	Genome accession	Genome size (nt)	CDS*	CheckV			Reference
				Quality†	Completeness (%)	Contamination (%)	
uvig_373971‡	–	16673	24	Genome-fragment	7.20	22.93	(73)
uvig_510342‡	–	18306	38	Genome-fragment	10.26	0	(73)
uvig_510345‡	–	15123	34	Genome-fragment	8.48	0	(73)
Zuo_2017_SRR5677802_NODE_1_length_175069_cov_4814.635012‡	–	175069	272	High-quality	100	0	(74)

782 * Annotated using the PHROG hmm database with Prokka.

783 † Complete, completeness assessed DTR (high-confidence); for all other annotations (i.e. high-quality, genome fragment) completeness was assessed by
784 AAI (high-confidence).

785 ‡ Detected using PhageClouds.

786 **Table 3.** Host range analysis of phage vB_KmiM-2di and phage vB_KmiM-4Dii

Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii	Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii
Ko1	*K. oxytoca*	+ †	+	GFKo2	*K. grimontii*	0	+
Ko2	*K. oxytoca*	+	+	GFKo3	*K. grimontii*	0	+
Ko3	*K. michiganensis*	+	++	GFKo4	*R. ornithinolytica*	++	++
Ko4	*K. oxytoca*	+	++	GFKo5	*K. grimontii*	+	+
Ko5	*K. michiganensis*	+	+	GFKo6	*K. oxytoca*	0	0
Ko6	*K. oxytoca*	+	+	GFKo7	*K. michiganensis*	+, d	++, d
Ko7	*K. oxytoca*	0	+	GFKo8	*K. michiganensis*	0	+
Ko8	*K. grimontii*	+	+	GFKo9	*R. ornithinolytica*	0	0
Ko9	*K. grimontii*	+	++	GFKo10	*K. michiganensis*	0	+
Ko10	*K. michiganensis*	0	+	GFKo11	*K. huaxiensis*	+	+
Ko11	*K. oxytoca*	++	++	GFKo12	*K. grimontii*	0	0
Ko12	*K. oxytoca*	+	++	GFKo13	*K. michiganensis*	0	0
Ko13	*K. michiganensis*	+, d	+, d	GFKo14	*K. michiganensis*	+	+
Ko14	*K. michiganensis*	0	+	GFKo15	*K. michiganensis*	+	++
Ko15	*K. oxytoca*	+	++	GFKo16	*K. oxytoca*	++	++
Ko16	*K. grimontii*	++	++	GFKo17	*R. ornithinolytica*	0	0

Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii	Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii
Ko17	*K. oxytoca*	+	++	GFKo18	*K. michiganensis*	+	+
Ko18	*K. michiganensis*	++	++	GFKo19	*K. michiganensis*	+	+
Ko19	*K. oxytoca*	++	++	GFKo20	*K. grimontii*	++	++
Ko20	*K. oxytoca*	+	0	GFKo21	*K. grimontii*	++	++
Ko21	*K. michiganensis*	+	+	GFKo22	*K. grimontii*	+	+
Ko22	*K. michiganensis*	+	+	GFKo24	*R. ornithinolytica*	0	+
Ko23	*K. michiganensis*	+	+	GFKo25	*R. ornithinolytica*	0	+
Ko24	*K. michiganensis*	+	+	GFKo26	*K. grimontii*	+	+
Ko25	*K. oxytoca*	0	+	GFKo27	*K. grimontii*	+	+
Ko26	*K. oxytoca*	++	++	GFKo28	*K. michiganensis*	+	++
Ko27	*K. grimontii*	+	++	GFKo29	*K. michiganensis*	++	+
Ko28	*K. michiganensis*	+	+	GFKo30	*K. michiganensis*	0	0
Ko29	*K. michiganensis*	+	+	GFKo31	*K. michiganensis*	+	+
Ko30	*K. grimontii*	0	+	GFKo32	*K. michiganensis*	0	+
Ko31	*K. michiganensis*	+	++	GFKo33	*K. grimontii*	0	0
Ko32	*K. michiganensis*	+	+	GFKo34	*K. grimontii*	+	+
Ko33	*K. michiganensis*	0	+	GFKo35	*K. michiganensis*	+	+

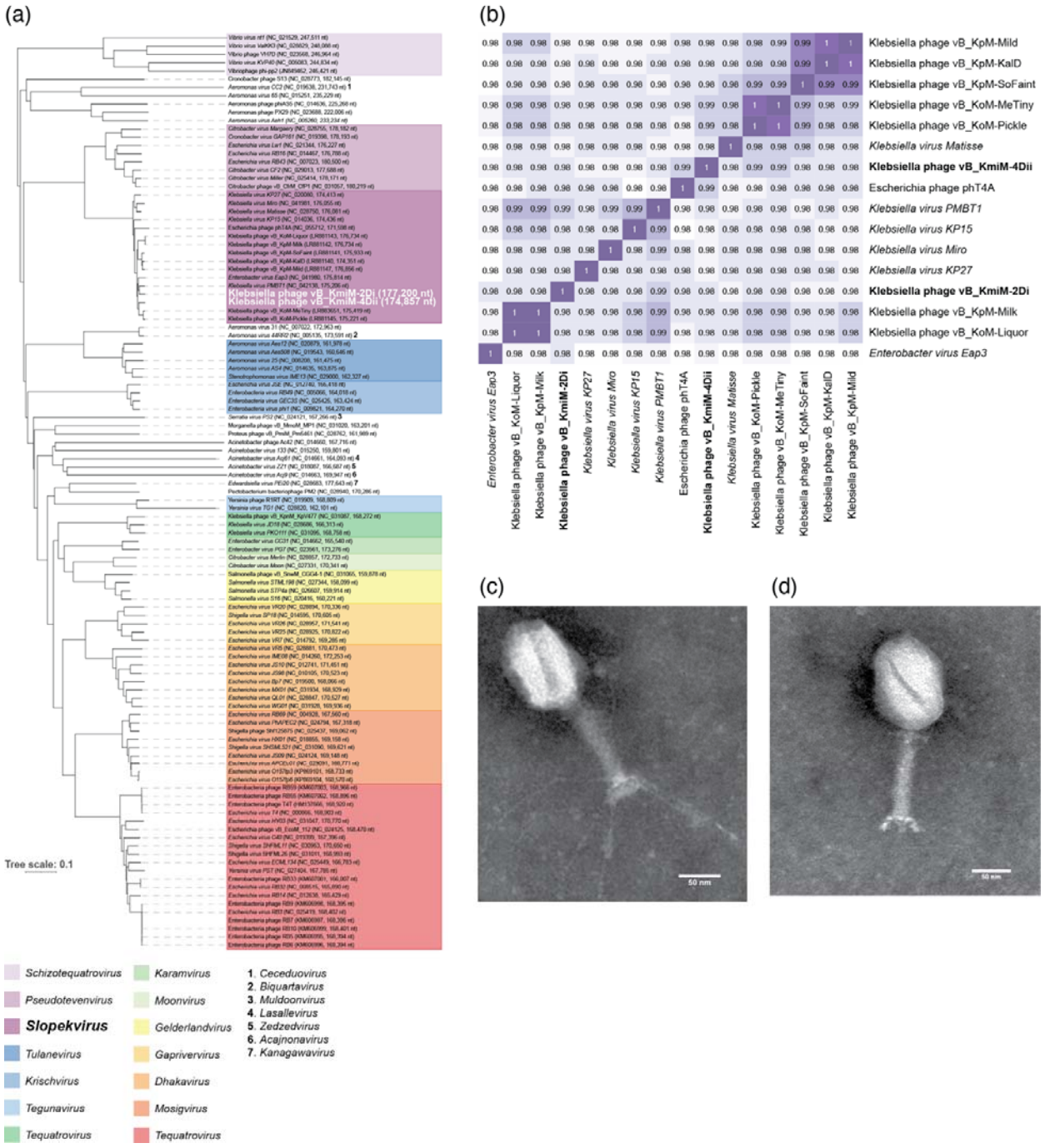
Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii	Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii
Ko34	*K. pneumoniae*	0	+	GFKo36	*K. michiganensis*	+	++
Ko35	*K. michiganensi*	+	0	GFKo37	*K. michiganensis*	+	+
Ko36	*K. michiganensis*	0	+	GFKo38	*K. michiganensis*	0	++
Ko37	*K. oxytoca*	+	+	GFKo39	*K. grimontii*	++	++
Ko38	*K. oxytoca*	+	+	GFKo40	*K. michiganensis*	+	+
Ko39	*K. michiganensis*	+	+	GFKo41	*K. michiganensis*	0	+,d
Ko40	*K. oxytoca*	0	0	GFKo42	*K. pneumoniae*	0	+
Ko41	*K. michiganensis*	0	0	GFKo43	*R. ornithinolytica*	0	0
Ko42	*K. grimontii*	+	+	GFKo44	*K. michiganensis*	0	0
Ko43	*K. michiganensis*	+	+	GFKo45	*K. michiganensis*	+	+
Ko44	*K. oxytoca*	+	+	GFKo46	*K. pneumoniae*	0	+
Ko45	*K. oxytoca*	+	+	GFKo47	*K. oxytoca*	0	0
Ko46	*K. michiganensis*	0	0	GFKo48	*K. pneumoniae*	0	+
Ko47	*K. oxytoca*	+	+	GFKo49	*K. pneumoniae*	0	0
Ko48	*K. oxytoca*	+	+	GFKo50	*K. huaxiensis*	0	0
Ko49	*K. michiganensis*	0	+	GFKo52	*R. terrigena*	++	++
Ko50	*K. michiganensis*	+	+				

Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii	Lab ID*	*rpoB* gene-based identity	vB_KmiM-2Di	vB_KmiM-4Dii
Ko51	*K. grimontii*	+	+				
Ko52	*K. michiganensis*	+	+				
Ko53	*K. oxytoca*	+	+				
Ko54	*K. oxytoca*	0	+				
Ko55	*K. oxytoca*	+	+				
Ko56	*K. michiganensis*	+	+				
Ko57	*K. oxytoca*	+	+				
Ko58	*K. michiganensis*	+	+				
Ko59	*K. michiganensis*	+	+				

787 * Ko prefix, clinical isolate; GFKo prefix, veterinary isolate.

788 † ++, complete lysis; +, hazy lysis; 0, no visible plaques; d, depolymerase activity.

789



790

791

792 **Fig. 1.** Genomic analyses of slopekviruses *Klebsiella* phage vB_KmiM-2Di and *Klebsiella* phage

793 vB_KmiM-4Dii and related viruses. (a) ViPTree-generated phylogenetic tree of slopekviruses and

794 their closest relatives. (b) ANIm analysis of vB_KmiM-2Di and vB_KmiM-4Dii genomes with members

795 of the genus *Slopekvirus*. (c) Phage vB_KmiM-2Di. (d) Phage vB_KmiM-4Dii. (c, d) Scale bar, 50 nm.

796

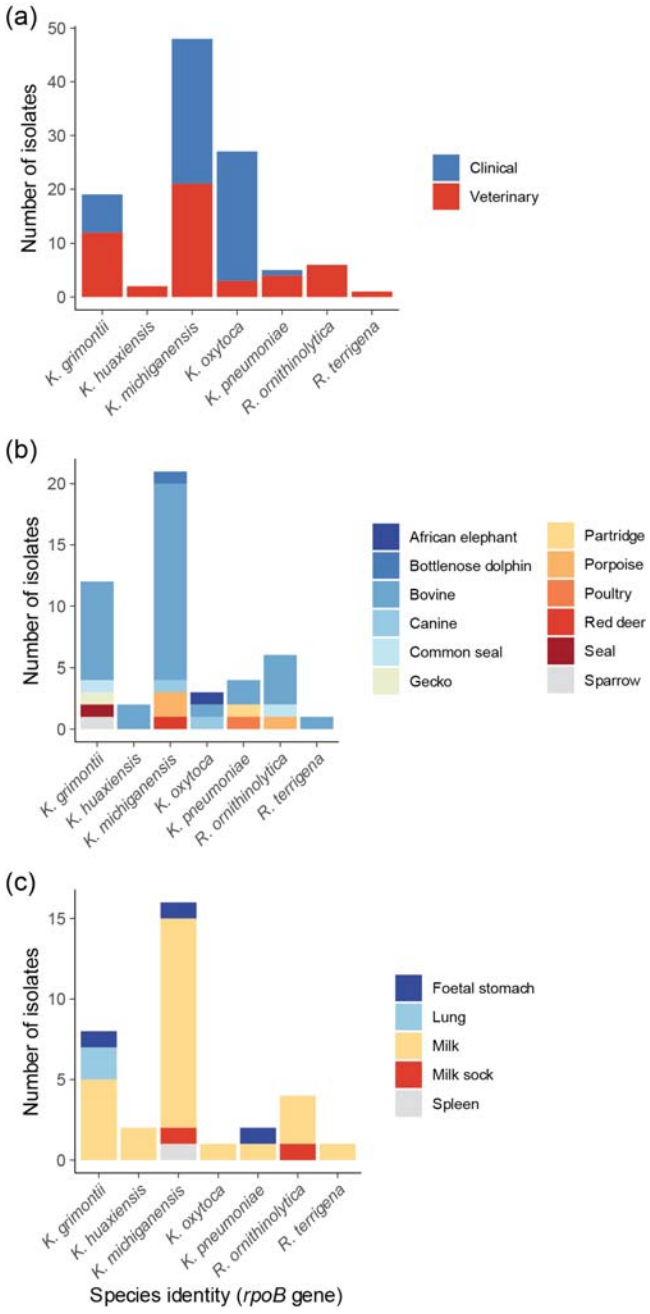


Fig. 2. Graphical representations of sources of clinical and veterinary isolates included in this study. (a) Distribution of clinical (*n*=59) and veterinary (*n*=49) isolates among the nine species of bacteria identified in this study. (b) Association of veterinary isolates (*n*=49) recovered from different animals with the nine species of bacteria identified in this study. (c) Association of bovine-associated isolates (*n*=34) with eight different species of bacteria identified in this study.

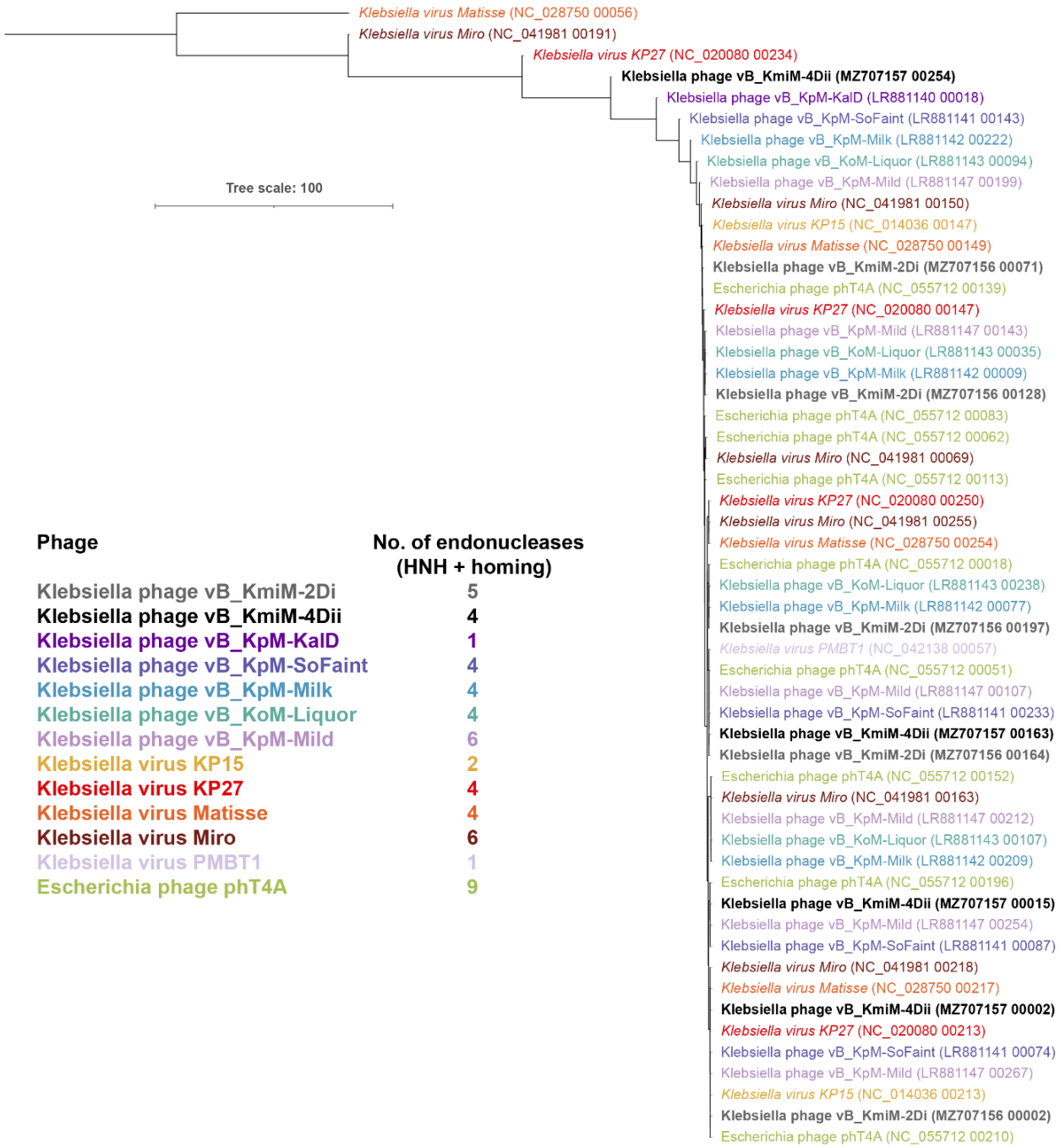


Fig. 3. Phylogenetic analysis (neighbour-joining tree) and summation of endonucleases encoded in slopekviruses. In the tree each gene product is coloured based on the phage that encodes it. Scale bar, average number of amino acid substitutions per position. The alignment, distance matrix and newick file associated with the analysis are available from figshare as **Supplementary Material**.

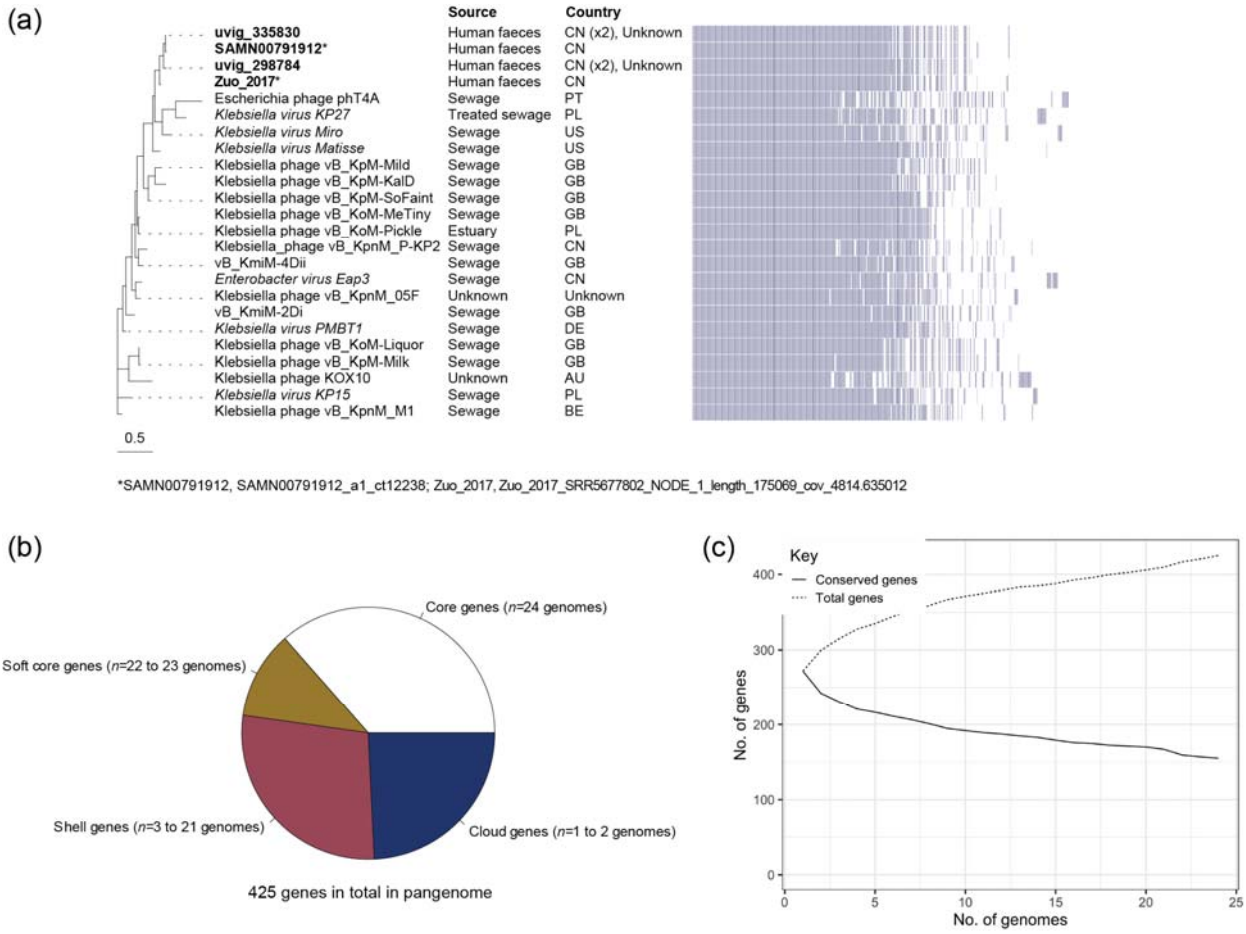
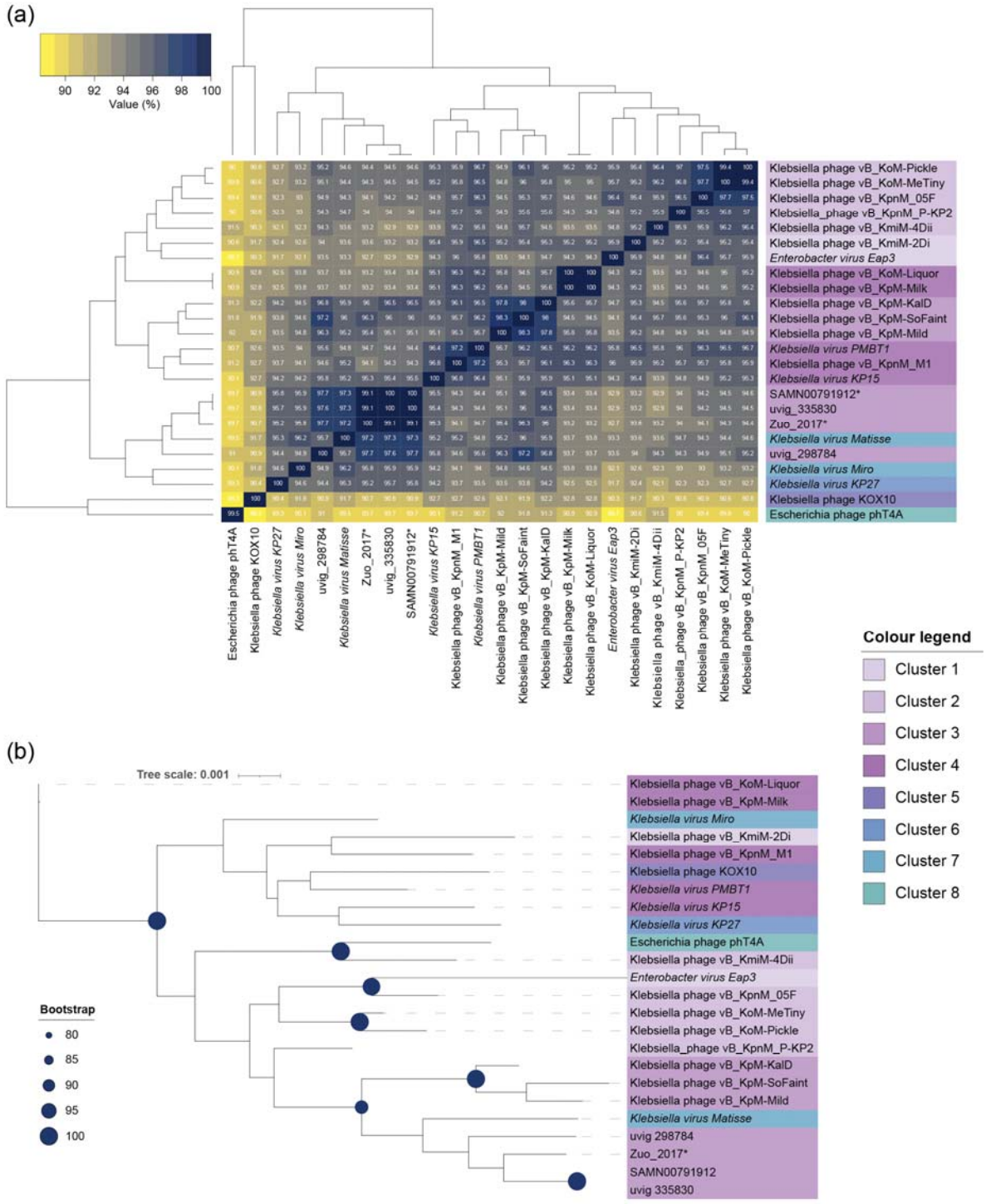


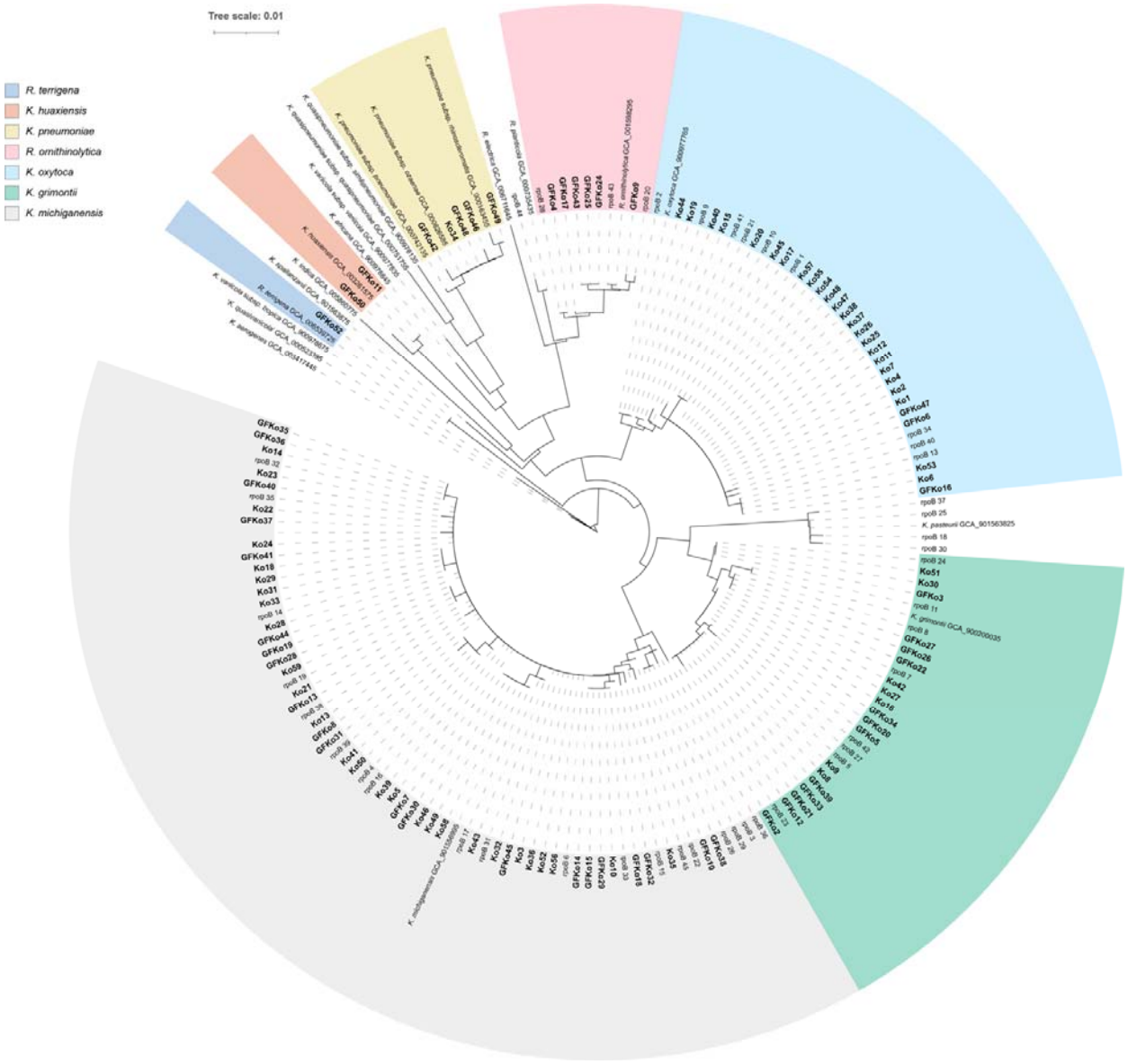
Fig. 4. Visualization of pangenome data for 24 slopekvirus genomes. (a) Roary matrix of the 425 genes representing the total pangenome. AU, Australia; BE, Belgium; CN, China; GB, Great Britain; DE, Denmark; PL, Poland; PT, Portugal; US, United State of America. (b) Pie chart showing the contribution of the 425 genes to the pangenome. (c) Plots showing how the number of total genes (dotted line) and conserved genes (dashed line) changes as more genomes are added to the slopekvirus pangenome.



*SAMN00791912, SAMN00791912_a1_ct12238; Zuo_2017, Zuo_2017_SRR5677802_NODE_1_length_175069_cov_4814.635012

Fig. 5. Determination of whether distinct species are represented with the genus *Slopekvirus*. (a) Bidirectional clustering heatmap visualizing VIRIDIC-generated similarity matrix for the 24 slopekvirus genomes. (b) Phylogenetic tree (maximum likelihood) generated by concatenating the

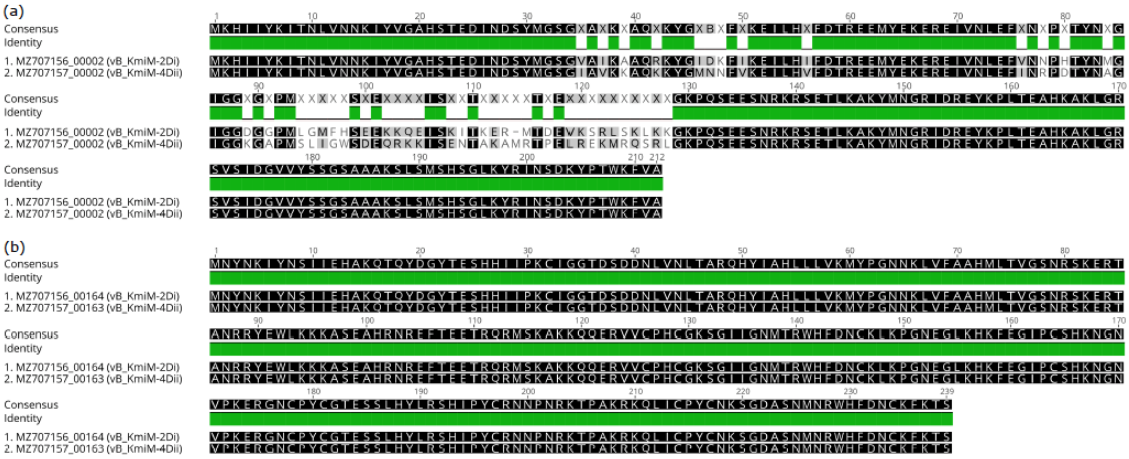
823 amino acid sequences of the 148 core genes from the pangenome that had ≥95 % identity and ≥70
824 % coverage across the 24 genomes. Bootstrap values are presented as a percentage of 100
825 replications. (a, b) Colouring of the data corresponds to the eight species clusters predicted by
826 VIRIDIC.



827

828

829 **Supplementary Fig. 1.** Phylogenetic tree (neighbour joining) showing relationship of clinical (*n*=59)
830 and veterinary (*n*=49) isolates with other *Klebsiella* (including *Raoultella*) spp. Accession numbers
831 shown for type species of *Klebsiella* spp. are for genome sequences from which *rpoB* gene
832 sequences were extracted. The tree was rooted using the *rpoB* gene sequence of *K. aerogenes* as an
833 outgroup. Scale bar, average number of nucleotide substitutions per position.



834

835

836 **Supplementary Fig. 2.** Amino acid alignments of homologous HENs identified in vB_KmiM-2Di and
837 vB_KmiM-4Dii. (a) MZ707156_00002 homologous with MZ707157_00002; 79.7 % pairwise amino
838 acid identity. (b) MZ707156_00164 homologous with MZ707157_00163; 100 % pairwise amino acid
839 identity.

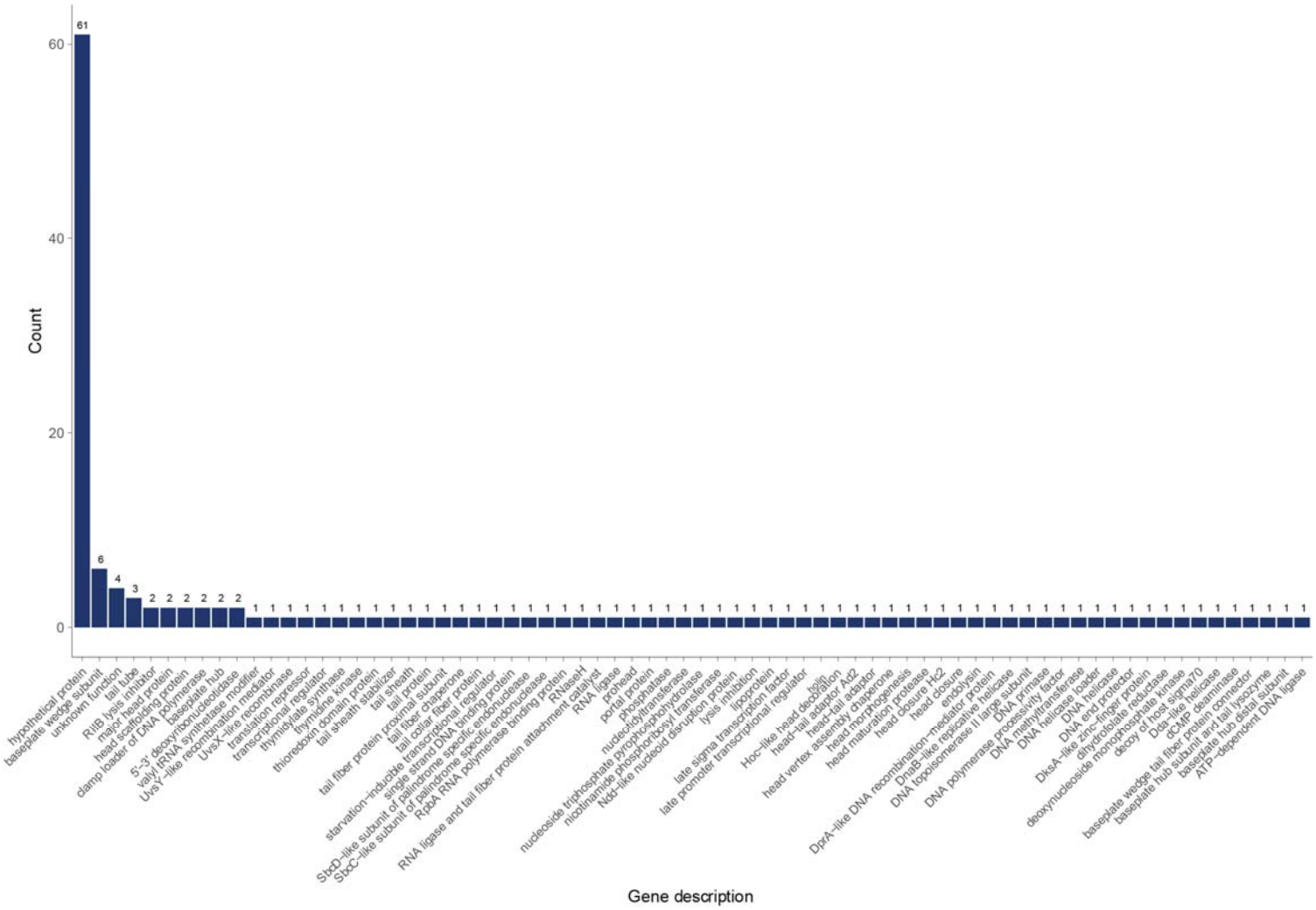
840

841

842

843

844 **Supplementary Fig. 3.** Mauve alignment of phages vB_KmiM-2Di and vB_KmiM-4Dii with
845 representatives of the genus *Slopekvirus*. The alignment shows that the genomes of the phages are
846 highly similar. Visual inspection revealed the HENs contributed to the diversity of the genomes.



847

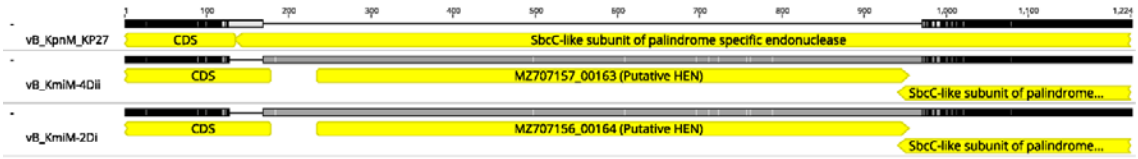
848 **Supplementary Fig. 4.** Summary of functions predicted for 148 core genes identified in the slopekvirus pangenome analysis.

849

850

851 **Supplementary Fig. 5.** Analysis of HEN YP_007348875.1 originally identified in vB_KpnM_KP27. (a)
852 Genome alignment of vB_KpnM_KP27 with vB_KmiM-2Di and vB_KmiM-4Dii revealed the presence
853 of a potentially homologous gene in vB_KmiM-4Dii that was absent in vB_KmiM-2Di. (b) Amino acid
854 alignment of two translated sequences showed 86 % pairwise amino acid identity between
855 YP_007348875.1 and MZ707157_00254.

856



Supplementary Fig. 6. Genomic regions encoding putative homologous HENs in vB_KmiM-2Di and vB_KmiM-4Dii. These HENs were found immediately upstream of a CDS encoding a putative SbcC-like subunit of a predicted palindrome-specific endonuclease. No putative homologue was identified in the genome of vB_KpnM_KP27.