RESEARCH

Genomic and functional analysis of *rmp* locus variants in *Klebsiella pneumoniae*

Margaret M. C. Lam[1*], Stephen M. Salisbury[2], Logan P. Treat[2], Ryan R. Wick[1], Louise M. Judd[1], Kelly L. Wyres[1,3], Sylvain Brisse[4], Kimberly A. Walker[2], Virginia L. Miller[2,5] and Kathryn E. Holt[1,6]

Abstract

Background *Klebsiella pneumoniae* is an opportunistic pathogen and a leading cause of healthcare-associated infections in hospitals, which are frequently antimicrobial resistant (AMR). Exacerbating the public health threat posed by *K. pneumoniae*, some strains also harbour additional hypervirulence determinants typically acquired via mobile genetic elements such as the well-characterised large virulence plasmid KpVP-1. The *rmpADC* locus is considered a key virulence feature of *K. pneumoniae* and is associated with upregulated capsule expression and the hypermucoid phenotype, which can enhance virulence by contributing to serum resistance. Typically such strains have been susceptible to all antimicrobials besides ampicillin; however, the recent emergence of AMR hypermucoid strains is concerning.

Methods Here, we investigate the genetic diversity, evolution, mobilisation and prevalence of *rmpADC*, in a dataset of 14,000 genomes from isolates of the *Klebsiella pneumoniae* species complex, and describe the RmST virulence typing scheme for tracking *rmpADC* variants for the purposes of genomic surveillance. Additionally, we examine the functionality of representatives for variants of *rmpADC* introduced into a mutant strain lacking its native *rmpADC* locus.

Results The *rmpADC* locus was detected in 7% of the dataset, mostly from genomes of *K. pneumoniae* and a very small number of *K. variicola* and *K. quasipneumoniae*. Sequence variants of *rmpADC* grouped into five distinct lineages (*rmp1*, *rmp2*, *rmp2A*, *rmp3* and *rmp4*) that corresponded to unique mobile elements, and were differentially distributed across different populations (i.e. clonal groups) of *K. pneumoniae*. All variants were demonstrated to produce enhanced capsule production and hypermucoviscosity.

Conclusions These results provide an overview of the diversity and evolution of a prominent *K. pneumoniae* virulence factor and support the idea that screening for *rmpADC* in *K. pneumoniae* isolates and genomes is valuable to monitor the emergence and spread of hypermucoid *K. pneumoniae*, including AMR strains.

Keywords *Klebsiella pneumoniae*, Hypermucoid, Hypermucoviscosity, Virulence, Hypervirulence, Genomic surveillance, Virulence plasmids

*Correspondence:
Margaret M. C. Lam
margaret.lam@monash.edu
Full list of author information is available at the end of the article

Lam *et al. Genome Medicine*   (2025) 17:36

Page 2 of 13

Background

A distinct pathotype of *Klebsiella pneumoniae*, often referred to as hypervirulent *K. pneumoniae* (hvKp), poses a significant public health challenge outside of hospital settings where it causes severe and sometimes life-threatening infections [1–3]. These are regarded as being distinct from 'classical *Kp*' strains, which typically cause opportunistic infections mostly within healthcare settings, and are often multidrug-resistant (MDR). Community-acquired infections can arise in otherwise healthy and immunocompetent individuals, although there are reportedly associations with comorbidities such as diabetes [4]. Common examples of infections include pyogenic liver abscess, endophthalmitis, pneumonia and meningitis, but they can also present as metastatic, multisite infections. Earlier reports of hypervirulent, community-acquired infections were largely confined to countries in Eastern Asia. Cases are now being more widely reported in other regions including Europe, North America and Australia, although often associated with individuals of East Asian descent [5].

The overwhelming majority of hvKp infections are associated with strains from distinct genetic backgrounds or lineages; these include clonal groups CG23, CG25, CG65, CG66, CG86 and CG380 [6, 7]. Several features are considered hallmark characteristics of these hvKp strains [7]. Most produce a K1 or K2 capsule, encoded by the *cps* (K) loci KL1 and KL2, respectively, and O1 lipopolysaccharide (OL1 locus). Many hvKp strains also exhibit hypermucoviscosity (HMV), which is defined by a positive string test and/or low OD_{600} measurements in a sedimentation assay, and is associated with presence of the *rmpADC* locus and/or *rmpA2* gene. Lastly, most hvKp also synthesise the siderophores aerobactin (*iuc* locus) and salmochelin (*iro* locus), in addition to the intrinsic siderophore enterobactin (*ent* locus) [7, 8]. The acquired siderophore loci, *rmpADC* and *rmpA2* (which also appears to be part of a locus including homologues of *rmpD* and *rmpC*), are typically mobilised by the large *K. pneumoniae* virulence plasmids (KpVP) that have been stably maintained for over 100 years in some hvKp clones including CG23 and CG86, although *iro* and *rmpADC* can also be mobilised via the chromosomal integrative conjugative element ICE*Kp1* [9–11].

In *K. pneumoniae*, the hypermucoid or HMV phenotype has long been associated with the gene *rmpA* although the exact mechanisms leading to the phenotype were unknown [12–14]. Based on observations and experimental evidence from earlier studies using knock-out mutants, it was proposed that the expression of *rmpA* upregulated expression of *cps* thereby increasing capsule production and subsequently resulting in HMV. Subsequent work has confirmed that the *rmpA* LuxR-like transcriptional regulator is part of a larger operon (herein called *rmp*) together with the *rmpD* and *rmpC* genes located downstream [15, 16]. This work has further clarified that enhanced capsule expression and HMV are two discrete features requiring *rmpC* (regulator of *cps*) and *rmpD* (encoding a small protein required for HMV), respectively. While HMV can be attained in the absence of elevated *cps* expression (i.e. in a *rmpC* mutant), it was not observed in capsule-defect mutants, suggesting that HMV does rely on the presence of some capsular components but these need not be hyperexpressed. Accordingly, it was recently shown that HMV is driven through elongation of the capsule polysaccharide chain, driven by a direct interaction between RmpD and Wzc (transmembrane protein), and presumptive indirect interaction with Wzy (capsule repeat unit polymerase), both of which are components of the core capsule synthesis and export machinery [17, 18]. The *rmpA* and *rmpA2* genes appear to be frequently subjected to insertions or deletions (indels) within a poly(G) tract that consequently encode a truncated and presumably non-functional product, and this has been suggested as a mechanism by which differential expression of the two genes is achieved [14].

The presence or absence of *rmpA* in clinical isolates has been investigated in dozens of studies focused on hypervirulent infections, and *rmpA* has been identified as one of several biomarkers for distinguishing hvKp strains from non-hvKp [19]. However, their detection shows a variable degree of correlation with HMV as measured by string test (51–98%). The functional impact of allelic variation in the *rmp* locus genes has not yet been explored, although this is likely to be important for explaining or predicting HMV based on *rmp* locus sequences. Here, we investigate the genetic diversity and distribution of the *rmp* locus in the *K. pneumoniae* species complex (KpSC), identify key variant lineages of *rmp* and their associated mobile genetic elements (MGEs), and demonstrate that representatives of each of the lineages are able to induce HMV and elevated capsule production when introduced into an hvKp isolate lacking its native *rmp* locus (KPPR1S Δ*rmp*).

Methods

Genome sequences and genotyping

The initial screening for *rmpADC* was conducted on the same 2733 *K. pneumoniae* species complex (KpSC) genomes included in a 2018 study [10] examining the aerobactin- and salmochelin-encoding loci *iuc* and *iro*, which are often co-localised with *rmp* on the same mobile genetic elements. The SRST2-table-from-assemblies.py Python script (https://github.com/rrwick/SRST2-table-from-assemblies) [20]

Lam *et al. Genome Medicine*   (2025) 17:36

Page 3 of 13

was used to screen these assemblies for the presence of existing *rmpA* alleles from the virulence database on BIGSdb-*Kp* (37 alleles as of October 2020; https://bigsdb.pasteur.fr/klebsiella/) [21] along with the reference *rmpD* and *rmpC* sequences from Walker et al. [15, 16] with BLAST+ v2.2.21 (≥90% coverage and identity), and novel alleles extracted with the --report_new_consensus flag. Unique alleles from 160/2733 genomes with a typeable *rmp* locus (defined as those in which all genes in the locus could be assigned an allele) were assigned allele numbers, and unique allele combinations were used to define unique '*rmp* sequence types (RmSTs)' using a multi-locus sequence typing (MLST) approach. These *rmpADC* alleles and RmST profiles were then incorporated along with *rmpA2* alleles into the genotyping pipeline Kleborate v2.0.0 and applied to screen 13,156 publicly available *Klebsiella* genomes [22]. From this dataset of 13,156 genomes, 944 genomes positive for *rmp* and/or *rmpA2* (including genomes with an incomplete *rmp* locus) were included for analysis in this study (see Additional file 1: Table S1 for genome accessions, isolate metadata and genotyping information). We also included *rmp*-positive genomes from three additional datasets: 4/208 isolates collected at an Australian hospital in 2002 [23], in addition to 36/392 isolates from the same hospital in 2020, and 8/276 isolates from the Burden of Antibiotic Resistance in Neonates from Developing Societies (BARNARDS) network [24]. Novel *rmpADC* alleles and RmST profiles were added to the BIGSdb-*Kp* database. Clonal groups, and designations of clones as hypervirulent or MDR, were done using previously defined ST-CG assignments [25].

Phylogenetic analyses

For each unique RmST, an alignment of the concatenated *rmpA*, *rmpD* and *rmpC* sequences was generated with MUSCLE v3.8.31 and used as an input for maximum likelihood phylogenetic inference with RAxML v8.2.9 [26] run five times with the generalised time-reversible GTR+gamma model. Similarly, phylogenies were also generated for *rmpA*, *rmpD* and *rmpC* genes individually using the same approach described for RmST. *Rmp* lineages were defined based on monophyletic groups of RmSTs mobilised by a unique mobile element (see below). Phylogenetic trees for genomes belonging to hypervirulent clonal groups CG23, CG65 and CG86 were generated using a core genome distance approach implemented in Pathogenwatch (pathogen.watch) [27]. In brief, the genomes for each CG of interest were uploaded to Pathogenwatch, matched/aligned against the *K. pneumoniae* reference core library (2,172,367 bp) to generate a matrix of pairwise distances that were then used to build a neighbour-joining tree.

Comparison of *rmp* genetic contexts

All contigs containing *rmp* were manually inspected in Bandage v0.8.1 [28] to determine whether the locus was located on the chromosome or on a previously described virulence plasmid (KpVP-1 reference pSGH10 with accession CP025081.1 for contigs with *rmp* lineage 1, i.e. *rmp1*, or KpVP-2 reference Kp52.145 plasmid II with accession FO834905.1 for *rmp2* contigs; both are reference sequences that have been widely used for lab and/or genomic studies). For any assemblies where the *rmp*-containing contig did not match to any of the lineages described above, the contig sequence was screened against the NCBI non-redundant nucleotide database and top hits (i.e. BLASTn results with the highest *E* value corresponding to hits with the highest coverage and identity) noted. A single annotated representative (complete sequence where possible) was selected for each *rmp* lineage to compare the overall genetic structures of the neighbouring regions of *rmp* (i.e. up to 15 kbp upstream and downstream of *rmp*). Annotations were performed with Prokka v1.14.6 [29] and comparisons between the annotated genes were visualised with clinker v0.0.21 (github.com/gamcil/clinker) [30].

Bacterial strains, plasmids and growth conditions

The *K. pneumoniae* strains and plasmids used in this study are listed in Table 1. The strains from which representative *rmp* was amplified were as follows: SGH10 for *rmp1*, 52.145 for *rmp2*, NCTC 13669 for *rmp2A*, KPPR1S for *rmp3* and NCTC 1936 for *rmp4*. Strains were grown at 37 °C in LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl; BD Difco, Ref# 244630). Saturated overnight

Table 1 Strains and plasmids used in this work

Strain or plasmid	Relevant genotype	Reference
E. coli		
DH5α	*F⁻ φ80ΔlacZM15 Δ(lacZYA-argF) U169 deoP recA1 endA1 hsdR17 (r_k⁻ m_k⁻)*	Invitrogen
K. pneumoniae		
KPPR1S	ATCC43816, Rif^r, Str^r	Palacios 2017 [31]
VK371	KPPR1S, Δ*rmpADC*	Walker 2019 [15]
VK646	KPPR1S, Δ*wcaJ*	Walker 2020 [16]
Plasmids		
pMWO-078	Sp^r; p15A *ori* cloning vector, *tetO*	Obrist 2012 [32]
pKW186	*rmpADC* from KPPR1S (*rmp3*)	Walker 2019 [15]
pKW215	*rmpADC* from SGH10 (*rmp1*)	This work
pKW216	*rmpADC* from Kp52.145 (*rmp2*)	This work
pKW222	*rmpADC* from NCTC 1926 (*rmp4*)	This work
pLPT059	*rmpADC* from NCTC 13669 (*rmp2A*)	This work

Lam *et al. Genome Medicine*   (2025) 17:36

Page 4 of 13

cultures were diluted to an OD_{600} of 0.2 and grown for 5.5 h. Antibiotics were used where appropriate: kanamycin (Kan), 50 μg/ml; rifampin (Rif), 30 μg/ml; spectinomycin (Sp), 50 μg/ml. All plasmids were introduced into *K. pneumoniae* by electroporation as previously described [15]. For expression of genes cloned into the pMWO-078 vector, 100 ng/ml anhydrous tetracycline (aTc) was added to the medium at the time of subculture.

Primers used for the construction of expression vectors are listed in Additional file 2: Table S2. All *rmp* loci were amplified from genomic DNA by PCR and then cloned into pMWO-078 [32] by Gibson assembly (NEB). The *rmp* expression vectors were then transformed into KPPR1S Δ*rmp*. KPPR1S (ST493; K2/O1) is a well-characterised streptomycin and rifampicin-resistant derivative of ATCC 43816 and has been used in previous studies to characterise *rmpD*, *rmpC* and HMV [15–17]. Wildtype KPPR1S, Δ*wcaJ* and Δ*rmp* strains transformed with pMWO-078 were also included as control strains for measuring HMV. The *wcaJ* mutant does not produce capsule and is HMV negative, while the *rmp* mutant produces capsule but is HMV negative as it does not produce RmpD.

Assessment of hypermucoviscosity

Saturated overnight cultures of strains containing *rmp* plasmids were subcultured in fresh LB for 5.5 h at 37 °C with 100 ng/ml anhydrotetracycline (ATc) to induce *rmp* expression. Cultures were normalised to 1 OD_{600}/ ml and centrifuged at $1000 \times g$ for 5 min. Mucoviscosity of cultures was determined by normalising OD_{600} of the culture supernatant to the starting culture as previously described, and an $OD_{600} = 0.25$ considered hypermucoviscous [33].

Uronic acid measurement

Uronic acid (UA) was measured following an established protocol [34–36] from cultures grown as described for assessing HMV. Briefly, UA was extracted from 500 μl of culture with 1% zwittergent in 100 mM citric acid at pH 2, precipitated with 1.2 ml 100% ethanol, and resuspended in 1.2 ml 12.5 mM sodium tetraborate dissolved in sulphuric acid. Twenty microlitres of 0.15% 3-phenylphenol in 0.5% NaOH was added and absorbance at 520 nm was measured. UA amounts were determined from a standard curve generated with glucuronolactone (from 1500 to 50 μM).

Results

### Prevalence of *rmp* in the *K. pneumoniae* species complex

Screening for the *rmp* locus in publicly available KpSC genomes ($n = 15$ species/subspecies) identified its presence in 992/13,993 genomes across only three KpSC species: *K. pneumoniae* (980/11,967, 8.2%), *K. quasipneumoniae* subsp. *similipneumoniae* (*Kqs*; 5/522, 1.0%) and *K. variicola* subsp. *variicola* (*Kv*; 7/626, 1.1%) (see Additional file 2: Table S3). The majority of these included complete *rmp* loci with intact coding sequences for *rmpA*, *rmpD* and *rmpC* (i.e. presumably functional variants, detected in 73% of *rmp*+ genomes); however, some loci were incomplete (i.e. deletion variants missing at least one gene within the locus or arising from assembly fragmentation) or had nonsense mutations resulting in premature stop codons truncating one or more of the encoded proteins (i.e. truncation variants) (see Additional file 2: Table S3). Fifteen genomes were identified with multiple *rmp* loci (two *rmp* each); the majority of these carried at least one functional *rmp* locus (73%; eight genomes with two functional *rmp* loci and three genomes with one functional *rmp*) while the remaining four genomes each had two non-functional *rmp* loci arising from deletion and/or truncation variants. BLASTn search of all non-redundant bacteria in NCBI with default parameters identified complete *rmp* loci in four non-*Klebsiella* genomes, in all cases *rmp* was present in plasmids sequenced from *Escherichia coli* transconjugants that had been mated with a *K. pneumoniae* strain carrying KpVP-1-like virulence plasmids (GenBank accessions MN200130.1, MN182750.1, MZ475697.1, CP068571.1). Incomplete *rmp* loci (44% coverage) were detected in an additional five plasmids that were also sequenced from *E. coli* transconjugants.

Genetic diversity of the *rmp* locus

Each of the genes comprising the *rmp* locus showed some degree of genetic diversity, with 85 *rmpA* alleles, 77 *rmpD* alleles and 42 *rmpC* alleles observed across *rmp*+ genomes. For the 825 genomes carrying an intact 'typeable' *rmp* locus (838 loci total accounting for genomes with two *rmp* loci), allelic variants could be assigned to all three genes, and resulted in 170 unique combinations which were each assigned a unique RmST (*rmp* sequence type). The *rmp* loci from the remaining 167 genomes were designated 'non-typeable' due to the locus being incomplete (i.e. missing at least one gene or comprising a fragmented gene to which an allele could not be designated).

Maximum likelihood phylogenetic analysis of the concatenated *rmpA*, *rmpD* and *rmpC* sequences belonging to each unique RmST revealed the grouping of sequences into five distinct lineages, which we labelled *rmp1*, *rmp2*, *rmp2A*, *rmp3* and *rmp4* (Fig. 1). These lineages were labelled as such to match with their associated *iro* and/or *iuc* lineages (i.e. *rmp1* is associated with *iuc1* and *iro1*). The nucleotide divergence between lineages ranged from 0.7 to 11% (mean 4.7%), decreasing to 0–5.9% within lineages (mean 0.4%). No

Lam *et al. Genome Medicine*   (2025) 17:36

Page 5 of 13
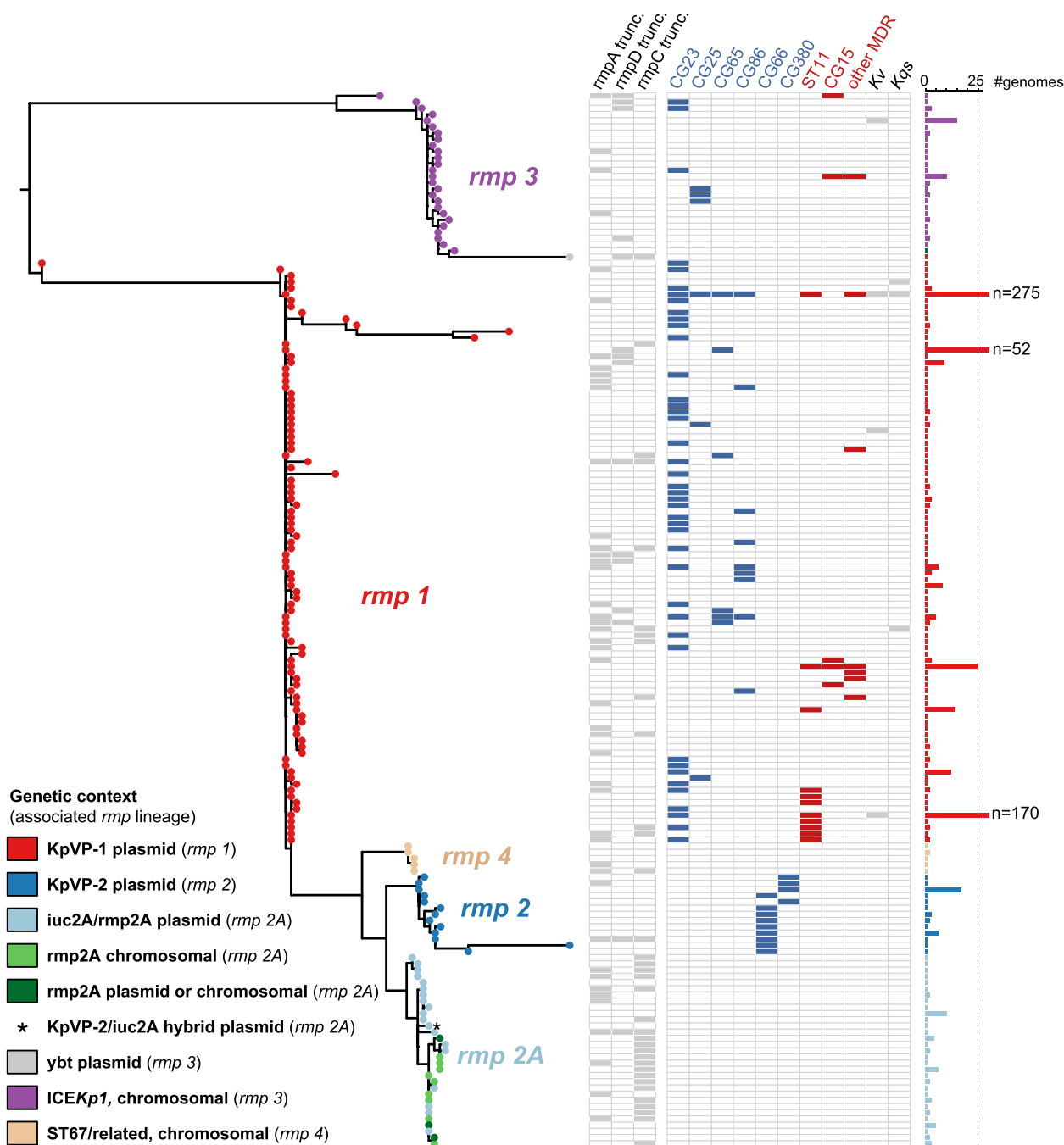
Fig. 1 Maximum-likelihood phylogenetic tree of *rmp* sequence types (RmSTs). Lineages are labelled and tips are coloured by the associated mobile genetic element according to the legend. Column labels are as follows for a particular RmST: presence or absence of truncations in the *rmpA*, *rmpD* and *rmpC*, detection within a hypervirulent (blue) or MDR (red) clone (definitions of clones from Wyres et al. [25]) or non-*K. pneumoniae* species (Kv, *Klebsiella variicola*; Kqs, *Klebsiella quasipneumoniae* subsp. *similipneumoniae*). Clones included in 'Other MDR': CG14, CG17, CG29, CG101, CG147, CG152 and CG307. The number of genomes from which each RmST was detected is shown in the bar graph on the right-hand side

rmp gene alleles were shared between lineages (see individual gene phylogenies in Additional file 2: Fig. S1), suggesting an absence of recombination between the lineages. The *rmp3* lineage was notably divergent from the other *rmp* lineages, with a mean divergence of 8.4% compared to 1.8% between all other lineages (Table 2). Further, the *rmpD* alleles of the *rmp3* lineage were longer than those of other lineages, measuring

Lam *et al. Genome Medicine*   (2025) 17:36

Page 6 of 13

Table 2 Nucleotide divergence within (bolded) and between *rmp* lineages. Values given are range (mean) and expressed as percentages

	rmp1	*rmp2*	*rmp2A*	*rmp3*	*rmp4*
rmp1	**0–5.9 (0.4)**				
rmp2	1.8–6.1 (2.4)	**0–2.3 (0.6)**			
rmp2A	1.6–4.9 (1.9)	0.7–3.2 (1.3)	**0–0.6 (0.3)**		
rmp3	4.8–11 (8.4)	7.1–9.4 (8.5)	6.8–8.9 (8.1)	**0–3.8 (0.5)**	
rmp4	1.6–4.7 (1.9)	1.3–3.0 (1.7)	1.2–1.8 (1.6)	7.0–9.1 (8.4)	**0–0.1 (0.05)**

176–177 bp in length compared to the 151–162 bp alleles observed in *rmpD* from the other lineages. No associations between *rmp* lineages with geography or host were observed.

Diversity of *rmp*-associated mobile genetic elements

Each of the *rmp* lineages were associated with unique plasmids or chromosomal contexts (see Fig. 1), many of which have been previously characterised in a study examining the aerobactin- and salmochelin-encoding loci *iuc* and *iro* [10], which are typically co-localised with *rmpA2* and *rmp* respectively on the same mobile elements. The dominant *K. pneumoniae* virulence plasmid, KpVP-1 (associated with *iuc1* and *iro1* loci) was associated with the *rmp1* lineage and accounted for the majority of *rmp* carriage (785 genomes, 79.1%). Next most common were the *iuc2A* virulence plasmids (associated with *iuc2A*) [10], carrying *rmp2A* lineage (77 genomes, 7.8%); followed by ICE*Kp1* carrying *rmp3* (58 genomes, 5.8%); and KpVP-2 virulence plasmids with *rmp2* (associated with *iuc2* and *iro2*, 45 genomes, 4.6%). Two additional genetic contexts were observed in this study: (i) the *rmp4* lineage was chromosomally encoded and associated exclusively with *K. pneumoniae* CG67 genomes (six genomes, identified as ST67 or single-locus variants thereof, also known as *K. pneumoniae* subspecies rhinoscleromatis [37]) and (ii) presence of *rmp3* into the yersiniabactin-encoding *ybt4*-type plasmids (three genomes; note partial deletion of *rmp* locus in two of these genomes). For the remaining genomes, one was associated with a KpVP-2/*iuc2A* hybrid plasmid as previously described [10], nine carried multiple *rmp* loci (due to presence of both KpVP-1 and ICE*Kp1*), and 17 were unresolved due to assembly issues. KpVP-1 was detected in 785 genomes, where it typically carried both *rmp* and *rmpA2* (669/785). However, variants of KpVP-1 were common (i.e. coverage of KpVP-1 reference ranged from 26.7% coverage in *rmp1*+ genomes): we observed 116/785 genomes with *rmp* only without *rmpA2* (20 and 32 of these lacking an intact *iro1* and *iuc1*, respectively). An additional 180 genomes carried

rmpA2 only without *rmp* (849 *rmpA2*+ genomes in total). Nine genomes appeared to have chromosomal integration of KpVP-1 (*n* = 7 ST11/LV, *n* = 1 ST23 and *n* = 1 ST383; see Additional file 1: Table S1).

The genetic context of *rmp* in representatives of each of the key *rmp* lineages is shown in Fig. 2. The *rmp* locus was located adjacent to the *iro* locus and *peg-344* gene (which has been mis-labelled as *pagO* in some studies) in most cases, with the exception of the *ybt4* plasmids (which lack *iro* and *peg-344* genes) and *iuc2A* plasmids (which carry *iroB* only). The *iro* locus is typically intact in KpVP-1, KpVP-2 and ICE*Kp1*, but is partially disrupted by insertion sequences (IS) in the ST67 (rhinoscleromatis) chromosome and *iuc2A* plasmids. IS flank the *rmp*/*iro* region in all contexts; however, the specific IS vary and it is not clear which have played a role in mobilising *rmp* and/or *iro*. Notably, all contexts include IS*3* on one or both ends of the *rmp*/*iro* region. The closely related *rmp1*, *rmp2*, *rmp2A* and *rmp4* lineages all have IS*3* in a conserved position adjacent to *rmpA*; this also appears to be conserved in the divergent lineage *rmp3*, suggesting it was likely present in the common ancestor of all *rmp* lineages and may have played a role historically in the mobilisation of *rmp*/*iro* between genetic backgrounds (see Fig. 2). However, all *rmp/iro* regions have additional IS near or within the locus, some of which may have contributed to further mobilisation and/or degradation of the distinct lineages over time. While the assembly of the *rmp* region in the *ybt4* plasmid was incomplete, an IS*3* fragment was detected next to *rmp.*

Hypermucoviscosity of *rmp* variants in a KPPR1S background

To determine if there were functional differences between the level of HMV conferred by the *rmp* lineages, we ectopically expressed variants of the *rmp* locus cloned into pMWO-078 in an HMV-negative *rmp* deletion mutant KPPR1S Δ*rmp*. Wildtype KPPR1S is a ST493 K2/O1 isolate containing *rmp3* (ICE*Kp1*), which is the most divergent *rmp* lineage (Fig. 1); KPPR1S Δ*rmp* therefore provides a good background to assess potential functional differences between *rmp* variants. HMV was assessed using a sedimentation assay. The Δ*wcaJ* and Δ*rmp* vector controls fully sedimented, whereas the KPPR1S vector control did not sediment well (Fig. 3A). Ectopical expression of the native locus, *rmp3*, conferred elevated HMV in Δ*rmp* compared to the KPPR1S vector control. This elevated phenotype is likely due to the overexpression of *rmp*, and this complemented strain therefore serves as a reference to compare the effects of expressing the other *rmp* loci. Expression of the *rmp* genes from *rmp1*, *rmp2* and *rmp2A* also conferred elevated HMV levels in Δ*rmp* (Fig. 3A), and there were no statistically significant differences in sedimentation resistance between these strains

Lam *et al. Genome Medicine* (2025) 17:36

Page 7 of 13
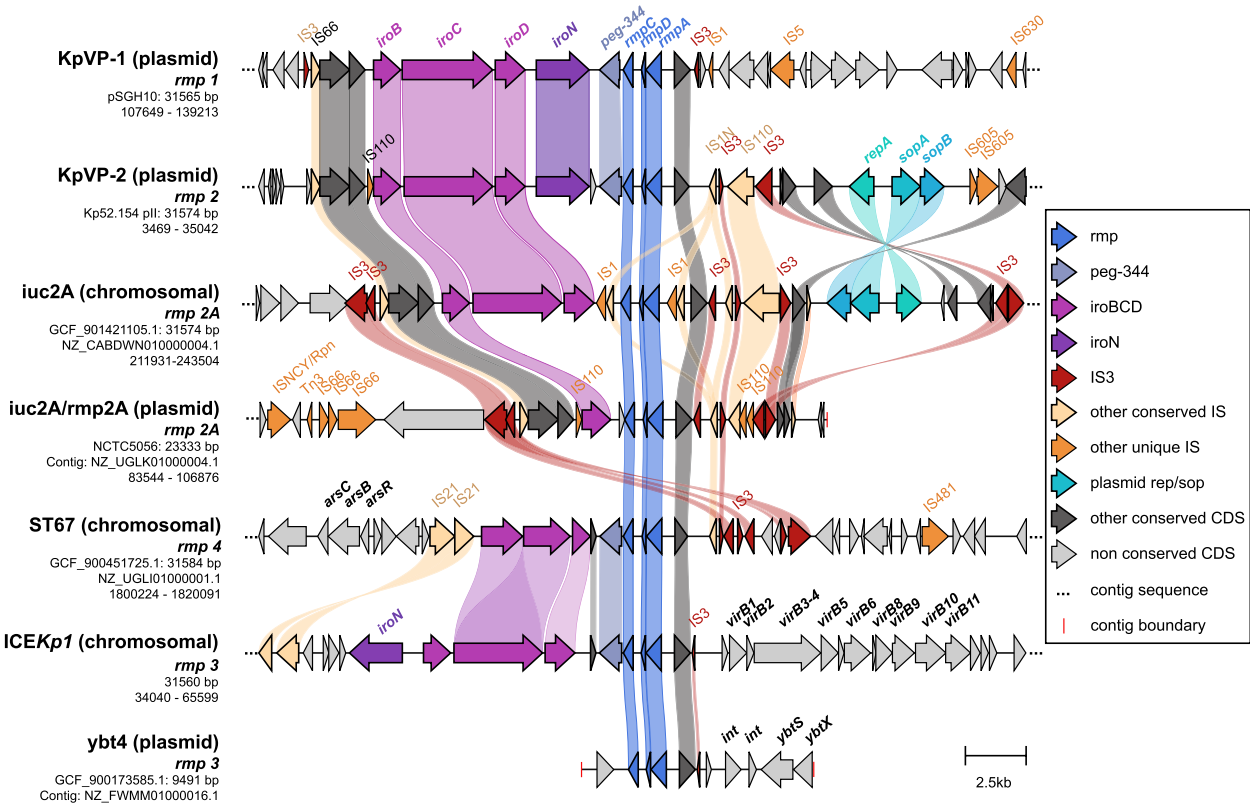
Fig. 2 Comparison of upstream and downstream regions of *rmp* associated with different mobile variants. Arrows represent coding sequences and those corresponding to genes of interest are labelled and coloured by functionality as per the legend. Other labelled loci include *arsCBR* (arsenic resistance), *virB1-virB11* (*virB*-type 4 secretion system) and *ybtSX* (part of the yersiniabactin locus). Shading corresponds to regions of similarity (amino acid sequence identity ≥ 30%) as identified by clinker
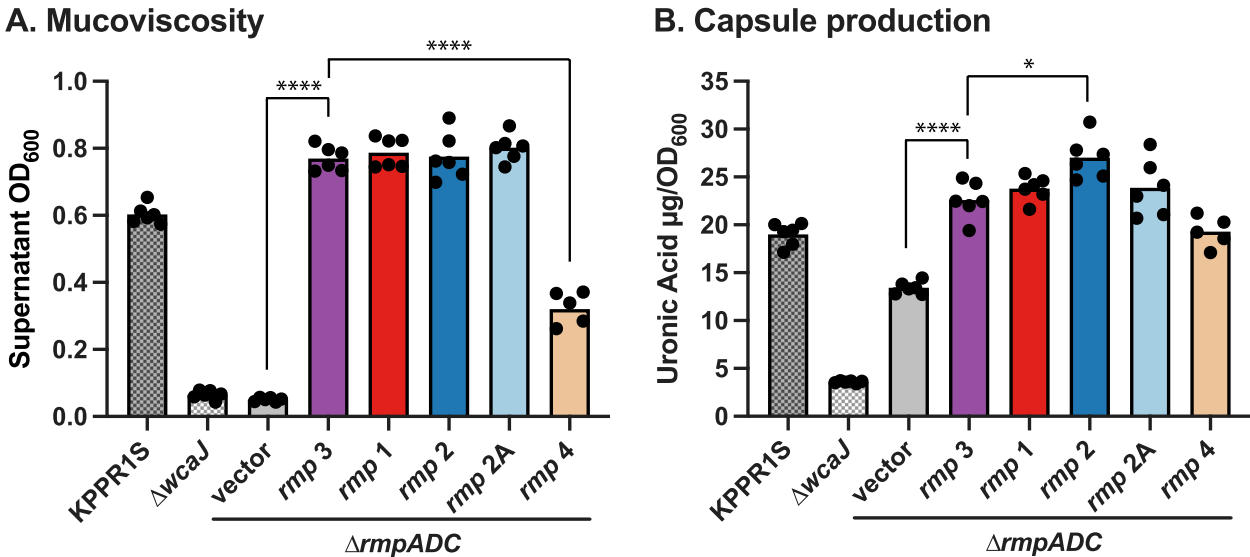
Fig. 3 Hypermucoviscosity and capsule production of ectopically expressed *rmp* loci from the five different lineages. **A** Mucoviscosity assay and **B** uronic acid assay of KPPR1S, Δ*wcaJ* and Δ*rmp* strains with vector (pMWO-078) or lineage-specific pRmp (see Table 1). Data from two technical replicates and three biological replicates were obtained following a 5.5-h induction of KPPR1S expression plasmid-borne *rmp* genes as described in Methods. One-way ANOVA with Tukey's post-test was used to determine significance. *, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$; comparisons not indicated were not significant

Lam *et al. Genome Medicine*   (2025) 17:36

Page 8 of 13

versus that expressing *rmp3*. However, expression of *rmp4* conferred a significantly lower level of HMV compared to the other lineages.

Capsule production of *rmp* variants in a KPPR1S background

Elevated capsule production is an *rmpC*-dependent phenotype in KPPR1S, independent of HMV, whereby RmpC upregulates capsule expression via the promoters upstream of the *galF* and *manC* genes in the K locus, and KPPR1S Δ*rmp* mutants display decreased capsule production [15, 16]. To further investigate any functional differences between *rmp* variants, we assayed the same transformants described above for uronic acid (UA) production, which is used as an indicator of the amount of capsule [34]. KPPR1S, Δ*wcaJ* (i.e. capsule-negative) and Δ*rmp* strains were included as controls. Expression of *rmp* from each of the lineages increased capsule production (Fig. 3B). Expression of *rmp1*, *rmp2A* and *rmp4* in the Δ*rmp* mutant increased capsule production to the same level observed when expressing the native *rmp3* locus (Fig. 3B). Expression of *rmp2* produced slightly but significantly higher uronic acid levels compared to *rmp3* ($P < 0.05$). Collectively, these data suggest that there are subtle but significant functional differences between the *rmp* lineages.

Distribution of *rmp* in the *K. pneumoniae* species complex

The *rmp* locus was detected in 129 unique *K. pneumoniae* STs, 3 *Kqs* STs and 5 *Kv* STs, representing 143 unique combinations of STs and *rmp* lineages. As expected, the prevalence of *rmp* is typically quite high within known hvKp clones (≥80% prevalence within any given ST assigned to a hypervirulent clonal group with the exception of CG25, which has recently been shown to comprise of two distinct lineages [38]; mean 83.6%) (Fig. 4). *Rmp* was also common in CG67 (rhinoscleromatis, 100% *rmp4*; 25% truncated) and CG91 (ozaenae, 79% *rmp2A* incomplete or truncated); both lineages have previously been defined as subspecies due to their distinct pathotypes [37]. However, relatively high *rmp* prevalence was also observed in some 'generalist' (i.e. non-hvKp and non-MDR) *K. pneumoniae* clones (for generalist clones with $n \geq 5$ genomes: 1.6–100% *rmp* prevalence, mean *rmp* prevalence = 40.2%, median *rmp* prevalence =

15.1% [IQR = 5.6–86.4%]). Conversely, the MDR clones together with CG36 and CG45, which are also clones that are commonly detected in healthcare settings, had relatively lower frequencies of *rmp* (mean prevalence = 11.4%).

KpVP-1-*rmp1* was the most widely disseminated and was not only detected in the hypervirulent clones from which they were initially characterised but also in MDR clones. Numbers of KpVP-1-*rmp1* were notably high in ST11 (211 genomes) and ST15 (50 genomes), reflecting recently reported MDR-hypervirulent convergent variants of these well-known MDR clones [39]. KpVP-1 also accounted for the majority of *rmp* acquisitions in the other KpSC species where the genetic context could be resolved. In comparison, the other mobile elements appeared in relatively fewer STs, but also appeared to be fixed in most of these clones that likely serve as native hosts to the respective mobile elements (see Fig. 4). Of the fifteen genomes that carried multiple *rmp* loci, ten were ST23 genomes that carried ICE*Kp1* (*rmp3*) on the chromosome and the KpVP-1 (*rmp1*) virulence plasmid.
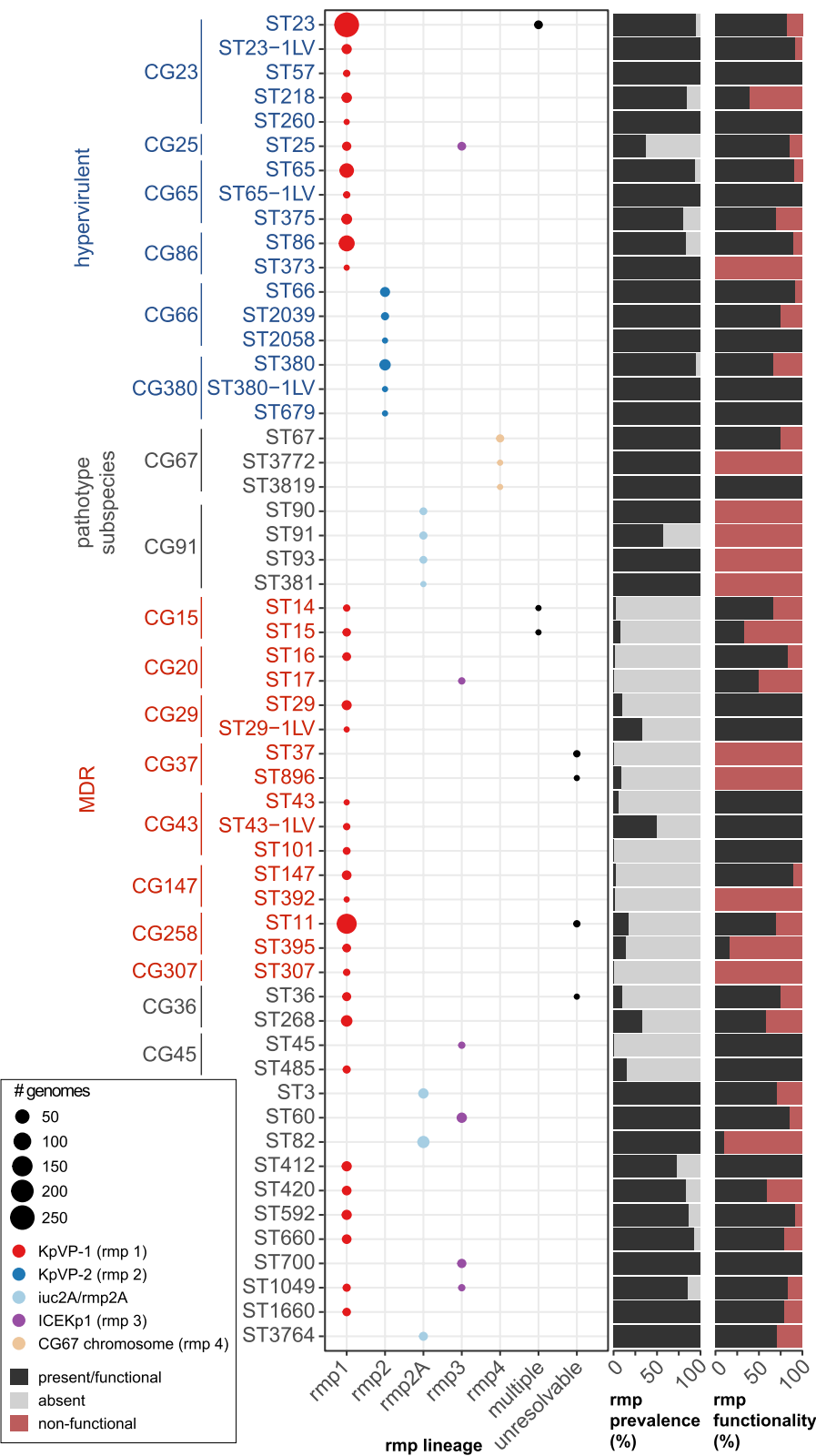
Amongst *rmp*+ genomes with a confident K locus call (917/992), 38 different K loci were detected (Fig. 5). Half (i.e. 19) were detected in two or fewer genomes. The most common K loci were KL1 (319 genomes), KL2 (203) and KL64 (140), which collectively account for 72.2% of genomes with a confident K locus, followed by KL57 (45) and KL20 (40). The dominant K loci are largely driven by their associations with over-represented STs (Fig. 5), many of which are known hvKp lineages such as KL1 in CG23 (which harbours KpVP-1) and KL2 in CG380 (KpVP2), CG65 and CG86 (KpVP-1). The next most common K locus, KL64, is associated with a sublineage of MDR clone ST11 that is widespread in China, known to have acquired variants of KpVP-1 [39].

Given previous reports and observations highlighting the common occurrence of truncations in *rmpA* due to insertions and deletions (indels) within a poly-G tract [14], we next examined the *rmp* locus for the occurrence of truncations. Truncations were detected in 30/85 *rmpA*, 9/77 *rmpD* and 16/42 *rmpC* alleles across 63, 16 and 55 genomes, respectively (Additional file 2: Fig. S1). The majority of truncations (45/55 truncated alleles) were caused by indels within homopolymer
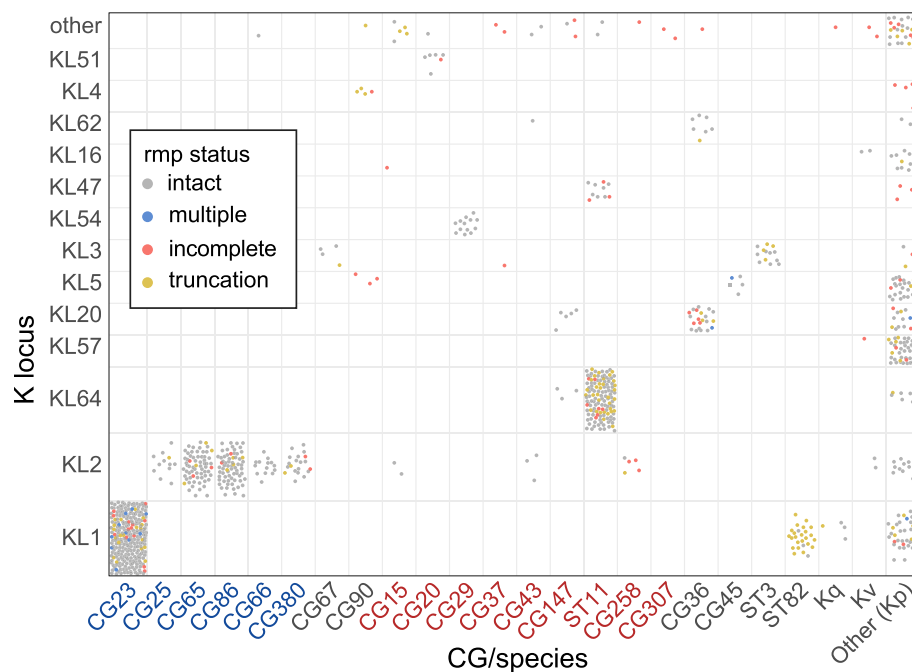
(See figure on next page.)

Fig. 4 Distribution of *rmp* across *K. pneumoniae* sequence types with ≥5 *rmp*+ genomes. Rows indicated *K. pneumoniae* sequence types (STs; as labelled), and grouped and labelled by clonal group (CG) where applicable; those corresponding to hypervirulent, pathotype subspecies or multidrug-resistant (MDR) clones are labelled accordingly. Other clones are considered 'generalist' clones. The bubble plot shows the number of *rmp*+ genomes assigned to a particular *rmp* lineage, coloured by the type of mobile element. The first bar plot shows prevalence of *rmp* (i.e. black: present, grey: absent) and the second shows functional status of the locus (i.e. black: intact and functional, red: non-functional due to deletions or truncations)

Lam *et al. Genome Medicine*   (2025) 17:36

Page 9 of 13

Fig. 4 (See legend on previous page.)

Lam *et al. Genome Medicine*   (2025) 17:36

Page 10 of 13

Fig. 5 Distribution of *rmp* within K locus groups and KpSC clonal groups (CG) or species. Each circle represents an *rmp+* genome with a confident K locus call (917 genomes) and is coloured by the functionality of the *rmp* locus as per the figure legend. Twenty-five K loci with 4 or fewer genomes were collapsed into 'Other': KL6, KL10, KL12, KL14, KL15, KL17, KL19, KL21, KL23, KL24, KL25, KL27, KL30, KL34, KL35, KL39, KL45, KL55, KL63, KL102, KL108, KL127, KL137, KL147 and KL149. Clonal group labels are coloured blue and red to indicate hypervirulent and MDR lineages, respectively. Kq: *Klebsiella quasipneumoniae*, Kv: *Klebsiella variiicola*

tracts resulting in premature stop codons arising from frameshift mutations (Additional file 3: Table S4). Further, some homopolymer regions had a higher frequency of mutations compared to others, including the previously reported poly-G tract in *rmpA* spanning nucleotide position 276 to 285 whereby indels were detected in 16 *rmpA* variants across 46 genomes. The frequency of truncated *rmp* loci compared to intact loci was lower within the hypervirulent clones and appeared to be more common outside of these clones (Figs. 4 and 5). Closer inspection of the phylogenies for hypervirulent clones CG23, CG65 and CG86 (Additional file 2: Figs. S2–S4) revealed a random distribution of truncated *rmp* alleles throughout these populations. Truncations in *rmpA2* (*n* = 745/849) were also common, outnumbering intact variants (*n* = 95 *rmpA2* intact; note *n* = 9 with truncation status not known due to fragmented *rmpA2* sequences) and were detected at high prevalence (>55 to 100%) across both hypervirulent and MDR clones.

Discussion

We previously reported briefly on the sequence variation of the *rmp* locus, which clustered into four lineages and were each associated with distinct MGEs [22]. In this study, we provide detailed insights into the phylogenetic relationships, genetic contexts, distribution and phenotypic functionality of *rmp* variants. Building on the initial *rmp* scheme, we characterised an additional *rmp4* lineage observed only in genomes belonging to a sublineage of *K. pneumoniae* (CG67), updating the number of *rmp* lineages to five.

The 1:1 correlation typically observed between *rmp* lineages and the MGEs that mobilise them (with few exceptions including *rmp3* on *ybt4* related plasmids, and the chromosomal versus plasmid variants of *rmp2A*) also extends to *iro*, shown here to be typically located adjacent to *rmp* (Fig. 2), and to *iuc*, which is often located elsewhere on the same plasmids [10]. This association highlights the co-evolution of these virulence loci, and presumably other common genes shared between the MGEs. While specific details or the order of events in the evolutionary history of these MGEs cannot be determined, similarities in the genetic contexts surrounding *rmp* do suggest a shared ancestry and the role of IS, particularly IS*3*, in the mobilisation of *rmp/iro* between different MGEs. Interestingly, the presence of IS*3* upstream of *rmpA2* has been flagged as being necessary for complete activation of the promoter for *rmpA2* and may likewise serve a similar purpose for *rmpA* [40].

Lam *et al. Genome Medicine*   (2025) 17:36

Page 11 of 13

We also observed three instances of a novel genetic context for *rmp*, where a variant of *rmp3* (ICE*Kp1*) had been introduced into a yersiniabactin plasmid (i.e. another virulence MGE) [41] likely via IS*3*-mediated transposition (Fig. 2). Outside of *K. pneumoniae*, *rmp* (KpVP-1 and ICE*Kp1*) was only detected in two other KpSC species at rare frequencies and did not appear to have been acquired naturally in any other species. *K. pneumoniae* is therefore likely to be the original host for *rmp*, although the same may not hold true for the other virulence loci given that *iro* and *iuc* variants have been detected in other non-*Klebsiella* species including *Enterobacter* spp. and *Escherichia coli* [10].

Even with the larger genome dataset screened in this present study, the prevalence of *rmp* overall (7.5%) and each of the MGEs was very similar to that reported in our earlier study characterising *iuc* (8.7%) and *iro* (7.2%) (i.e. 14,000 versus 2700 genomes). The *rmp1* lineage (KpVP-1) was by far the most dominant accounting for 80% of the *rmp* burden, in part driven by their association and maintenance via clonal expansion within the key hypervirulent clones such as CG23 [11], but also due to the more recent transmission and spread within MDR clones such as ST11 and ST15 (Figs. 1 and 4). Based on growing reports of 'hypervirulent ST11' from China, it is possible that KpVP-1 and variants thereof are also being stably maintained within this clone following its acquisition. Variants of KpVP-1 which include deletion or recombinant variants, and chromosomal integrations, are commonly reported in ST11 and have been proposed to confer a reduced fitness cost [39, 42, 43]. A more recent study has highlighted at least two plasmid variants of KpVP-1 (referred to as 'PIUC1-IncFIB(K)$_{37}$' and 'pIUC1-IncFIB(Mar)') that account for the majority of hypervirulent carbapenem-resistant *K. pneumoniae* in China [44]. In another study investigating the virulence plasmid in a collection of ST11 genomes, the authors noted as low as 25% coverage of their chosen KpVP-1 reference (pK2044, accession AP006726) across recombinant KpVP-1 plasmids, and 19–41% coverage for chromosomal integrations of KpVP-1 [41]. A similar range in coverage was also observed in this dataset not only in ST11 but also other clones, and manual inspection of these assemblies suggested the presence of several KpVP-1 plasmid variants as well as chromosomal integrations. With the exception of KpVP-1-*rmp1* and ICEKp1-*rmp3*, which were detected in multiple STs/CGs, the remaining MGEs appeared to be stably conserved within a small number of clones in which they were detected (93.9–100% prevalence; Fig. 4).

Each of the *rmp* lineages was shown to be functional, resulting in HMV and elevated capsule production when a representative of each was introduced into a single strain background, KPPR1S Δ *rmp*. Notably, this ST493 strain has a K2 capsule, which is one of two dominant capsule types amongst known hvKp clones such as CG86 (Fig. 5). We also observed variability in the extent of HMV and capsule production for different *rmp*, which further reiterates that HMV and capsule production are two separable traits [15, 16]; for example, the expression of *rmp4* does not restore HMV to the same level as KPPR1S *rmp3* but does restore capsule production (Fig. 3). It is unclear if expressing these *rmp* (or other representatives of the same lineages) in other KpSC strains with different serotypes will have the same impact, especially given the extensive diversity of Wzc in different K loci and its interaction with RmpD. In another study, expression of KpVP-1 *rmp* in strains with different capsules (KL1, KL2 and KL64) yielded differences in HMV and capsule production [45]. These insights will be particularly useful for predicting the impacts of *rmp* acquisition in other clones, including those considered to be MDR, and is the subject of ongoing work by our team.

Loss of function mutations of *rmpA* arising from indels within the poly(G) tract has been well documented in many studies [14], and while this site (i.e. bases 267 to 285 of the *rmpA_2* reference) does account for the majority of *rmpA* truncations in this dataset, indels within additional homopolymer tracts in *rmpA*, *rmpD* and *rmpC* were also observed (Additional file 3: Table S4). Loss of function mutations in the *rmp* genes were observed in at least 145 genomes (14.6% of *rmp+* genomes, including three with multiple *rmp*), most often in the non-hypervirulent clones. Further, we observed parallel evolution of the same loss of function mutations on multiple occasions within different hypervirulent clones (e.g. allele *rmpA_4* in CG23, CG65 and CG86; Additional file 2: Figs. S2–S4). Taken together, these findings highlight the potential reversibility of homopolymer tract mutations, which may take place to help alleviate the negative selection pressure following acquisition of KpVP-1 or other MGEs. Importantly, these loss of function mutations also need to be carefully considered when interpreting data based solely on PCR detection of *rmpA/A2*, and may partly explain the discrepancies in the literature reporting on the association between *rmp* presence and HMV, although other reasons also include inconsistencies and the unreliability of string testing, which is heavily influenced by temperature dependencies and strain genetic background [46].

Conclusions

Our findings reveal that, similar to the other key virulence loci co-localised on the same MGEs, genetic variation within the *rmp* locus is highly structured and this information can be harnessed to track novel *rmp* acquisitions. To this end, detection and genotyping of the *rmp* genes (i.e. RmST typing), alongside the reporting of

locus disruptions, has already been implemented in our genotyping tool for KpSC genomes, Kleborate (github. com/klebgenomics/Kleborate). The tool also outputs a virulence score that is currently calculated from the presence of various siderophores (*ybt* and iuc) and colibactin (*clb*), and does not take into account *rmp* (or *iro*). Given the apparent role of *rmp* in hypervirulence, the virulence score in addition to *rmp* presence should both be considered when assessing the virulence of a given strain or genome. Ongoing research investigating the expression of *rmp* variants in different strain and capsule backgrounds will yield additional important insights into the function of *rmp*, particularly for novel strains such as those from the MDR clones that acquire them, and can be used to further improve the genotyping output for *rmp*.

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s13073-025-01461-5.

Additional file 1: Table S1 Strain information and Kleborate genotyping output for genomes included in this study

Additional file 2: Table S2 Primers used in this study. Table S3 Distribution of *rmp* loci by species. Fig. S1 Phylogenetic relationships of the A. *rmpA*, B. *rmpD* and C. *rmpC* genes. Fig. S2 Distribution of *rmpADC* allelic variants and KpVP-1 associated virulence loci *rmpA2*, *iuc* and *iro* in *Klebsiella pneumoniae* clonal group 23. Fig. S3 Distribution of *rmpADC* allelic variants and KpVP-1 associated virulence loci *rmpA2*, *iuc* and *iro* in *Klebsiella pneumoniae* clonal group 65. Fig. S4 Distribution of *rmpADC* allelic variants and KpVP-1 associated virulence loci *rmpA2*, *iuc* and *iro* in *Klebsiella pneumoniae* clonal group 86

Additional file 3: Table S4 Description of truncated allelic variants of *rmpA*, *rmpD* and *rmpC*

Authors' contributions
Acquisition of data: MMCL, SMS, LPT, LMJ, KAW Analysis and interpretation of data: MMCL, SMS, LPT, KLW, SB, KAW, VLM, KEH Writing of code: RRW and KEH Initial drafting of manuscript: MMCL with input from KEH and KLW Conceptualisation of study: KEH with input from VLM and KAW All authors read and approved the final manuscript.

Data availability
All whole-genome sequences analysed in this study are publicly available on NCBI, and the accession numbers are listed in Additional file 1: Table S1. The RmST scheme is available in the *K. pneumoniae* BIGSdb database (https:// bigsdb.pasteur.fr/klebsiella/) [21] and in the Kleborate distribution (https:// github.com/klebgenomics/Kleborate) [47].

Declarations

Author details
[1]Department of Infectious Diseases, School of Translational Medicine, Monash University, The Burnet Institute, Level 285 Commercial Rd, Melbourne 3004, Australia. [2]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC, USA. [3]Centre to Impact AMR, Monash University, Melbourne, Australia. [4]Biodiversity and Epidemiology of Bacterial Pathogens, Institut Pasteur, Université Paris Cité, Paris, France. [5]Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC, USA. [6]Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, UK.

References
1. Podschun R, Ullmann U. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin Microbiol Rev. 1998;11:589–603.
2. Shon AS, Bajwa RPS, Russo TA. Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. Virulence. 2013;4:107–18.
3. Russo TA, Marr CM. Hypervirulent *Klebsiella pneumoniae*. Clin Microbiol Rev. 2019;32:e00001–19.
4. Lee IR, et al. Differential host susceptibility and bacterial virulence factors driving *Klebsiella* liver abscess in an ethnically diverse population. Sci Rep. 2016;6:29316.
5. Choby JE, Howard-Anderson J, Weiss DS. Hypervirulent *Klebsiella pneumoniae* – clinical and molecular perspectives. J Intern Med. 2020;287:283–300.
6. Struve C, et al. Mapping the evolution of hypervirulent *Klebsiella pneumoniae*. MBio. 2015;6:1–12.
7. Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. Nat Rev Microbiol. 2020;18:344–59.
8. Lam MMC, Wick RR, Judd LM, Holt KE, Wyres KL. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. Microb Genomics. 2022;8(3):000800.
9. Lin TL, Lee CZ, Hsieh PF, Tsai SF, Wang JT. Characterization of integrative and conjugative element ICE*Kp1*-associated genomic heterogeneity in a *Klebsiella pneumoniae* strain isolated from a primary liver abscess. J Bacteriol. 2008;190:515–26.
10. Lam MMC, et al. Tracking key virulence loci encoding aerobactin and salmochelin siderophore synthesis in *Klebsiella pneumoniae*. Genome Med. 2018;10:77.
11. Lam MMC, Wyres KL, et al. Population genomics of hypervirulent *Klebsiella* pneumoniaeclonal-group 23 reveals early emergence and rapid global dissemination. Nat Comms. 2018;9(1):2703.
12. Nassif X, Fournier J, Arondel J, Sansonetti PJ. Mucoid phenotype of *Klebsiella pneumoniae* is a plasmid-encoded virulence factor. Infect Immun. 1989;57:546–52.
13. Hsu CR, Lin TL, Chen YC, Chou HC, Wang JT. The role of *Klebsiella pneumoniae rmpA* in capsular polysaccharide synthesis and virulence revisited. Microbiology. 2011;157:3446–57.
14. Cheng HY, et al. RmpA regulation of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* CG43. J Bacteriol. 2010;192:3144–58.

Lam *et al. Genome Medicine*   (2025) 17:36

Page 13 of 13

15. Walker KA, et al. A *Klebsiella pneumoniae* regulatory mutant has reduced capsule expression but retains hypermucoviscosity. MBio. 2019;10:e00089–e119.

16. Walker KA, Treat LP, Sepúlveda VE, Miller VL. The small protein RmpD drives hypermucoviscosity in *Klebsiella pneumoniae*. MBio. 2020;11:e01750–e1820.

17. Ovchinnikova OG, et al. Hypermucoviscosity regulator RmpD interacts with Wzc and controls capsular polysaccharide chain length. MBio. 2023;14:e00800–e823.

18. Whitfield C, Wear SS, Sande C. Assembly of bacterial capsular polysaccharides and exopolysaccharides. Annu Rev Microbiol. 2020;74:521–43.

19. Russo TA, et al. Identification of biomarkers for the differentiation of hypervirulent *Klebsiella pneumoniae* from classical *K. pneumoniae*. J Clin Microbiol. 2018;56(9):e00776–18.

20. Wick RR. SRST2-table-from-assemblies.py Python script. https://github.com/rrwick/SRST2-table-from-assemblies.

21. BIGSdb-Pasteur. *Klebsiella pneumoniae* species complex genotyping databases. https://bigsdb.pasteur.fr/klebsiella/.

22. Lam MMC, et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nat Commun. 2021;12:4188.

23. Jenney AW, et al. Seroepidemiology of *Klebsiella pneumoniae* in an Australian tertiary hospital and its implications for vaccine development. J Clin Microbiol. 2006;44:102–7.

24. Sands K, et al. Characterization of antimicrobial-resistant Gram-negative bacteria that cause neonatal sepsis in seven low- and middle-income countries. Nat Microbiol. 2021;6:512–23.

25. Wyres KL, et al. Distinct evolutionary dynamics of horizontal gene transfer in drug resistant and virulent clones of *Klebsiella pneumoniae*. PLoS Genet. 2019;15: e1008114.

26. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22:2688–90.

27. Argimón S, et al. Rapid genomic characterization and global surveillance of *Klebsiella* using Pathogenwatch. Clin Infect Dis. 2021;73:S325–35.

28. Wick RR, Schultz MB, Zobel J, Holt KE. Genome analysis Bandage: interactive visualization of de novo genome assemblies. 2015;31:3350–2.

29. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

30. Gilchrist CLM, Chooi Y-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics. 2021;37:2473–5.

31. Palacios M, Broberg CA, Walker KA, Miller VA. A serendipitous mutation reveals the severe virulence defect of a *Klebsiella pneumoniae fepB* mutant. mSphere. 2017;2(4):e00341–17. https://doi.org/10.1128/msphere.00341-17.

32. Obrist MW, Miller VL. Low copy expression vectors for use in *Yersinia* sp. and related organisms. Plasmid. 2012;68:33–42.

33. Bachman, M. A. *et al.* Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung infection. MBio. 2015;6. https://doi.org/10.1128/mbio.00775-15.

34. Lawlor MS, Hsu J, Rick PD, Miller VL. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. Mol Microbiol. 2005;58:1054–73.

35. Blumenkrantz N, Asboe-Hansen G. New method for quantitative determination of uronic acids. Anal Biochem. 1973;54:484–9.

36. Domenico P, Schwartz S, Cunha BA. Reduction of capsular polysaccharide production in *Klebsiella pneumoniae* by sodium salicylate. Infect Immun. 1989;57:3778–82.

37. Brisse S, et al. Virulent clones of *Klebsiella pneumoniae*: identification and evolutionary scenario based on genomic and phenotypic characterization. PLoS One. 2009;4(3):e4982.

38. Pei N, et al. Large-scale genomic epidemiology of *Klebsiella pneumoniae* identified clone divergence with hypervirulent plus antimicrobial-resistant characteristics causing within-ward strain transmissions. Microbiol Spectr. 2022;10:e02698–e2721.

39. Liao W, Liu Y, Zhang W. Virulence evolution, molecular mechanisms of resistance and prevalence of ST11 carbapenem-resistant *Klebsiella pneumoniae* in China: a review over the last 10 years. J Glob Antimicrob Resist. 2020;23:174–80.

40. Yi-Chyi L, Hwei-Ling P, Hwan-You C. RmpA2, an activator of capsule biosynthesis in *Klebsiella pneumoniae* CG43, regulates K2 *cps* gene expression at the transcriptional level. J Bacteriol. 2003;185:788–800.

41. Lam MMC, et al. Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICE*Kp* in *Klebsiella pneumoniae* populations. Microb Genom. 2018;4(9):e000196.

42. Wang R, et al. Increase in antioxidant capacity associated with the successful subclone of hypervirulent carbapenem-resistant *Klebsiella pneumoniae* ST11-KL64. Nat Commun. 2024;15:67.

43. Wang S, et al. Global evolutionary dynamics of virulence genes in ST11-KL47 carbapenem-resistant *Klebsiella pneumoniae*. Int J Antimicrob Agents. 2024;64: 107245.

44. Sun, Z. *et al.* The pivotal role of IncFIB(Mar) plasmid in the emergence and spread of hypervirulent carbapenem-resistant *Klebsiella pneumoniae*. Sci. Adv. 2025;11:eado9097.

45. Xuemei Y, Xiaoxuan L, Wai-Chi CE, Rong Z, Sheng C. Functional characterization of plasmid-borne rmpADC homologues in *Klebsiella pneumoniae*. Microbiol Spectr. 2023;11:e03081–e3122.

46. Le MN, et al. Genomic epidemiology and temperature dependency of hypermucoviscous *Klebsiella pneumoniae* in Japan. Microb Genomics. 2022;8(5):mgen000827.

47. Holt KE, et al. Kleborate. https://github.com/klebgenomics/Kleborate.

Publisher's Note