Evaluating the Genome and Resistome of Extensively Drug-Resistant *Klebsiella pneumoniae* using Native DNA and RNA Nanopore Sequencing

Miranda E. Pitt[1,*], Son H. Nguyen[1], Tânia P.S. Duarte[1], Mark A.T. Blaskovich[1], Matthew A. Cooper[1], Lachlan J.M. Coin[1,*]

[1]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, 4072, Australia

* To whom correspondence should be addressed. Email: miranda.pitt@imb.uq.edu.au, l.coin@imb.uq.edu.au

ABSTRACT

Klebsiella pneumoniae frequently harbour multidrug resistance and current methodologies are struggling to rapidly discern feasible antibiotics to treat these infections. While rapid DNA sequencing has been proposed for prediction of resistance profile; the role of rapid RNA sequencing has yet to be fully explored. The MinION sequencer can sequence native DNA and RNA in real-time, providing an opportunity to contrast the utility of DNA and RNA for prediction of drug susceptibility. This study interrogated the genome and transcriptome of four extensively drug-resistant (XDR) *K. pneumoniae* clinical isolates. The majority of acquired resistance (≥75%) resided on plasmids including several megaplasmids (≥100 kbp). DNA sequencing identified most resistance genes (≥70%) within 2 hours of sequencing. Direct RNA sequencing (with a ~6x slower pore translocation) was able to identify ≥35% of resistance genes, including aminoglycoside, β-lactam, trimethoprim and sulphonamide and also quinolone, rifampicin, fosfomycin and phenicol in some isolates, within 10 hours of sequencing. Polymyxin-resistant isolates showed a heightened transcription of *phoPQ* (≥2-fold) and the *pmrHFIJKLM* operon (≥8-fold). Expression levels estimated from direct RNA sequencing displayed strong correlation (Pearson: 0.86) compared to qRT-PCR across 11 resistance genes. Overall, MinION sequencing rapidly detected the XDR *K. pneumoniae* resistome and direct RNA sequencing revealed differential expression of these genes.

INTRODUCTION

Klebsiella pneumoniae is one of the leading causes of nosocomial infections, with reports of mortality rates as high as 50% (1-5). This opportunistic pathogen frequently exhibits multidrug resistance which severely limits treatment options (6). A high abundance of resistance is commonly encoded on plasmids, accounting for the rapid global dissemination of resistance (1,6). Common therapeutic options for multidrug-resistant infections include carbapenems, fosfomycin, tigecycline and polymyxins (7). However, resistance is also rapidly developing against these antibiotics (6). Alarmingly, pandrug-resistant (PDR) *K. pneumoniae* have emerged which are resistant to all commercially available antibiotics (8,9).

One of the major contributors to the advent of antibiotic resistance is the inability for current detection methodologies to readily and accurately assess bacterial infections in particular, the resistance profile (10). This has resulted in the unnecessary use of antibiotics for viral infections and ineffective antibiotics being administered for resistant infections. Rapid sequencing has been proposed as a way to determine pandrug resistance profiles, including approaches which utilise high accuracy short reads, as well as those which exploit real-time single-molecule sequencing such as Oxford Nanopore Technologies (ONT). The ONT MinION platform is a portable single-molecule sequencer which can sequence long fragments of DNA and stream the sequence data for further data processing in real-time, detecting the presence of bacterial species and acquired resistance genes (11-15). Moreover, the long reads coupled with the ability to multiplex samples has immensely aided with the assembly of bacterial genomes (16-18). This capability allows for the rapid determination of whether resistance is residing on the chromosome or plasmid/s. Of particular interest are high levels of resistance encoded on plasmids, as these genes can rapidly be transferred throughout the bacterial population via horizontal gene transfer.

ONT has recently released a direct RNA sequencing capability, which sequences native transcripts. Other sequencing technologies rely on fragmentation, cDNA conversion and PCR steps which create experimental bias and hinder the accuracy of determining gene expression (19,20). The ability for MinION sequencing to read long fragments enables full length transcripts to be investigated. To date, only a few direct RNA sequencing publications exist which include eukaryote transcriptomes, primarily yeast (*Saccharomyces cerevisiae* (19,21)) and recently, human (BioRxiv: https://doi.org/10.1101/459529). This sequencing has additionally been implemented in viral transcriptomics (22, BioRxiv: https://doi.org/10.1101/300384, BioRxiv: https://doi.org/10.1101/373522). Only one prior study by Smith AM *et al.* has applied this sequencing to bacterial 16S ribosomal RNA (rRNA) to detect epigenetic modifications (BioRxiv: https://doi.org/10.1101/132274). Bacterial transcription differs significantly from eukaryotes in that transcription and translation occur simultaneously. As a result, bacterial mRNA transcripts lack poly(A) tails and alternative splicing (23). The poly(A) tail is critical for the library preparation for ONT sequencing thus, we have established a methodology for adding this component onto transcripts.

In this study, we applied MinION sequencing to interrogate both the genome and the transcriptome (via direct RNA sequencing) for XDR *K. pneumoniae* clinical isolates. Of interest was to compare the potential for RNA sequencing to provide a better correlation to the resistance phenotype than DNA sequencing. These isolates have previously undergone 'traditional' whole genome sequencing (Illumina) and antimicrobial susceptibility testing (24). Three strains were selected from this cohort which exhibited resistance to all 24 classes or combinations of antibiotics

tested, a high abundance of antibiotic resistance genes (≥26) and differing lineages (ST11 (16_GR_13), ST147 (1_GR_13) and ST258 (2_GR_12)). Additionally, these isolates harbour polymyxin resistance which is facilitated by a disruption in or upstream of *mgrB*. MgrB is the negative regulator of PhoPQ and mutation results in the up-regulation of *pmrC* and the *pmrHFIJKLM* operon (25-27). This enables the addition of phosphoethanolamine and/ or 4-amino-4-deoxy-L-arabinose (Ara4N) onto the basal component of lipopolysaccharide, lipid A. These modifications perturb the key electrostatic interaction between lipid A and polymyxins that is critical for their activity (28,29). These pathways associated with polymyxin resistance were further explored using direct RNA sequencing. An additional polymyxin-susceptible XDR isolate (ST258; 20_GR_12) was selected to determine the differential expression associated with polymyxin resistance. This research aimed to assemble these genomes, discern the differential expression of resistance genes and ascertain the time required for detection. Furthermore, we sought to compare DNA and RNA sequencing as modalities for the rapid identification of acquired antibiotic resistance.

MATERIAL AND METHODS

Bacterial strains and growth conditions

Clinically acquired XDR *K. pneumoniae* strains were sourced through the Hygeia General Hospital, Athens, Greece (24). Antimicrobial susceptibility assays (Supplementary Table S1), sequence typing and detection of acquired resistance genes for these isolates have previously been determined (24). Strains were stored at -80°C in 20% (v/v) glycerol and the same stock was used as per the prior study (24). When required for extractions, glycerol stocks were grown on lysogeny broth (LB) agar plates and 6 morphologically similar colonies were selected for inoculation. The inoculum was grown in LB overnight at 37°C shaking at 220 rpm. This overnight inoculum was used for both DNA and RNA extractions.

DNA extraction and high molecular weight DNA isolation

DNA was extracted from 10 ml of overnight culture using the DNeasy Blood and Tissue Kit (Qiagen) according to manufacturer's guidelines, with the addition of an enzymatic lysis buffer pre-treatment (60 mg/ml lysozyme). High molecular weight (HMW) DNA from the prior extraction was selected using the MagAttract HMW DNA Kit (Qiagen) as per manufacturer's instructions. Subtle changes included a further proteinase K treatment on the DNA extracts at 56°C for 10 min followed by supplementation of RNase A (1 mg) for 15 min at room temperature. Several attempts at direct DNA extraction from bacterial cells were undertaken using the MagAttract HMW DNA kit, however, were unsuccessful with these isolates. Due to several issues with potential carbohydrate contamination (260/230 ratio: ≤0.3), 2_GR_12 was also purified with the Monarch® PCR & DNA Cleanup Kit (New England BioLabs) using the protocol to isolate fragments >2000 bp. DNA and RNA contamination was quantitated using Qubit®2.0 (Thermo Fisher Scientific) and purity determined with a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). DNA fragment sizes were determined using the Genomic DNA ScreenTape & Reagents (Agilent) and sizes from 200 to >60000 bp were analyzed on a 4200 TapeStation System (Agilent) (Supplementary Figure S1).

99 **RNA extraction, mRNA enrichment and poly(A) ligation**

100 The overnight culture was sub-cultured in 10 ml of cation-adjusted Muller Hinton Broth (caMHB) to reflect conditions
101 used for minimum inhibitory concentration (MIC) assays. Cultures were grown to mid-log phase (OD_{600} = 0.5-0.6).
102 RNA was extracted via the PureLink™ RNA Mini Kit (Thermo Fisher Scientific) as per manufacturer's protocols
103 which included using Homogenizer columns (Thermo Fisher Scientific). To remove DNA contamination, the TURBO
104 DNA-free™ kit was implemented. A minor adjustment was an increased concentration of TURBO DNase (4 U)
105 incubated at 37°C for 30 min. The RNeasy Mini Kit (Qiagen) clean up protocol was additionally used to purify and
106 concentrate RNA samples. Ribosomal RNA was depleted via the MICROB*Express*™ Bacterial mRNA Enrichment
107 Kit (Thermo Fisher Scientific). Minor protocol changes included adding ≥2 µg of DNA depleted RNA and the
108 enriched mRNA was precipitated for 3 h at -20°C. Poly(A) ligation was performed using the Poly(A) Polymerase
109 Tailing Kit (Astral Scientific) as per the manufacturer's alternative protocol (4 U input of Poly(A) Polymerase). The
110 input RNA concentration was ≥800 ng and RNA samples were incubated at 37°C for 1 h. Poly(A) ligated RNA was
111 purified using Agencourt AmpureXP (Beckman Coulter Australia) beads (1:1 ratio). RNA and DNA contamination
112 was quantitated using the Qubit®2.0 (ThermoFisher Scientific) and purity determined with a NanoDrop 1000
113 Spectrophotometer (Thermo Fisher Scientific). RNA fragment size was checked using an Agilent RNA 6000 Pico kit
114 and run on a 2100 Bioanalyzer (Agilent Technologies) for the initial RNA extract, post ribosomal RNA depletion and
115 after poly(A) ligation (Supplementary Figure S2).

116 **RNA extraction, mRNA enrichment and poly(A) ligation**

117 RNA libraries (≥600 ng poly(A) ligated RNA) were prepared using the Direct RNA Sequencing kit (SQK-RNA001).
118 The Rapid Barcoding Sequencing kit (SQK-RBK001) was used for HMW DNA samples (1_GR_13, 16_GR_13,
119 20_GR_12; 300 ng input each). Isolate 2_GR_12 (300 ng input) was prepared separately using the Rapid Sequencing
120 Kit (SQK-RAD003). Libraries were sequenced with MinION R9.4 flowcells and the raw data (fast5 files) were base-
121 called using Albacore 2.1.1. RNA reads were additionally base-called with Albacore 2.2.7.

122 **Real-time resistome detection emulation**

123 The real-time emulation was performed post sequencing and the time required to detect antibiotic resistance was
124 determined as previously described (14). Briefly, this pipeline aligns Albacore base-called reads via BWA-MEM
125 (ArXiv: https://arxiv.org/abs/1303.3997) to an antibiotic resistance gene database. Antibiotic resistance genes were
126 obtained from the ResFinder 3.0 database (30). This dataset comprises of 2131 genes which were clustered based on
127 90% identity to form 611 groups or gene families. The detection of false positives is reduced using the multiple
128 sequence alignment software kalign2 (31), a probabilistic Finite State Machine (32) and once the alignment score
129 reached a threshold, the resistance gene was reported.

130

131

132 **Assembly of genomes**

133 To assemble genomes with both Illumina and ONT reads, SPAdes v3.10.1 (33) were utilised. Hybrid assemblers
134 included npScarf (34) and Unicycler v0.3.1 (35). Assemblers using only ONT reads included Canu v1.5 (excluding
135 reads <500bp) (36) and the combination of Minimap2 v2.1-r311 and Miniasm v0.2-r168-dirty; Racon (git commit
136 834442) were used in both cases to polish the assemblies (37,38). Consensus sequences were determined using Mauve
137 (snapshot_2015-02-13) to construct the final assembly (39). The output from each assembly software is reported in
138 Supplementary Table S2. Genomes were annotated using the Rapid Annotation using Subsystem Technology (RAST)
139 which also provided a list of virulence genes (40). The location of acquired antibiotic resistance genes were determined
140 using ResFinder 3.0 (30) and plasmids were identified via PlasmidFinder 1.3 (41). To discern if plasmid sequences
141 have previously been reported, contigs underwent a BLASTn analysis against the National Center for Biotechnology
142 Information (NCBI) database (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

143 **RNA alignment and expression profiling**

144 Base-called RNA reads were converted to DNA (uracil bases changed to thymine) and aligned using BWA-MEM
145 (parameters: -k 11 -W20 -r10 -A1 -B1 -O1 -E1 -L0 –Y) to the updated genome assemblies. Due to the lack of introns
146 and full length transcripts being obtained, BEDTools coverage (42) was used to ascertain the relative expression of
147 resistance genes. This was normalized to the number of counts obtained for the housekeeping gene, *rpsL* (43), to
148 compare against qRT-PCR results. Read alignments were further visualised using Integrative genomics viewer (IGV)
149 2.3.59 (44).

150 **Whole transcriptome differential gene expression**

151 To identify genes which were differentially expressed between a pair of samples (x and y), we used a beta-binomial
152 distribution to calculate the probability of observing less than or equal to x_g reads mapping to gene g in sample x,
153 conditional on the total number of reads mapping to all genes (sum_g(x_g)), the number of reads in sample y mapping
154 to gene g (y_g) as well as the total number of reads mapping to all genes in sample y (sum_g (y_g). This was calculated
155 in R using the pbetabinom.ab function in the VGAM package, with q = x_g, size = sum_g'(x_g'), alpha = y_g +1; beta
156 = sum_g'(y_g') - y(g) +1. Genes for which this probability was less than a predefined threshold were deemed to be
157 significantly under expressed in sample x given sample y. A similar statistic was used to check for over-expression.

158 **Quantitative real-time reverse transcriptase PCR (qRT-PCR)**

159 First strand synthesis to generate cDNA (1 µg total DNase-depleted RNA) was performed using SuperScript III
160 (Thermo Fisher Scientific) which was also used for MinION direct RNA sequencing library preparations. Primers
161 used are displayed in Supplementary Table S4. Samples were prepared in triplicate via the SYBR Select Master Mix
162 (Thermo Fisher Scientific) and expression detected using a ViiA 7 Real-time PCR system (Thermo Fisher Scientific).
163 Cycling conditions include: Hold 50°C (2 min), 95°C (2 min) followed by 50 cycles of: 95°C (15 sec), 55°C (1 min).
164 A melt curve was included to determine the specificity of the amplification and a no template control to detect

contamination or primer dimers. Results were analysed with QuantStudioTM Real-Time PCR Software, triplicates were averaged, normalised to the housekeeping gene *rpsL* and relative expression determined via the $2^{-\Delta\Delta CT}$ method (45).

Data availability

Whole genome sequencing of the 4 clinical isolates, including the recent assembly, has been deposited under BioProject PRJNA307517 (www.ncbi.nlm.nih.gov/bioproject/PRJNA307517). ONT DNA sequencing data has been deposited on the Sequence Read Archive (www.ncbi.nlm.nih.gov/sra/) under study SRP133040. Accession numbers are as follows: 1_GR_13 (SRR6747887), 2_GR_12 (SRR6747886), 16_GR_13 (SRR6747885) and 20_GR_12 (SRR6747884). ONT direct RNA sequencing data (pass and fail reads) have been deposited on the Sequence Read Archive (www.ncbi.nlm.nih.gov/sra/) under study SRP133040. Accession numbers are as follows: 1_GR_13 (SRR7719054), 2_GR_12 (SRR7719055), 16_GR_13 (SRR7719052) and 20_GR_12 (SRR7719053).

RESULTS

Discerning the location of acquired resistance in the genome

Utilising the capacity for MinION sequencing to read long fragments of DNA, the location of antibiotic resistance genes were clearly resolved (Table 1). All genomes were circular with the exception of 2_GR_12 where 3 plasmids remained linear. This was partly due to difficulties extracting DNA and not retaining long fragments (Supplementary Figure S1). Amongst the four isolates, the chromosome size ranged between 5.1-5.5 Mb which encoded resistance genes *blaSHV-11*, *fosA* and *oqxAB*. The majority of resistance (≥75%) mapped to plasmids.

At least one megaplasmid, defined as a plasmid larger than 100 kbp, was detected in all isolates (Table 1). These commonly harboured the replicon IncA/C2 or InFIB and IncFIIK. The IncA/C2 plasmid was present in all samples except 20_GR_12. This plasmid contained up to 16 resistance genes which conferred resistance towards aminoglycosides, β-lactams, phenicols, rifampicin, sulphonamides, tetracyclines and trimethoprim, with the exception of 16_GR_13. Isolate 16_GR_13 lacked trimethoprim resistance on its IncA/C2 plasmid. The plasmids containing both replicons IncFIB and IncFIIK differed vastly between all four replicates. All contained $IncFIB_{pKpn3}$ and IncFIIK, however, 1_GR_13 differed with $IncFII_{pKP91}$. Additionally, a differing IncFIB replicon was detected on a separate contig in 1_GR_13 (pKPHS1) and 2_GR_12 (pQil). The only instance where another dual replicon was identified was in 1_GR_13 which harboured both IncR and IncN. This plasmid contained aminoglycoside, β-lactam, trimethoprim, macrolide and sulphonamide resistance. 1_GR_13 also contained a 5.5 kb circular contig which was annotated as a phage genome. Various regions of these megaplasmids were unique to these isolates compared to prior sequences deposited on NCBI (Supplementary Table S5).

The ColRNAI plasmid was present in all except 1_GR_13 which encoded aminoglycoside and quinolone resistance (*aac(6')-Ib, aac(6')-Ib-cr*) (Table 1). The ColRNAI plasmid in 2_GR_12 and 20_GR_12 was 13841 bp in size and shared 75% similarity between the two isolates. This plasmid differed in 16_GR_13 which contained no

resistance genes and 35% the size. The same IncX3 plasmid (43380 bp) was apparent in isolates 2_GR_12 and 20_GR_12. Unique to 16_GR_13 was the IncL/ $M_{pOXA-48}$ plasmid containing *blaOXA-48* and the 50979 bp IncN plasmid in 20_GR_12 with resistance against 5 classes (aminoglycoside (*aph(3")-Ib*, *aph(6)-Id*), β-lactam (*blaTEM-1A*), sulphonamide (*sul2*), tetracycline (*tet(A)*), trimethoprim (*dfrA14*)) of antibiotics.

Multiple copies of acquired resistance genes were apparent across plasmids in several isolates. For 1_GR_13, up to three copies were present of genes *aadA24*, *aph(3')-Ia*, *aph(6)-Id, dfrA1, dfrA14, strA* and *sul1* (Table 1). In 2_GR_12, *sul1* and *blaTEM-1A* were duplicated and for 16_GR_13, only *sul1* was represented twice.

Real-time detection emulation of resistance genes via DNA sequencing

The vast majority (≥70%) of resistance genes were detected via DNA sequencing within the first 2 hours (Figure 1, Supplementary Table S3). These genes confer resistance towards aminoglycosides, β-lactams, fosfomycin, macrolides, phenicols, quinolones, rifampicin, sulphonamides, tetracyclines and trimethoprim. 20_GR_12 lacked acquired resistance genes for macrolides, phenicols and rifampicin, however, all other classes were detected within 2 hours. All isolates, except 2_GR_12, were sequenced for 21 hours which was sufficient to obtain the complete genome assembly. Only a few additional genes were detected after the first 10 hours across isolates (Supplementary Table S3). For 2_GR_12, an extended run of 41 hours detected no further genes after 20 hours. Overall, the presence of these resistance genes corresponded to a resistant phenotype towards aminoglycosides, β-lactams, fosfomycin, phenicols, quinolones, sulphonamides (sulfamethoxazole), tetracyclines and trimethoprim (Supplementary Table S1). As macrolides and rifampicin are not routinely used to treat *K. pneumoniae* infections, no breakpoints exist according to CLSI and EUCAST guidelines, hence, MICs were not determined.

Post 2 hours of sequencing, several genes not observed in the final assembly via ResFinder 3.0 were detected (Supplementary Table S3). These were predominantly genes attributed to aminoglycoside, β-lactam, rifampicin and phenicol resistance. Furthermore, resistance genes to additional differing classes were detected including fusidic acid and vancomycin. This was evident in 2_GR_12 (*fusB*) and 16_GR_13 (*fusB*, *vanR*). However, these genes had less than 30 reads and their phred-scale mapping quality scores (MAPQ) were less than 10 (misplaced probability greater than 0.1). Furthermore, the majority of genes not observed in the final assembly nor observed in Illumina data exhibited a MAPQ score of ≤10 which may indicate that a more stringent threshold is required to negate false positives. However if this threshold increases, true positives would not be detected including *aadA1*, *aadA2* and *ARR-2* in 2_GR_12 and *blaOXA-48*, *blaCTX-M-15* and *ARR-2* in 16_GR_13.

Several genes found in the final assembly were not detected in the real-time emulation analysis (Supplementary Table S3). This was mainly observed for aminoglycoside resistance encoding genes. For 1_GR_13, this included *aadA1, ant(2")-Ia*, *aph(6)-Id* and *aadA24.* Similarly, 2_GR_12 and 20_GR_12 lacked *aph(3")-Ib* and *aph(6)-Id.* 2_GR_12 additionally had the absence of *ant(2")-Ia.* Detection of *ant(2")-Ia*, *aph(3")-Ib*, *aph(6)-Id* was not present in 16_GR_13. 16_GR_13 further lacked *catB4* (phenicol) and *tet(A)* (tetracycline). Various phenicol resistance genes were reported in the real-time emulation however, the incorrect gene was identified which may represent sequencing errors accumulated over time and high similarity to other phenicol resistance genes. The tetracycline resistance gene, *tet(A)*, was interestingly not reported in this emulation with 190 reads and the majority of reads exhibiting a high

mapping confidence (MAPQ = 60, equivalent to an error probability of $1x10^{-6}$). This gene was only detected after 10 hours for 1_GR_13 and 2_GR_12 and this result may be influenced by the presence of only 1 copy of *tet(A)* encoded on a low copy number megaplasmid (between 1 to 1.5, see Table 1).

Direct RNA sequencing resistance detection

The time required to detect resistance was further interrogated using RNA sequencing. Rapid detection was apparent for several resistance genes via direct RNA sequencing (Figure 1). This was evident for genes conferring resistance to aminoglycosides, β-lactams, sulphonamides and trimethoprim for all four isolates. Resistance towards these antibiotics was commonly detected within 6 hours. In some instances, quinolone, rifampicin, fosfomycin and phenicol resistance was detected. This result remained similar when all reads or passed reads alone were analysed. The most significant difference when analysing all reads was the detection of *fosA* in 1_GR_13 and *ARR-2* and *fosA* in 2_GR_12. Consistently absent from this analysis were genes attributed to macrolide (*mph(A)*) and tetracycline (*tet(A)*, *tet(G)*) resistance, however, isolates exhibited high levels of resistance to tetracycline (>64 µg/ml) (Supplementary Table S1). This may indicate that isolates require antibiotic exposure to enable transcription of these genes. Commonly no new genes were detected after 12 hours of sequencing with the exception of *fosA* in 2_GR_12. Although *fosA* was detected when including the failed reads, a low MAPQ score (≤10) was apparent. Similar to the DNA real-time detection, several genes not found in the final assembly were identified (Supplementary Table S3). With the exception of 20_GR_12, this included *aadB* and *strB* for all isolates. Additional genes detected included *ARR-7* in 1_GR_13, *strA* in 2_GR_12 and for 16_GR_13, *blaCTX-M-64*, *blaOXA-436* and *strA*. Similar genes or gene families were identified when comparing DNA and direct RNA sequencing. Overall, genes were detected more readily via DNA rather than RNA sequencing, possibly due to a lack of RNA expression in the absence of the antibiotic to which resistance is encoded. There were only a few instances where RNA sequencing detected resistance more quickly than DNA sequencing which included *aac(3')-IIa* in 16_GR_13 and *sul2* and *catA1* in 2_GR_12. Similar results were apparent when investigating data yield rather than time (Supplementary Figure S4).

Levels of expression of resistance genes

RNA sequencing accumulated over approximately 40 hours yielded between 0.9 and 1.7 million reads for these isolates (Supplementary Figure S3). However, only 10 to 14% of these reads successfully passed base-calling which was similar when using either Albacore 2.1.1 or 2.2.7. The low proportion of base-called reads reflects the fact that base-calling algorithms have not yet been optimised for direct RNA sequencing, and even less so for bacterial RNA sequencing. When aligning passed reads to the final assembly, ≥98% of reads were mappable, however, ≤40% of these had a MAPQ score ≥10. When all reads were aligned, ≤22% mapped to the genome and ≤5% exhibited a MAPQ score ≥10. A proportion of these reads were found to map to rRNA including 1_GR_13 (18%), 2_GR_12 (37%), 16_GR_13 (24%) and 20_GR_12 (23%). Overall, at least 58% of genes (with at least 1 read mapping to the gene) were identified to be expressed across isolates (1_GR_13 (68%), 2_GR_12 (58%), 16_GR_13 (75%) and 20_GR_12 (69%).

Amongst the four isolates, levels of expression for resistance genes on the chromosome (*blaSHV-11*, *fosA* and *oqxAB*) were low (≤122 counts per million (cpm) mapped reads) (Figure 2). The remaining resistance genes were located on plasmids. Resistance genes exhibiting high levels of expression (300 cpm) were apparent in 1_GR_13 (*blaTEM-1B*, *blaVIM-27*, *sul1*, *aph(3')-Ia*), 2_GR_12 (*aac(6')-Ib*, *catA1*, *blaKPC-2*), 16_GR_13 (*aac(6')Ib-cr*, *aac(3)-IIa*, *blaCTX-M-15*, *blaTEM-1B*, *blaOXA-48*) and 20_GR_12 (*blaKPC-2*, *aac(6')Ib*). Counts for *aac(6')-1b* and *aac(6')-1b-cr* in 2_GR_12 and 20_GR_12 were grouped. The gene *aac(6')-1b-cr* is a shortened version of *aac(6')-1b* and both were identified in the same genome position, hence, only *aac(6')-1b* is displayed in Figure 2. Relative expression did not differ significantly when analysing passed reads alone or all reads. All highly expressed genes were detected within 6 hours as per the real-time detection emulation. As anticipated, low levels of expression were observed for fosfomycin (*fosA*), tetracycline (*tet(A)*, *tet(B)*) and macrolide (*mph(A)*) resistance.

A subset of 11 resistance genes which represent resistance across various classes of antibiotics were investigated to validate differential gene expression in these RNA extractions via qRT-PCR. These included resistance towards aminoglycosides (*aac(6')Ib*, *strA*), β-lactams (*blaKPC-2*, *blaOXA-10*, *blaTEM-1*), phenicols (*cmlA1*), trimethoprim (*dfrA14*), fosfomycin (*fosA*), quinolone (*oqxA*), sulphonamides (*sul2*) and tetracyclines (*tet(A)*). A similar trend was observed between direct RNA sequencing and qRT-PCR results (Spearman's rank correlation coefficient: 0.83; Pearson correlation: 0.86) (Figure 3). The highest expression of a resistance gene was observed for *blaKPC-2* although only one copy was present in a lower copy number plasmid in 2_GR_12 and 20_GR_12 (Figure 2, Figure 3 and Table 1). Additionally, low levels of expression for *fosA* and *tet(A)* were apparent despite exhibiting resistance towards fosfomycin and tetracycline (Figure 3, Supplementary Table S1). Direct RNA sequencing was unable to detect low levels of expression whilst qRT-PCR could detect these genes (Figure 3).

Across the transcriptome, antibiotic resistance genes were identified to harbour high differential expression between isolates (Figure 4). Virulence genes were comparable across these strains similar to all remaining or background genes. The top differentially expressed genes were determined (Supplementary Figure S5) and several were associated with polymyxin resistance pathways. Heightened expression was seen in polymyxin-resistant isolates 1_GR_13, 2_GR_12, 16_GR_13 in comparison to the single susceptible isolate in particular, genes associated with Ara4N synthesis. These genes include 4-deoxy-4-formamido-L-arabinose-phosphoundecaprenol deformylase (ArnD), UDP-4-amino-4-deoxy-L-arabinose formyltransferase and UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferase.

Transcriptional biomarkers for polymyxin resistance

Three of the isolates harboured resistance towards polymyxins via disruptions in *mgrB* which included 1_GR_13, 2_GR_12 and 16_GR_13. 1_GR_13. These isolates have an insertion sequence (IS) element, IS*Kpn26*-like, at nucleotide position 75 in the same orientation as *mgrB*. 2_GR_12 also contained an insertion at the same position, however, in the opposite orientation and additional mutations in *phoP* (A95S) and *phoQ* (N253T). 16_GR_13 possessed an IS element, IS*1R*-like, 19 bp upstream of *mgrB*. Direct RNA sequencing revealed only low level expression of *mgrB* in isolates (1_GR_13 (78.4 cpm), 2_GR_12 (16.3 cpm), 16_GR_13 (0 cpm), 20_GR_12 (2.3 cpm)). The expression levels of various genes associated with this pathway were verified via qRT-PCR which include

genes *phoP*, *phoQ*, *pmrA*, *pmrB*, *pmrC*, *pmrD*, *pmrE*, *pmrH* and *pmrK* (Figure 5). Direct RNA sequencing revealed a slight increase in transcription of *phoPQ* (≥2-fold) relative to the expression in 20_GR_12. A ≥13-fold increase in expression was observed for *pmrH* and ≥8-fold elevation for *pmrK*. Similar trends for expression were also reported using qRT-PCR (Figure 5B).

DISCUSSION

XDR *K. pneumoniae* pose as a major threat to modern medicine with rapid diagnostics critical to discern appropriate treatment options (1,6). The MinION sequencing technology employed in this study has great potential to detect antibiotic resistance in a timely manner, as shown with four XDR *K. pneumoniae* isolates. This method was able to resolve both the assembly of plasmids harbouring high levels of resistance (through DNA sequencing) and the expression from the resistome in the absence of antibiotic treatment (through RNA sequencing).

The ability for ONT to sequence long fragments of DNA has significantly aided the assembly of bacterial genomes and plasmids (16-18). In this study, multiple megaplasmids (≥100 kbp) were identified which were previously unresolved via Illumina sequencing (24). These harboured replicons IncA/C2 or a dual replicon, IncFIIK and IncFIB. The IncA/C, IncF and IncN plasmids have been commonly associated with multidrug resistance (46). Although several plasmids in this study revealed similarity to previously reported isolates via NCBI, various sequences deviated. In particular, the IncA/C2 plasmid exhibited multiple regions unique to these isolates. Several IncA/C2 megaplasmids have been previously described which harbour various resistance genes, however, the extent of resistance in our study has yet to be unveiled (47,48). Prior studies have shown the IncFIIK and IncFIB replicons to localise on the same plasmid and also megaplasmids with multidrug resistance (6). The $IncFIB_{pQil}$ plasmid in this study contained various β-lactam resistance genes (*blaKPC-2*, *blaOXA-9*, *blaTEM-1A*) which has been identified previously (49). Similarly, *blaOXA-48* segregated with the IncL/M replicon (50,51), however, deviations in this plasmid were identified.

The real-time analysis capability entailed in MinION sequencing has the potential to rapidly determine the antibiotic resistance profile. Previously, this device has been utilised to rapidly assemble bacterial genomes, discern species and detect antibiotic resistance (12-15). This study investigated the potential time required to discern resistance via a real-time emulation as previously described (17). The majority (≥70%) of resistance genes were detected via DNA sequencing within 2 hours. However, several genes that were not identified in the final assembly were apparent after 2 hours. This may be attributed to the high similarity (≥80%) amongst various genes, in particular, those associated with aminoglycoside, β-lactam, rifampicin and phenicol resistance. Furthermore, the error rate associated with ONT sequencing and the accumulation of these errors over time may result in the false detection of these genes. Nanopore DNA sequencing currently has an accuracy ranging from 80 to 90% which limits its ability to detect mutations (17). Various resistance genes only differ by a few nucleotides which significantly impacts the resistance phenotype and the antibiotics which can be utilised to treat these infections. Furthermore, direct RNA sequencing has an average error rate of 12% (21). Hence, it is essential for the technology to increase its accuracy in order to correctly and rapidly diagnose antibiotic resistance.

Investigating the transcriptome of these isolates can potentially elucidate the correlation between genotype and the subsequent resistant phenotype. One of the advantages of RNA sequencing is that it can identify conditions in which a resistance gene is present but not expressed, potentially resulting in a susceptible phenotype. However, if expression is only induced in the presence of an antibiotic, the absence of RNA transcripts may falsely suggest susceptibility. Direct RNA sequencing revealed high levels of transcription from genes associated with aminoglycoside, β-lactam, sulphonamide and trimethoprim resistance within 6 hours. The detection of quinolone, rifampicin, and phenicol resistance correlated to the levels of transcription within samples. All isolates exhibited low levels of expression for fosfomycin, macrolide and tetracycline resistance, despite exhibiting phenotypic resistance to fosfomycin and tetracycline. Whether this transcription is due to prior exposure to these antibiotics in the clinic and the longevity of this expression post exposure warrants further investigation. The changes in transcription levels in response to antibiotic exposure also need to be assessed in future experiments. Furthermore, the time required to detect resistance may be hindered by the slower translocation speed associated with direct RNA sequencing (70 bases/ second) compared to DNA sequencing (450 bases/ second). Furthermore, insufficient rRNA depletion and low base-calling of data could be impacting the detection of this low level expression.

Another variable to consider when evaluating differential expression is the operon or promoter which can further be explored via cloning. In particular, the highest levels of expression were observed for *blaKPC-2* in 2_GR_12 and 20_GR_12. Alterations in the promoter region have previously been reported to influence high levels of expression (52). Furthermore, despite low levels of transcription for fosfomycin (*fosA*) and tetracycline (*tet(A)*, *tet(G)*), phenotypically these isolates consistently retain resistance (24). FosA, an enzyme involved in the degradation of fosfomycin, is commonly encoded chromosomally in *K. pneumoniae* and a combination of expression and enzymatic activity contributes to resistance (53). Genes *tet(A)* and *tet(G)* encode efflux pumps which, in the absence of tetracycline, are lowly expressed (54). Detecting inducible resistance such as tetracycline resistance highlights one of the advantages of investigating the transcriptome. Additionally, copy number of plasmids can further alter the levels of expression detected for these resistance genes.

In this study we also investigated pathways attributed to polymyxin resistance. Three of these strains exhibited an IS element upstream of within *mgrB*, the negative regulator of PhoPQ (25,26). Elevated expression was apparent for *phoPQ* and also the *pmrHFIJKLM* operon in our polymyxin-resistant isolates harbouring a disruption in *mgrB*. This has previously been witnessed for other *K. pneumoniae* isolates harbouring *mgrB* disruptions and is a potential transcriptional marker for polymyxin resistance (27,43,55,56). However, this study is limited to four isolates and one mechanism associated with polymyxin resistance. Other pathways have previously been identified including the role of other TCSs such as CrrAB (57). The ability to use relative expression of key genes to detect polymyxin resistance requires further validation, including an increased sample size of resistant and non-resistant isolates. Furthermore, additional functional experiments such as complementation assays would be required in order to validate the contribution of a certain mutation to the transcriptome and subsequent resistance.

This study has utilised MinION sequencing to assemble four XDR *K. pneumoniae* genomes and has revealed several unique plasmids harbouring multidrug resistance. The vast majority of this resistance was detectable within 2 hours of sequencing, though a number of resistance genes were identified that were not present in the final assembly.

Exploiting this analysis in real-time would allow for a rapid diagnostic, however, the presence of a resistance gene does not necessarily indicate resistance is conferred and requires additional phenotypic characterisation. This research also established a methodology and analysis for bacterial direct RNA sequencing. The differential expression of resistance genes were successfully detected via this technology and can be exploited for bacterial transcriptomics. Once base-calling algorithms have been optimised, this could allow for a whole transcriptome interrogation of full length transcripts regulated by operons, where more than one gene is co-expressed in a transcript, and the evaluation of the poorly characterised epitranscriptome. This research established a methodology and analysis for bacterial direct RNA sequencing. The differential expression of resistance genes were successfully detected via this technology and can be exploited for bacterial transcriptomics. Overall, this study has begun to unravel the association between genotype, transcription and subsequent resistant phenotype in these XDR *K. pneumoniae* clinical isolates, establishing the groundwork for developing a diagnostic that can rapidly determine bacterial resistance profiles.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr Ilias Karaiskos and Dr Helen Giamarellou for providing the bacterial strains in this study. We also acknowledge Dr Evangelos Bellos for his guidance on the RNA sequencing analysis and Dr Devika Ganesamoorthy for the initial advice on the direct RNA sequencing library preparation.

FUNDING

LC is an NHMRC career development Fellow APP1103384). MAC is an NHMRC Principal Research Fellow (APP1059354) and currently holds a fractional Professorial Research Fellow appointment at the University of Queensland with his remaining time as CEO of Inflazome Ltd. a company headquartered in Dublin, Ireland that is developing drugs to address clinical unmet needs in inflammatory disease by targeting the inflammasome. MEP is an Australian Postgraduate Award scholar. MATB is supported in part by a Wellcome Trust Strategic Award 104797/Z/14/Z. This work was supported by the Institute for Molecular Bioscience Centre for Superbug Solutions (610246).

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

REFERENCES

1. Martin, R.M. and Bachman, M.A. (2018) Colonization, Infection, and the Accessory Genome of *Klebsiella pneumoniae*. *Front Cell Infect. Microbiol.*, **8**, 4.
2. Magill, S.S., Edwards, J.R., Bamberg, W., Beldavs, Z.G., Dumyati, G., Kainer, M.A., Lynfield, R., Maloney, M., McAllister-Hollod, L., Nadle, J. *et al.* (2014) Multistate point-prevalence survey of health care-associated infections. *N. Engl. J. Med.*, **370**, 1198-1208.
3. Kalanuria, A.A., Ziai, W. and Mirski, M. (2014) Ventilator-associated pneumonia in the ICU. *Crit. Care*, **18**, 208.
4. Talha, K.A., Hasan, Z., Selina, F. and Palash, M.I. (2009) Organisms associated with ventilator associated pneumonia in intensive care unit. *Mymensingh Med. J.*, **18**, S93-97.
5. Podschun, R. and Ullmann, U. (1998) *Klebsiella spp.* as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin. Microbiol. Rev.*, **11**, 589-603.
6. Navon-Venezia, S., Kondratyeva, K. and Carattoli, A. (2017) *Klebsiella pneumoniae*: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol. Rev.*, **41**, 252-275.
7. Karaiskos, I. and Giamarellou, H. (2014) Multidrug-resistant and extensively drug-resistant Gram-negative pathogens: current and emerging therapeutic approaches. *Expert Opin. Pharmacother.*, **15**, 1351-1370.
8. Chen, L., Todd, R., Kiehlbauch, J., Walters, M. and Kallen, A. (2017) Notes from the Field: Pan-Resistant New Delhi Metallo-Beta-Lactamase-Producing *Klebsiella pneumoniae* - Washoe County, Nevada, 2016. *MMWR Morb. Mortal Wkly Rep.*, **66**, 33.
9. Zowawi, H.M., Forde, B.M., Alfaresi, M., Alzarouni, A., Farahat, Y., Chong, T.M., Yin, W.F., Chan, K.G., Li, J., Schembri, M.A. *et al.* (2015) Stepwise evolution of pandrug-resistance in *Klebsiella pneumoniae*. *Sci. Rep.*, **5**, 15082.
10. Sommer, M.O.A., Munck, C., Toft-Kehler, R.V. and Andersson, D.I. (2017) Prediction of antibiotic resistance: time for a new preclinical paradigm? *Nat. Rev. Microbiol.*, **15**, 689-696.
11. Gardy, J.L. and Loman, N.J. (2018) Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat. Rev. Genet.*, **19**, 9-20.
12. Lemon, J.K., Khil, P.P., Frank, K.M. and Dekker, J.P. (2017) Rapid Nanopore Sequencing of Plasmids and Resistance Gene Detection in Clinical Isolates. *J. Clin. Microbiol.*, **55**, 3530-3543.
13. Votintseva, A.A., Bradley, P., Pankhurst, L., Del Ojo Elias, C., Loose, M., Nilgiriwala, K., Chatterjee, A., Smith, E.G., Sanderson, N., Walker, T.M. *et al.* (2017) Same-Day Diagnostic and Surveillance Data for Tuberculosis via Whole-Genome Sequencing of Direct Respiratory Samples. *J. Clin. Microbiol.*, **55**, 1285-1298.
14. Cao, M.D., Ganesamoorthy, D., Elliott, A.G., Zhang, H., Cooper, M.A. and Coin, L.J. (2016) Streaming algorithms for identification of pathogens and antibiotic resistance potential from real-time MinION(TM) sequencing. *Gigascience*, **5**, 32.

15. Quick, J., Ashton, P., Calus, S., Chatt, C., Gossain, S., Hawker, J., Nair, S., Neal, K., Nye, K., Peters, T. *et al.* (2015) Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of *Salmonella*. *Genome Biol.*, **16**, 114.

16. Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. (2017) Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb. Genom.*, **3**, e000132.

17. Li, R., Xie, M., Dong, N., Lin, D., Yang, X., Wong, M.H.Y., Chan, E.W. and Chen, S. (2018) Efficient generation of complete sequences of MDR-encoding plasmids by rapid assembly of MinION barcoding sequencing data. *Gigascience*, **7**, 1-9.

18. George, S., Pankhurst, L., Hubbard, A., Votintseva, A., Stoesser, N., Sheppard, A.E., Mathers, A., Norris, R., Navickaite, I., Eaton, C. *et al.* (2017) Resolving plasmid structures in Enterobacteriaceae using the MinION nanopore sequencer: assessment of MinION and MinION/Illumina hybrid data assembly approaches. *Microb. Genom.*, **3**, e000118.

19. Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A. *et al.* (2018) Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods*, **15**, 201-206.

20. Ozsolak, F. and Milos, P.M. (2011) RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.*, **12**, 87-98.

21. Jenjaroenpun, P., Wongsurawat, T., Pereira, R., Patumcharoenpol, P., Ussery, D.W., Nielsen, J. and Nookaew, I. (2018) Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of *Saccharomyces cerevisiae* CEN.PK113-7D. *Nucleic Acids Res.*, **46**, e38.

22. Moldovan, N., Tombacz, D., Szucs, A., Csabai, Z., Balazs, Z., Kis, E., Molnar, J. and Boldogkoi, Z. (2018) Third-generation Sequencing Reveals Extensive Polycistronism and Transcriptional Overlapping in a *Baculovirus*. *Sci. Rep.*, **8**, 8604.

23. Sorek, R. and Cossart, P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.*, **11**, 9-16.

24. Pitt, M.E., Elliott, A.G., Cao, M.D., Ganesamoorthy, D., Karaiskos, I., Giamarellou, H., Abboud, C.S., Blaskovich, M.A.T., Cooper, M.A. and Coin, L.J.M. (2018) Multifactorial chromosomal variants regulate polymyxin resistance in extensively drug-resistant *Klebsiella pneumoniae*. *Microb. Genom.*, doi: 10.1099/mgen.1090.000158.

25. Cannatelli, A., Giani, T., D'Andrea, M.M., Di Pilato, V., Arena, F., Conte, V., Tryfinopoulou, K., Vatopoulos, A. and Rossolini, G.M. (2014) MgrB inactivation is a common mechanism of colistin resistance in KPC-producing *Klebsiella pneumoniae* of clinical origin. *Antimicrob. Agents Chemother.*, **58**, 5696-5703.

26. Olaitan, A.O., Diene, S.M., Kempf, M., Berrazeg, M., Bakour, S., Gupta, S.K., Thongmalayvong, B., Akkhavong, K., Somphavong, S., Paboriboune, P. *et al.* (2014) Worldwide emergence of colistin resistance in *Klebsiella pneumoniae* from healthy humans and patients in Lao PDR, Thailand, Israel, Nigeria and France owing to inactivation of the PhoP/PhoQ regulator mgrB: an epidemiological and molecular study. *Int. J. Antimicrob. Agents*, **44**, 500-507.

27. Wright, M.S., Suzuki, Y., Jones, M.B., Marshall, S.H., Rudin, S.D., van Duin, D., Kaye, K., Jacobs, M.R., Bonomo, R.A. and Adams, M.D. (2015) Genomic and transcriptomic analyses of colistin-resistant clinical isolates of *Klebsiella pneumoniae* reveal multiple pathways of resistance. *Antimicrob. Agents Chemother.*, **59**, 536-543.

28. Helander, I.M., Kato, Y., Kilpelainen, I., Kostiainen, R., Lindner, B., Nummila, K., Sugiyama, T. and Yokochi, T. (1996) Characterization of lipopolysaccharides of polymyxin-resistant and polymyxin-sensitive *Klebsiella pneumoniae* O3. *Eur. J. Biochem.*, **237**, 272-278.

29. Velkov, T., Deris, Z.Z., Huang, J.X., Azad, M.A., Butler, M., Sivanesan, S., Kaminskas, L.M., Dong, Y.D., Boyd, B., Baker, M.A. *et al.* (2014) Surface changes and polymyxin interactions with a resistant strain of *Klebsiella pneumoniae*. *Innate Immun.*, **20**, 350-363.

30. Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M. and Larsen, M.V. (2012) Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.*, **67**, 2640-2644.

31. Lassmann, T., Frings, O. and Sonnhammer, E.L. (2009) Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.*, **37**, 858-865.

32. Allison, L., Wallace, C.S. and Yee, C.N. (1990) When is a string like a string? In: Artificial Intelligence and Mathematics. *Ft. Lauderdale FL.*

33. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D. *et al.* (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.*, **19**, 455-477.

34. Cao, M.D., Nguyen, S.H., Ganesamoorthy, D., Elliott, A.G., Cooper, M.A. and Coin, L.J. (2017) Scaffolding and completing genome assemblies in real-time with nanopore sequencing. *Nat. Commun.*, **8**, 14515.

35. Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.*, **13**, e1005595.

36. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H. and Phillippy, A.M. (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.*, **27**, 722-736.

37. Li, H. (2016) Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics*, **32**, 2103-2110.

38. Vaser, R., Sovic, I., Nagarajan, N. and Sikic, M. (2017) Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res.*, **27**, 737-746.

39. Darling, A.E., Tritt, A., Eisen, J.A. and Facciotti, M.T. (2011) Mauve assembly metrics. *Bioinformatics*, **27**, 2756-2757.

40. Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, **9**,75.

41. Carattoli, A., Zankari, E., Garcia-Fernandez, A., Voldby Larsen, M., Lund, O., Villa, L., Moller Aarestrup, F. and Hasman, H. (2014) *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.*, **58**, 3895-3903.

42. Quinlan, A.R. (2014) BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinformatics*, **47**, 11.12.11-34.

43. Cannatelli, A., D'Andrea, M.M., Giani, T., Di Pilato, V., Arena, F., Ambretti, S., Gaibani, P. and Rossolini, G.M. (2013) *In vivo* emergence of colistin resistance in *Klebsiella pneumoniae* producing KPC-type carbapenemases mediated by insertional inactivation of the PhoQ/PhoP mgrB regulator. *Antimicrob Agents Chemother*, **57**, 5521-5526.

44. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24-26.

45. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.

46. Carattoli, A. (2009) Resistance plasmid families in Enterobacteriaceae. *Antimicrob. Agents Chemother.*, **53**, 2227-2238.

47. Desmet, S., Nepal, S., van Dijl, J.M., Van Ranst, M., Chlebowicz, M.A., Rossen, J.W., Van Houdt, J.K.J., Maes, P., Lagrou, K. and Bathoorn, E. (2018) Antibiotic Resistance Plasmids Cointegrated into a Megaplasmid Harboring the *blaOXA-427* Carbapenemase Gene. *Antimicrob. Agents Chemother.*, **62**, pii: e01448-17.

48. Papagiannitsis, C.C., Dolejska, M., Izdebski, R., Giakkoupi, P., Skalova, A., Chudejova, K., Dobiasova, H., Vatopoulos, A.C., Derde, L.P., Bonten, M.J. *et al.* (2016) Characterisation of IncA/C2 plasmids carrying an In416-like integron with the *blaVIM-19* gene from *Klebsiella pneumoniae* ST383 of Greek origin. *Int. J. Antimicrob. Agents*, **47**, 158-162.

49. Chen, L., Chavda, K.D., Melano, R.G., Jacobs, M.R., Koll, B., Hong, T., Rojtman, A.D., Levi, M.H., Bonomo, R.A. and Kreiswirth, B.N. (2014) Comparative genomic analysis of KPC-encoding pKpQIL-like plasmids and their distribution in New Jersey and New York Hospitals. *Antimicrob. Agents Chemother.*, **58**, 2871-2877.

50. Poirel, L., Bonnin, R.A. and Nordmann, P. (2012) Genetic features of the widespread plasmid coding for the carbapenemase OXA-48. *Antimicrob. Agents Chemother.*, **56**, 559-562.

51. Potron, A., Poirel, L. and Nordmann, P. (2014) Derepressed transfer properties leading to the efficient spread of the plasmid encoding carbapenemase OXA-48. *Antimicrob. Agents Chemother.*, **58**, 467-471.

52. Cheruvanky, A., Stoesser, N., Sheppard, A.E., Crook, D.W., Hoffman, P.S., Weddle, E., Carroll, J., Sifri, C.D., Chai, W., Barry, K. *et al.* (2017) Enhanced *Klebsiella pneumoniae* Carbapenemase Expression from a Novel Tn4401 Deletion. *Antimicrob. Agents Chemother.*, **61**.

53. Klontz, E.H., Tomich, A.D., Gunther, S., Lemkul, J.A., Deredge, D., Silverstein, Z., Shaw, J.F., McElheny, C., Doi, Y., Wintrode, P.L. *et al.* (2017) Structure and Dynamics of FosA-Mediated Fosfomycin Resistance in *Klebsiella pneumoniae* and *Escherichia coli*. *Antimicrob. Agents Chemother.*, **61**.

54. Saenger, W., Orth, P., Kisker, C., Hillen, W. and Hinrichs, W. (2000) The Tetracycline Repressor-A Paradigm for a Biological Switch. *Angew Chem. Int. Ed. Engl.*, **39**, 2042-2052.

55. Cheng, Y.H., Lin, T.L., Pan, Y.J., Wang, Y.P., Lin, Y.T. and Wang, J.T. (2015) Colistin resistance mechanisms in *Klebsiella pneumoniae* strains from Taiwan. *Antimicrob. Agents Chemother.*, **59**, 2909-2913.

56. Haeili, M., Javani, A., Moradi, J., Jafari, Z., Feizabadi, M.M. and Babaei, E. (2017) MgrB Alterations Mediate Colistin Resistance in *Klebsiella pneumoniae* Isolates from Iran. *Front Microbiol.*, **8**, 2470.

57. Baron, S., Hadjadj, L., Rolain, J.M., Olaitan, A.O. (2017) Molecular mechanisms of polymyxin resistance: knowns and unknowns. *Int. J. Antimicrob. Agents*, **48**,583-591.

Table 1. Final assembly of XDR *K. pneumoniae* isolates and location of antibiotic resistance genes

Isolate	ST	Contig	Length (bp)	Coverage	Contig ID[a]	Resistance Genes[b]
1_GR_13	147	1	**5181675**	1	C	*blaSHV-11*, *fosA*, *oqxA*, *oqxB*
		2	**192771**	1.95	P: IncA/C2	*aadA1*, *ant(2'')-Ia*, *aph(6)-Id*, *ARR-2*, *blaOXA-10*, *blaTEM-1B*, *blaVEB-1*, *cmlA1*, *dfrA14*, *dfrA23*, *rmtB, strA*, *sul1*, *sul2, tet(A)*, *tet(G)*
		3	**168873**	2	P: $IncFIB_{pKpn3}$, $IncFII_{pKP91}$	*aadA24*, *aph(3')-Ia*, *aph(6)-Id*, *dfrA1*, *dfrA14*, *strA*
		4	**108879**	1.53	P: $IncFIB_{pKPHS1}$	-
		5	**55018**	14.10	-	-
		6	**53495**	2.36	P: IncR, IncN	*aadA24*, *aph(3')-Ia*, *aph(6)-Id*, *blaVIM-27*, *dfrA1*, *mph(A)*, *strA, sul1*
2_GR_12	258	1	**5466424**	1	C	*blaSHV-11*, *fosA*, *oqxA*, *oqxB*
		2	197872	1.3	P: $IncFIB_{pKpn3}$, IncFIIK	*aadA2*, *aph(3')-Ia*, *catA1*, *dfrA12*, *mph(A)*, *sul1*
		3	175636	1.49	P: IncA/C2	*aadA1*, *ant(2'')-Ia*, *aph(3'')-Ib*, *aph(6)-Id*, *ARR-2*, *blaOXA-10*, *blaTEM-1A*, *blaVEB-1*, *cmlA1, dfrA14*, *dfrA23*, *rmtB*, *sul1*, *sul2*, *tet(A)*, *tet(G)*
		4	95481	1.61	P: $IncFIB_{pQil}$	*blaKPC-2*, *blaOXA-9*, *blaTEM-1A*
		5	**43380**	1.91	P: IncX3	*blaSHV-12*
		6	**13841**	4	P: ColRNAI	*aac(6')-Ib, aac(6')Ib-cr*
16_GR_13	11	1	**5426917**	1	C	*blaSHV-11*, *fosA*, *oqxA*, *oqxB*
		2	**187670**	0.88	P: $IncFIB_{pKpn3}$; IncFIIK	*aac(3)-IIa*, *aac(6')Ib-cr*, *aadA2*, *aph(3')-Ia, blaCTX-M-15*, *blaOXA-1*, *catB4, dfrA12*, *mph(A)*, *sul1*
		3	**155161**	0.99	P: IncA/ C2	*aadA1*, *ant(2'')-Ia*, *aph(3'')-Ib*, *aph(6)-Id*, *ARR-2*, *blaOXA-10*, *blaTEM-1B*, *blaVEB-1*, *cmlA1*, *rmtB*, *sul1*, *sul2*, *tet(A)*, *tet(G)*
		4	**63589**	1.49	P: IncL/ $M_{pOXA-48}$	*blaOXA-48*
		5	**5234**	188.49	-	-
		6	**4940**	97.77	P: ColRNAI	-
20_GR_12	258	1	**5395894**	1	C	*blaSHV-11*, *fosA*, *oqxA, oqxB*
		2	**170467**	1.77	P: $IncFIB_{pKpn3}$; IncFIIK	*aph(3')-Ia, blaKPC-2*, *blaOXA-9*, *blaTEM-1A*
		3	**50979**	1.42	P: IncN	*aph(3'')-Ib*, *aph(6)-Id*, *blaTEM-1A*, *dfrA14*, *sul2, tet(A)*
		4	**43380**	1.78	P: IncX3	*blaSHV-12*
		5	**13841**	10.82	P: ColRNAI	*aac(6')-Ib*, *aac(6')Ib-cr*

[a] Contig identity indicating chromosome (C) and plasmid (P: replicon (determined via PlasmidFinder 1.3)) sequences

[b] Resistance genes determined via ResFinder 3.0 and displayed in alphabetical order. **Bold** indicates a circular contig.

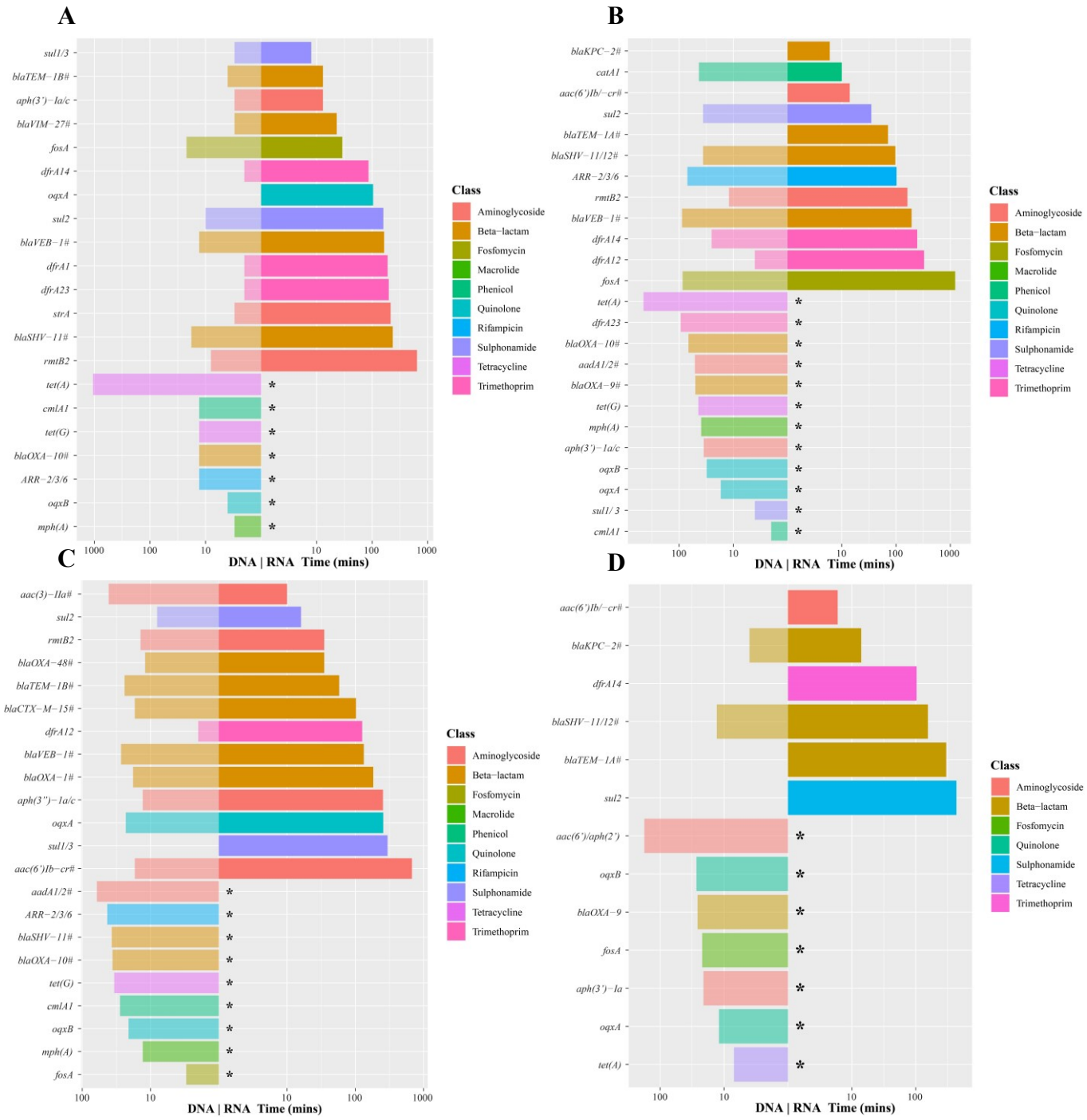

Figure 1. Time required to detect antibiotic resistance genes via the real-time emulation analysis using MinION DNA sequencing and direct RNA sequencing data. (**A**) 1_GR_13, (**B**) 2_GR_12, (**C**) 16_GR_13 and (**D**) 20_GR_12. Legend colours identify the class of antibiotic to which the gene confers resistance, / on y-axis indicates reads detected more than one resistance gene and # is a family of genes detected (>3). An asterisk (*) indicates the inability for direct RNA sequencing to detect this gene. Albacore 2.2.7 base-called sequences were used and all reads (pass and fail) were included in this analysis.

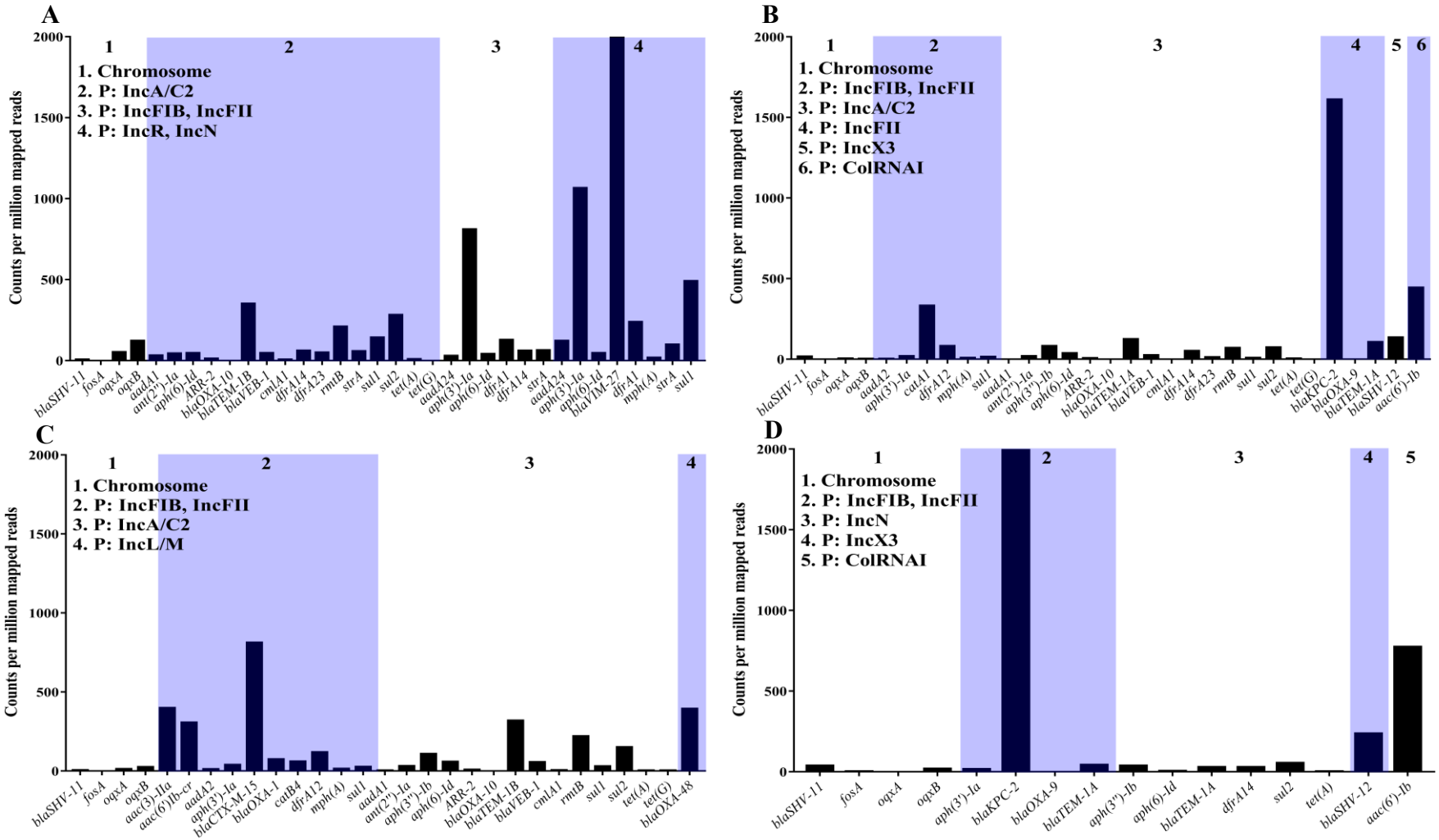

Figure 2. Expression of resistance genes determined using counts per million mapped reads. Due to differing levels of rRNA depletions across samples, reads mapping to rRNA were removed. Strains investigated include (**A**) 1_GR_13, (**B**) 2_GR_12, (**C**) 16_GR_13 and (**D**) 20_GR_12. X-axis depicts the resistance genes and are grouped based on the location in the genome where P indicates a plasmid followed by replicon identity. Albacore 2.2.7 base-called sequences were used and all reads (pass and fail) were included in this analysis.

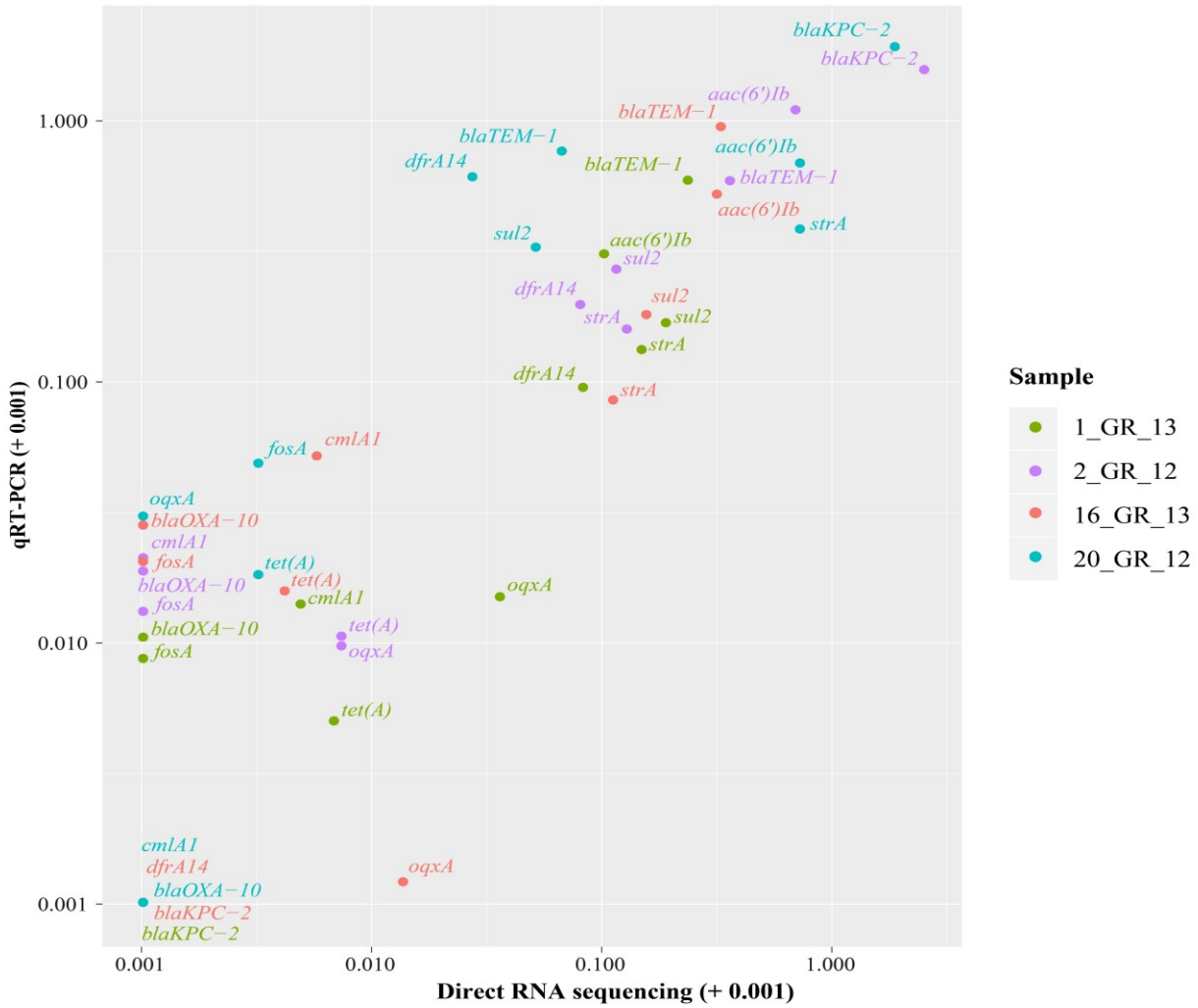

Figure 3. Correlation between resistance genes detected via direct RNA sequencing and validated using qRT-PCR. Relative expression was calculated via normalizing to the housekeeping gene, *rpsL* for both direct RNA sequencing (log2(gene/*rpsL*) and qRT-PCT ($2^{-\Delta\Delta CT}$). Due to high similarity between certain genes, several primers recognise more than one gene. These include *aac(6')Ib: aac(6')Ib-cr, aadA24*; *strA: aph(3")-Ib* and *blaTEM-1: blaTEM-1A, blaTEM-1B.*

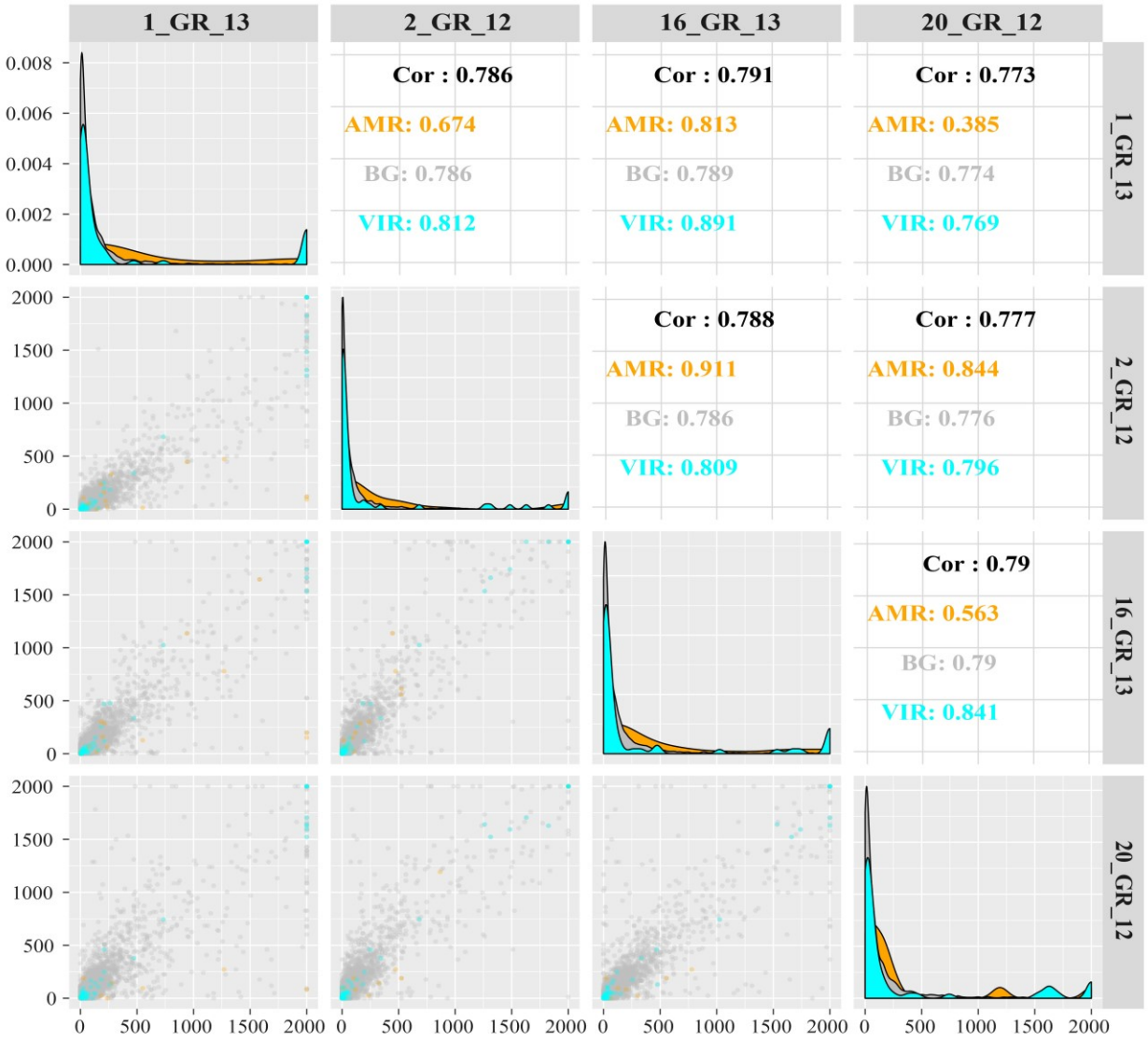

Figure 4. Correlation between gene expression in counts per million mapped reads (capped at 2000 cpm) observed in direct RNA sequencing between the four XDR *K. pneumoniae* isolates. Top panels display spearman correlation coefficients. The diagonal panel displays the histogram of gene expression levels (in cpm) for each sample. Colours indicate categorization of gene: antimicrobial resistance genes (AMR) (as per ResFinder 3.0), virulence genes (VIR) (determined via RAST) and all other genes or background genes (BG) are displayed.

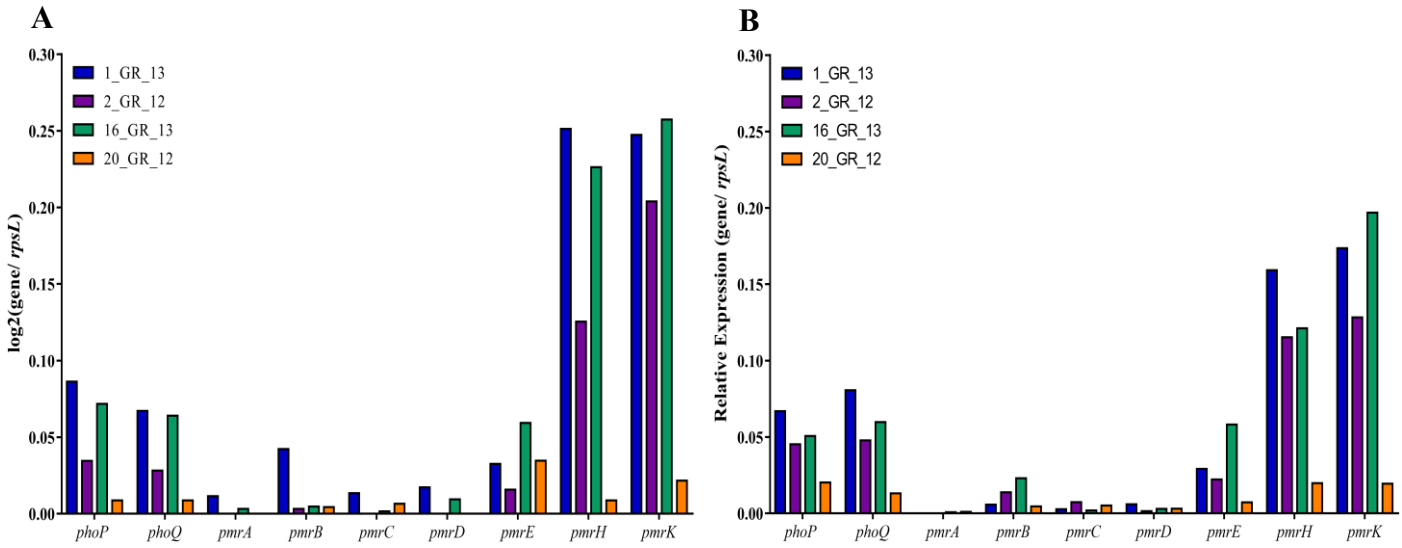

Figure 5. Expression of genes associated with the polymyxin resistance pathway. Comparison between (**A**) direct RNA sequencing and (**B**) qRT-PCR. Isolates harbouring resistance to polymyxins (MIC: >2 µg/mL) include 1_GR_13, 2_GR_12 and 16_GR_13.