Multi-species *bla*$_{NDM}$ outbreak in multiple tertiary and a primary healthcare facility in Merseyside, UK, driven by a combination of multi-species plasmids and small clonal outbreaks.

Caitlin Duggan[1]; Charlotte Brookfield[2]; David Lawrie[2]; Victoria Owen[2]; Timothy Neal[2]; James Cruise[2,3]; Alice J Fraser[1]; Fabrice E Graf[4,] Daire Cantillon[1]; Joseph M Lewis[2,4,5]; Thomas Edwards[1,*]; Eva Heinz[4,6,*]

1: Department of Tropical Disease Biology, Liverpool School of Tropical Medicine, Liverpool, UK.
2: NHS University Hospitals of Liverpool Group, Liverpool, UK.
3: Clatterbridge Cancer Centre NHS Foundation Trust, Liverpool, UK.

4: Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK.

5: Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, Liverpool, UK.
6: Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, UK.

***Correspondence:** eva.heinz@strath.ac.uk, thomas.edwards@lstmed.ac.uk

Keywords: beta-lactamases, carbapenem resistant Enterobacterales, NDM, plasmid-mediated resistance, epidemiology, clonal outbreaks.

Data availability statement: All sequence data is available under bioproject PRJEB89345, a detailed overview of accessions is given in Table S1. All underlying data and methods are provided in the publication or as supplementary information.

Abstract

Background: The *bla*$_{NDM}$ carbapenemase gene provides Gram-negative bacteria the ability to hydrolyse almost all β-lactam antibiotics. A 90% year on year increase from 2022 to 2023 in prevalence of *bla*$_{NDM}$ in clinical isolates was identified in a clinical microbiology laboratory serving Merseyside, UK hospitals and GP practices in 2023. We analysed isolates identified to disclose the species and plasmid diversity and to understand the mechanism of dissemination of *bla*$_{NDM}$ across Merseyside hospitals.

Methods: Data regarding the sample type, the isolation date and location of samples was collected and analysed to establish initial transmission hypotheses; 64 Gram-negative bacterial isolates which tested positive for *bla*$_{NDM}$ using routine diagnostics underwent short read whole-genome sequencing. The samples were derived from seven Merseyside hospitals and a Merseyside GP practice, of these, 24 were also long-read sequenced to

produce local reference assemblies and enable us to confidently resolve resistance gene locations and plasmids.

Results: Overall, six species were identified, including *Enterobacter hormaechei* (n=22) *Klebsiella pneumoniae* (n=21), *Escherichia coli* (n=15) *Citrobacter freundii* (n=3), *Citrobacter brakii* (n=1) and *Enterobacter kobei* (n=1). Most multi-isolate species belonged to a number of sequence types (IQR, 1-10), apart from *K. pneumoniae* where 18/21 were ST101. Four variants of bla_{NDM} were identified: bla_{NDM-1} (n=43), bla_{NDM-5} (n=12), bla_{NDM-6} (n=1), bla_{NDM-14} (n=1), and in six isolates bla_{NDM} was not identified by sequencing. Long-read sequencing of 24 representative isolates found that bla_{NDM-1} was commonly encoded by a IncHI2/IncHI2A plasmid (82.4%, n=14) whereas bla_{NDM-5}, (n=4) was encoded on different plasmids in each isolate but had the same conserved gene cassette including bla_{NDM-5} in all instances. Potential evidence of transposition events was noted in two isolates, where bla_{NDM-1} was found on a small (8.5kbp) untypable plasmid.

Conclusions: This study captures an outbreak of bla_{NDM} positive *Enterobacterales* in Merseyside hospitals and highlights the complex epidemiology of spread mediated both by expansion of successful lineages (*K. pneumoniae* ST101) and plasmids. A dominant IncHI2/IncHI2A plasmid that has disseminated bla_{NDM-1} across different species and subsequently led to small clonal outbreaks within-hospitals. This highlights the cross-species movement of AMR elements as an important factor in AMR dissemination, and the importance of both species-specific surveillance to monitor local outbreaks, but also species-agnostic surveillance to include plasmids spreading across species.

Introduction

In 2019, it is estimated that 1.27 million global deaths were attributed to bacterial antimicrobial resistance (AMR), with the highest number of deaths were attributed to AMR caused by *Escherichia coli* and *Klebsiella pneumoniae* (1). The use of β-lactams antibiotics have been a mainstay option to treat infections caused by *Enterobacterales*, including the use of penicillins, and cephalosporins, in particular third generation cephalosposrins. However, increasing resistance to these agents has led to an increased reliance on carbapenems, which are β-lactams not degraded by extended spectrum β-lactamase (ESBL) enzymes and thus an important, often last-resort, treatment option (2). There are few options beyond carbapenems, and carbapenemase resistance genes are increasingly widespread, facilitated by plasmid mediated horizontal gene transfer of the resistance genes, combined with the rapid global expansion of high-risk clones associated with these (3, 4).

Carbapenem resistant *Enterobacterales* are now an issue of critical priority on the World Health Organization's Bacterial Priority Pathogens List 2024 (5).

Carbapenemase-producing *Enterobacterales* (CPE) infections within hospitals in high-income countries are associated with worse clinical outcomes compared to sensitive isolates (6). The incidence of carbapenemase-producing organisms (CPO) is rising within England, with 7,438 cases of acquired carbapenemases reported to UKHSA in 2024, in comparison with 666 cases reported 2021 (7-9). Treatment options for CPE are limited: either repurposing of older, toxic agents (e.g. polymyxins) or newer beta lactam/beta lactamase combinations which may not be available.

The New Delhi metallo-β-lactamase (bla_{NDM}) carbapenemase gene belongs to the Ambler class B of carbapenemases has especially limited treatment options as many newly available inhibitors (such as vaborbactam) may not have activity (10, 11). Between 2022 and 2023, there was a 90% year on year increase in the number of inpatient samples positive for bla_{NDM} in Liverpool's central clinical microbiology laboratory (12). This study aimed to combine whole genome sequencing (WGS) and clinical metadata to understand the genomic epidemiology of this increase.

Methods

Ethics statement

No identifiable data was analysed during this project. As an anonymised use of routinely collected data and following UK Health Research Agency (HRA) guidance the study does not require formal research ethics committee review. It was approved by HRA (IRAS ID 345990).

Study setting

Liverpool Clinical Laboratories provides a clinical microbiology service to the Royal Liverpool and Aintree University Hospitals (large acute teaching hospitals including busy emergency departments and regional infectious diseases, stroke and trauma services), Broadgreen hospital, the Liverpool Women's, Liverpool Heart and Chest Hospitals (specialist obstetrics/gynaecology and cardiology/cardiothoracics/respiratory medicine hospitals respectively), the Walton Centre (specialist neurosurgical/neurology hospital) and Clatterbridge Cancer Centre – Liverpool (specialist oncology and haemato-oncology centre including stem cell transplant service). Together these hospitals provide general and emergency care to 630,000 people across Merseyside and specialist care to two million people in the North-West.

Clinical sample collection and processing

Isolates used in this study came from CPE colonisation screening, and clinical samples where infection was suspected or were collected from the hospital environment. Patients were screened using local CPE screening guidelines, following UKHSA (13) recommendations (17). The screening swab was cultured with meropenem selection, and the Gene-Xpert Carba-R assay (Cepheid, United States) was used to identify the CPE gene present in any samples which grew colonies. Clinical samples in which infection was suspected were cultured using standard techniques following UK standards for microbiology investigations (SMI), and speciated with MALDI-TOF mass spectrometer (Bruker, United States) followed by antimicrobial sensitivity testing (AST) using disc diffusion following EUCAST guidelines. For any isolates with a zone size smaller than the EUCAST carbapenemase screening diameter, Xpert Carba-R assay was used to define presence of carbapenemase genes. Following processing, all isolates were stored on microbank beads at -80C.

From January 2023 to December 2023, 68 coliform, Gram-negative bacterial isolates were identified to be positive for the bla_{NDM} gene, and 64 of these could be retrieved for further investigation.

Short-read sequencing

For short-read sequencing, a single colony from each of the isolates was sub-cultured overnight in a Luria-Bertani (LB) broth at 37°C and at 200rpm. DNA was extracted using the DNEasy™ Blood and Tissue Kit (Qiagen, Germany) or the Wizard® gDNA Purification Kit (Promega, United States), following manufacturers' protocol including pre-treatment for Gram-negative bacteria for the DNEasy™ Blood and Tissue Kit (Qiagen, Germany). DNA was quantified using the Qubit™ Fluorometer (Thermo Fisher Scientific, United States) using the Broad Range Assay Kit (Thermo Fisher Scientific, United States) and where possible normalised to 50ng/µL. 35µL of DNA from each isolate was then transported on dry ice to GeneWiz (Azenta Life Sciences, Germany) who completed library preparation and 150bp paired-end sequencing on a Novaseq (Illumina, United States).

Long-read sequencing

A single colony from each of the 24 isolates selected for long-read sequencing (Table S1) was sub-cultured overnight in a LB broth at 37°C and at 200rpm. Two of the isolates (NDM108, NDM159), were eventually resequenced after plating on LB agar (8µg/mL meropenem) to prevent bla_{NDM} loss. After overnight incubation, a single colony was then sub-cultured in LB broth with 8µg/mL of meropenem and incubated overnight at 37°C and at

200rpm. The MasterPure Complete DNA and RNA Purification Kit (LGC Biosearch Technologies, United Kingdom) was used to extract DNA, following the total nucleic acids purification protocol for cell samples. For NDM152, the Wizard DNA Purification Kit (Promega, United States) was used instead due to poor DNA recovery using the MasterPure Complete DNA and RNA Purification Kit (LGC Biosearch Technologies, United Kingdom). DNA was quantified using the Qubit™ Fluorometer (Thermo Fisher Scientific, United States) with the Broad Range Assay Kit (Thermo Fisher Scientific, United States) and the 4150 TapeStation System (Agilent, United States) using the Genomic DNA ScreenTape assay (Agilent, United States). The Native Barcoding Kit 24 V14 (SQK-NBD114.24) (ONT, United Kingdom) and protocol from ONT was used, with all recommended third-party consumables and QC checks. The barcoded library was added to the R10.4.1 Flow Cell on a MinION device (ONT, United Kingdom) before sequencing for 72-hours. Repeat sequencing was completed using the same gDNA and the same ONT kit and reagents, with two isolates per library on a R10.4.1 Flongle (ONT, United Kingdom) for a complete 24-hour sequencing run.

Short-read only *de novo* assemblies

FASTQ files were trimmed using the Trimmomatic (v.0.39) (14) paired end option, with the TruSeq3-PE FASTA file of adaptor sequences and recommended parameters. All FASTQ files were processed through FastQC (v.0.11.9) (15) and collated into MultiQC (v.1.21) (16) for checking. Shovill (v.1.1.0) (https://github.com/tseemann/shovill) was used as the *de novo* short-read assembler for reads. QC for assemblies was assessed using CheckM (v.1.2.2) (17) and Busco (v.5.6.1) (18). Isolates that had a value greater than 5% for contamination in CheckM was inspected manually using the Artemis Comparison Tool (V.13.0.16) (19) using other isolates within the collection as references. GAMBIT v.1.0.0 (20) was used to confirm the species of the isolates, by comparing them to a NCBI derived database produced by GAMBIT (downloaded 29th April 2024).

Long-read *de novo* assemblies

All long-reads were basecalled using the guppy super accurate model option on a GPU (v.6.4.6), and demultiplexed using a CPU (v.6.5.7). The Hybracter (v.0.7.3) (21) assembly pipeline was used to assemble all long-read sequences, using Flye (v.2.9.3) (22) as the assembler, and polish the assemblies with the long and short-read sequences described above. Due to low coverage from long-reads in eight isolates, Unicycler (v.0.5.0 or v.0.5.1) (23) was used as an assembler with the same filtered and trimmed long and short reads, produced from the Hybracter pipeline, as input. Assemblies were polished with Medaka (v.1.11.3) (https://github.com/nanoporetech/medaka), Polypolish (v.0.6.0) (24) and Pypolca (v.0.3.1) (25) with the --careful parameter. All draft assemblies were checked using CheckM

(v.1.2.2) (17); continuity of the presence of AMR genes and plasmid replicons were checked between the short-read and long-read derived assemblies using StarAMR (v0.10.0) (26).

Genomic analysis

All final assemblies were annotated using Prokka (v.1.14.6) (27). Mash (v.2.2.2) (28) was used to produce a file of the estimated genome distances between all isolates for use in hierarchical clustering within heatmaps (28). MLST were called using ariba (v.2.14.6) (29) with the PubMLST typing schemes for *Citrobacter freundii*, *Enterobacter cloacae* complex (ECC), *E. coli* (Achtman's seven-gene scheme (30)), and *K. pneumoniae* (all downloaded 9th May 2024).

The Panaroo (v.1.5.0) (31) pan-genome pipeline was used to calculate the pan genome for each of the assemblies using the '--clean-mode strict' and '-a core' options. From the core gene alignment produced, SNP-sites (v2.5.1) (32) was used to extract only SNPs within the alignment. This fasta file was passed to IQ-TREE (v.2.3.3) (33) to produce a best fit maximum likelihood tree file for the genus using a general time reversible model for all trees. A SNP-based phylogram of ST101 *K. pneumoniae* isolates was generated using Snippy (v.4.6.0) (https://github.com/tseemann/snippy), with the chromosome from the final long-read assisted genome assembly of NDM144 as the reference, and the trimmed short-reads for the rest of the ST101 isolates. Pairsnp (v.0.3.1) (https://github.com/gtonkinhill/pairsnp) calculated the pairwise SNP distances of each ST101 isolate. The core alignment produced by Snippy was then cleaned using snippy-clean_full_aln (34). Gubbins (v.3.3.1) (35) was run using the GTRGAMMA model. SNP-sites (v.2.5.1) (32) was then used to produce a FASTA file of only SNPs between the isolates after recombination were removed. RAxML-NG (v.1.2.2) (36) was used to produce a general time reversible model, maximum likelihood tree. Pyjar (https://github.com/simonrharris/pyjar), as implemented for python (v.3.7), was used to produce a joint ancestral reconstruction tree from the best maximum likelihood tree outputted from RAxML-NG. Based on community standards, a sub-lineage threshold of 5 SNPs and a threshold of 3 ancestral nodes to the root was selected for rPinecone (v.0.1.0). The SNP scaled tree, generated using the pyjar derived tree file and alignment, was then visualised in R (v.4.3.1) using the R packages ggtree (v.3.10.0), ggplot2 (v.3.5.0), and patchwork (v.1.3.0) (37-39).

Kleborate (v.2.3.2) output, with the implemented SRST2 (40), was used to predict virulence genes for each *K. pneumoniae* isolate (41). The CARD and SRST2 databases (download dates 14th February 2024 and 3rd May 2024 respectively), were used when running ariba (v.2.14.6) to look for AMR genes in the isolates and the ResFinder database (downloaded 3rd May 2024) was used alongside the other databases to check for consistency in the bla_{NDM}

variant identified (29). From the ariba AMR results using the SRST2 database, any results for the *catBx* gene were excluded from data analysis (42). To assess the presence of plasmid replicons, the PlasmidFinder database (43) (download 24 April 2024) was used with ariba. Ariba results for both AMR genes and plasmid replicons were visualised in R using the pheatmap package (v.1.0.12).

Abricate (v1.0.1) (https://github.com/tseemann/abricate) was used with both the CARD and ARG-ANNOT databases (download June 2024) to confirm the bla_{NDM} variant and the contig it was on. StarAMR (v0.10.0) (26) was used to confirm the plasmid replicons present and the plasmid replicons of the plasmid of the bla_{NDM} contig. Bandage (v.0.8.1) (44) was used to export the nucleotide sequence of the plasmid of interest to Proksee (45) to visualise in a circularised figure. Within Proksee, the CARD Resistance Gene Identifer tool (tool v.1.2.1) was used to annotate the plasmid figure with AMR genes; these genes were manually checked for consistency with Abricate results. The BLAST tool (tool v.1.3.1) was used to compare and visualise the plasmid sequences of different isolates.

To compare regions of certain isolates of interest, Artemis (v.16.0.17) (46) was used on the annotated gff file to select a region upstream and downstream from the gene of interest. This region was then re-annotated with Prokka (v.1.14.6) (27), before being compared and visualised with Clinker (v.0.0.29) (47).

Mucoid phenotypic characterisation

Each *K. pneumoniae* isolate was incubated overnight at 37°C on a freshly prepared LB agar plate, and an inoculation loop was used to stretch a single colony, repeated three times. The string test results were evaluated according to Shon, Bajwa and Russo (48).

Results

Outbreak driven by six different bacterial species

We characterised 64 *Enterobacterales* isolates routinely collected during 2023 from seven tertiary healthcare facilities and a singular primary healthcare facility in Merseyside; one isolate was excluded from further analysis as it was a mixed culture of both *En. hormaechei* and *E. coli,* resulting in 63 isolates used for detailed analyses. In this time, screening practices to identify CPEs remained consistent. There were no changes to the infection prevention and control (IPC) cleaning measures within hospitals, but standard practices were reinforced if CPEs were identified.

Isolates were a mixture of colonisation screening samples (88.9%, n=56), clinical samples from patients admitted to one of the study healthcare facilities (n=6), and a single environmental isolate (NDM107, *Enterobacter kobei*) from an en-suite shower trap within a patient room in Hospital 2 (Figure 1A; Figure 1B). WGS of all isolates determined that these belong to six species (*C. brakki*, n=1; *C. freundii*, n=3; *En. hormaechei*, n=22; *En. kobei*, n=1; *E. coli*, n=15; *K. pneumoniae*, n=21). All species containing more than one isolate were identified in multiple healthcare facilities throughout 2023 (Figure 1A).

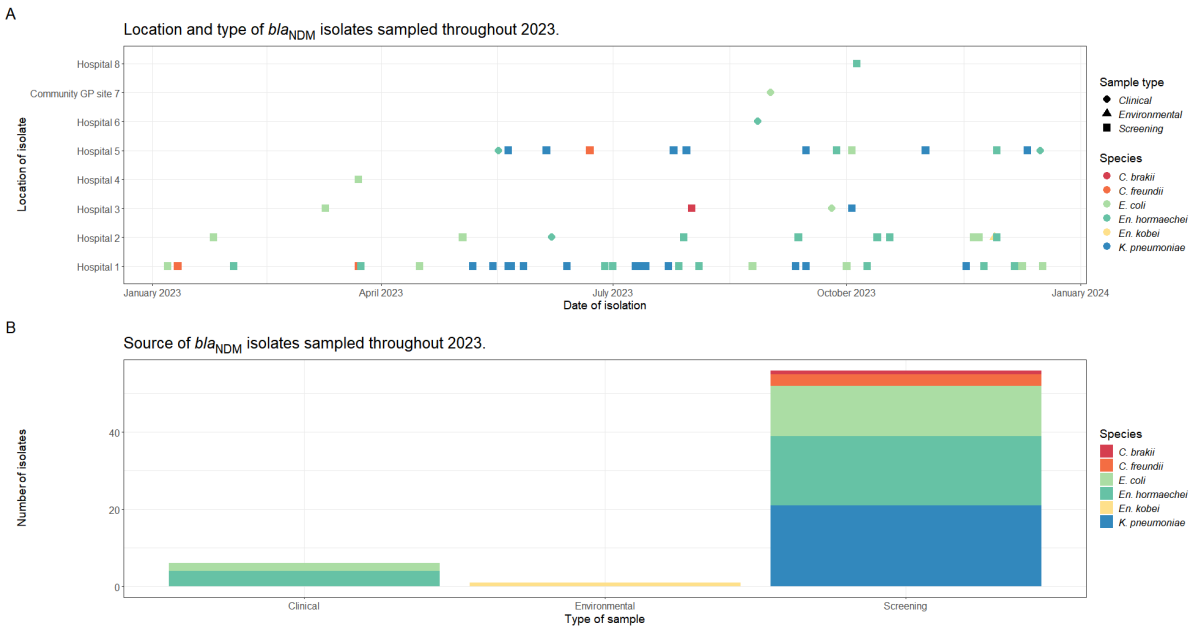

Figure 1: Characterisation of isolates sampled throughout 2023 from tertiary and primary healthcare facilities in Merseyside, UK. **A:** Timeline of the isolation of *bla*$_{NDM}$ isolates, between the seven hospitals and a community GP facility in Merseyside, each dot represents an isolate on the day that it was sampled from the patient, data is stratified in shape by its sample type and in colour by its WGS identified species. **B**: Source of *bla*$_{NDM}$ isolates, stratified by their species as identified by WGS, the y-axis represents the total number of isolates throughout 2023 to have that sample type.

Population structure reveals mix of clonal and single-isolate lineages

The *Citrobacter* spp. isolates consisted of three *C. freundii* isolates and one ST370 *Citrobacter brakki*, as shown in Figure 1. Two of the *C. freundii* were screening samples belonging to ST597 (NDM149; NDM154) taken from two hospitals (Hospital 1; Hospital 5) within three months of each other and the other, NDM152, was ST112.

The *Enterobacter* isolates, used for detailed analysis, belonging to the *Enterobacter cloacae* complex (ECC) were one ST56 *En. kobei*, and twenty-two *En. hormaechei* (95.7%) (Figure 2). The *En. hormaechei* isolates represented 10 STs; with ST111 and ST118 the most

prevalent STs at four isolates each. ST111 was observed across the span of two months in three of the study hospitals (Hospital 1; Hospital 5; Hospital 6) in three screening and one clinical isolate (Hospital 6). ST118 *En. hormaechei* isolates were observed over the course of eight months in three of the hospitals (Hospital 1; Hospital 2; Hospital 5) in three screening samples and one clinical sample. ST106 and ST161 both had three isolates; ST106 was observed in two screening samples from the same hospital (Hospital 2) three months apart and a clinical sample two months after the last screening ST106 sample was identified in Hospital 5 and ST161 was identified in three screening isolates from two study hospitals (Hospital 1; Hospital 5) in less than two months. ST104 and ST145 contained two isolates each, ST104 isolates were seen in the span of nine months and were both screening isolates from Hospital 1, and ST145 was seen in both a clinical and screening isolate from Hospital 2 and Hospital 1 (respectively) within a month. The remaining four STs were singletons (Figure 2; Table S1).

The fifteen *E. coli* isolates, sampled from fourteen patients, belonged to eleven STs (Table S1). The ST648 had four isolates, sampled from three of the study hospitals (Hospital 1-3) across two months (September – November 2023), three of these isolates were screening and one was a clinical sample (Figure 1; Table S1). ST405 had two isolates which were both screening samples from two separate hospitals with two months between their isolation; the rest of the isolates were singletons (n=8) (Figure 2; Table S1).

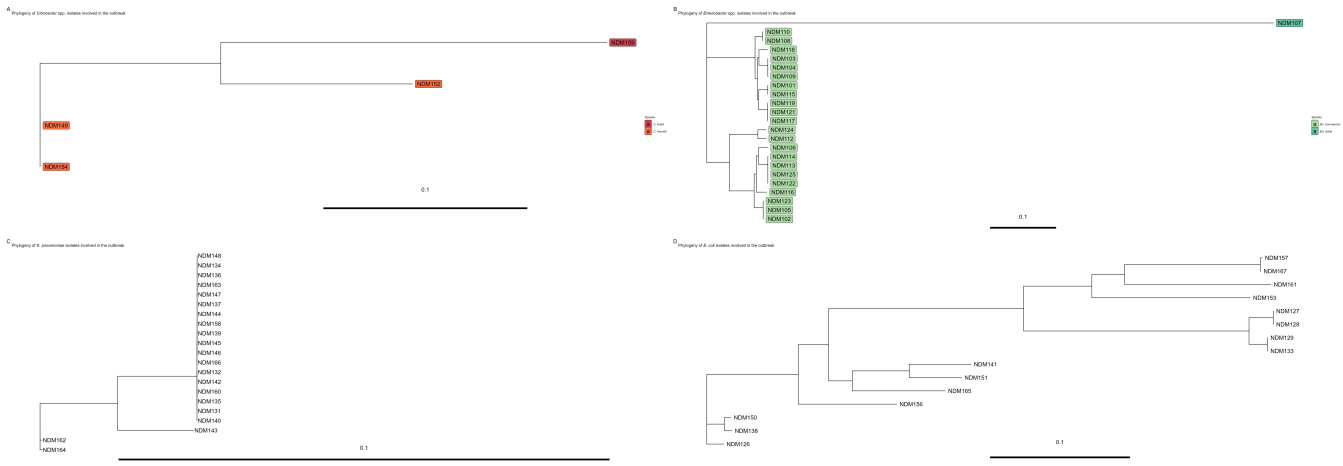

Figure 2: Phylogenetic structure of isolates within study. **A**: Maximum likelihood phylogeny of *Citrobacter* spp. isolates using a general time likelihood model; isolates are colour-coded according to species. **B**: Maximum likelihood phylogeny of *Enterobacter* spp. isolates using a general time likelihood model; isolates are colour-coded according to species. **C**: Maximum likelihood phylogeny of *K. pneumoniae* isolates within the study using a general time likelihood model. **D**: Maximum likelihood phylogeny of *E. coli* isolates within the study using a general time likelihood model.

85.7% (n=18) of the 21 *K. pneumoniae* isolates belonged to ST101, two were ST147 isolates identified in screening samples from Hospital 5 over the course of three months, and one was ST3702 (Figure 2). No *K. pneumoniae* isolates showed mucoid phenotype using the string test, and only the two ST147 isolates encode 1/5 (NDM162) and 4/5 (NDM164) virulence genes, respectively, as predicted by Kleborate (Table S2). NDM162 had the yersinabactin siderophore region; NDM164 encodes for yersinabactin, aerobactin and a partial region of the *rmpA2* gene (Table S2). To better resolve the ST101 *K. pneumoniae* population structure, we performed a whole-genome SNP analysis, illustrating two sublineages and three isolates without a sublineage within the sequence type (Figure 3A). The number of pairwise SNPs between the ST101 *K. pneumoniae* isolates ranged from 2 to 69, as shown in Figure 3b. Sublineage one contained four isolates and sublineage two contained eleven isolates. All ST101 K. pneumoniae isolates were screening samples; isolates belonging to sublineage 1 were sampled from Hospital 1 only between across three months of 2023 (Figure 3C). Isolates belonging to sublineage 2 were sampled from three of the hospitals within the study (Hospital 1, Hospital 3, and Hospital 5), between May and October of 2023, six days after the second sublineage 2 ST101 isolate was identified at Hospital 1 another sublineage 2 ST101 isolate was identified at Hospital 5 (Figure 3C)

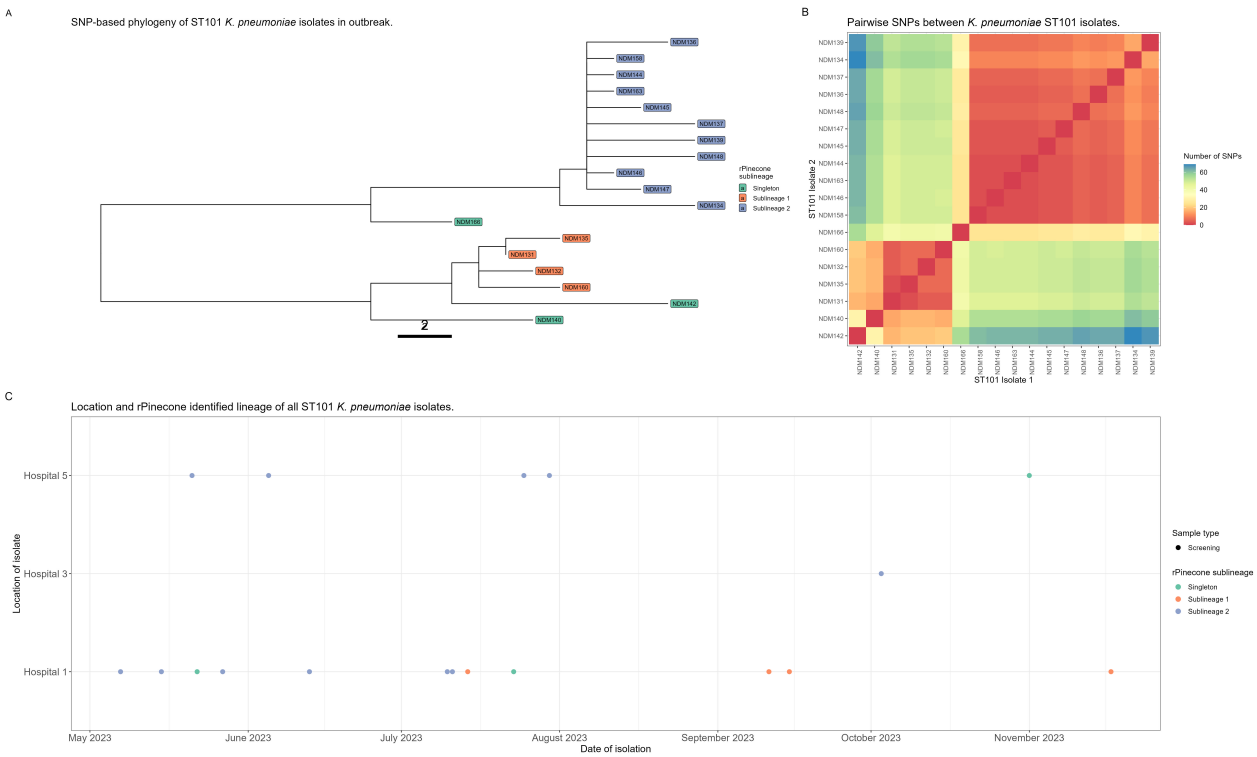

Figure 3: SNP based analysis of ST101 *K. pneumoniae* isolates where the first ST101 isolate sampled, NDM144, was used as a reference for all analyses. **A:** SNP based phylogeny, isolates are colour coded by their sublineages that were determined via rPinecone analysis (new sublineages were identified by 5 SNP threshold and the major sublineage threshold was set to 3 ancestral nodes), the branch lengths within the phylogram is indicative of the number of SNPs. **B:** Pairwise SNP comparison of all ST101 isolates, isolates are ordered along both axes by hierarchical clustering

using their estimated genome distances from Mash. **C**: Timeline of the isolation of ST101 *K. pneumoniae* isolates, between the three hospitals (Hospital 1, Hospital 3, Hospital 5) they were identified within between May and November 2023; isolates are stratified by their rPinecone analysis lineages.

Acquired resistance gene profiles differ between species

Assessing acquired resistance genes against all classes detected 59 genes in total and through clustering shows the presence of some species specific AMR genes (bla_{ACT-2}; *oqxB*), however many AMR genes are observed across multiple species (bla_{TEM-1D}; *sul2*) (Figure 4A; Table S3), with *K. pneumoniae* isolates encoding for an average of 14.9 resistance genes against 10 classes; *E. coli* isolates carried an average of 10.9 resistance genes against 10 classes; *En. hormaechei* isolates carried an average of 12.2 resistance genes against 10 classes and *C. freundii* carried an average of 10.3 resistance genes against 7 classes (Figure 4A). Whilst the isolates were originally selected for bla_{NDM} presence based on hospital diagnostics, six of the isolates in our study had no predicted bla_{NDM} gene following sequencing, which could be due to loss of the resistance gene in the course of re-growing between the initial diagnostics and sequencing.

Additional carbapenemase genes of bla_{IMP-1} (n=1), bla_{VIM-1} (n=1), bla_{VIM-1} and $bla_{OXA-48-like}$ (n=1), and $bla_{OXA-48-like}$ (n=1) were also observed in NDM108 (*En. hormaechei*), NDM159 (*C. brakii*), NDM161 (*E. coli*) and NDM165 (*E. coli*) respectively (Figure 4A). In three of these isolates (NDM108, NDM159 and NDM161) the bla_{NDM} gene could not be detected in our sequence data. We could observe several bla_{NDM} gene variants, with bla_{NDM-1} as most common variant in 4/6 species (43/63 isolates), being identified within samples from six of the study hospitals (Hospital 1, Hospital 2, Hospital 3, Hospital 5, Hospital 6, Hospital 8) between January and December of 2023 (Figure 4B; Figure 4C). However, in *E. coli* the bla_{NDM-5} variant was the most prevalent, as shown in Figure 4B. The bla_{NDM-5} variant was identified in samples from five hospitals (Hospital 1-5), and in a sample from a community GP, in samples isolated from January to December 2023 (Figure 4C). Two *K. pneumoniae* isolates furthermore encoded bla_{NDM-6} and bla_{NDM-14}, however the SRST2 database identified the bla_{NDM-14} gene as the bla_{NDM-8}.

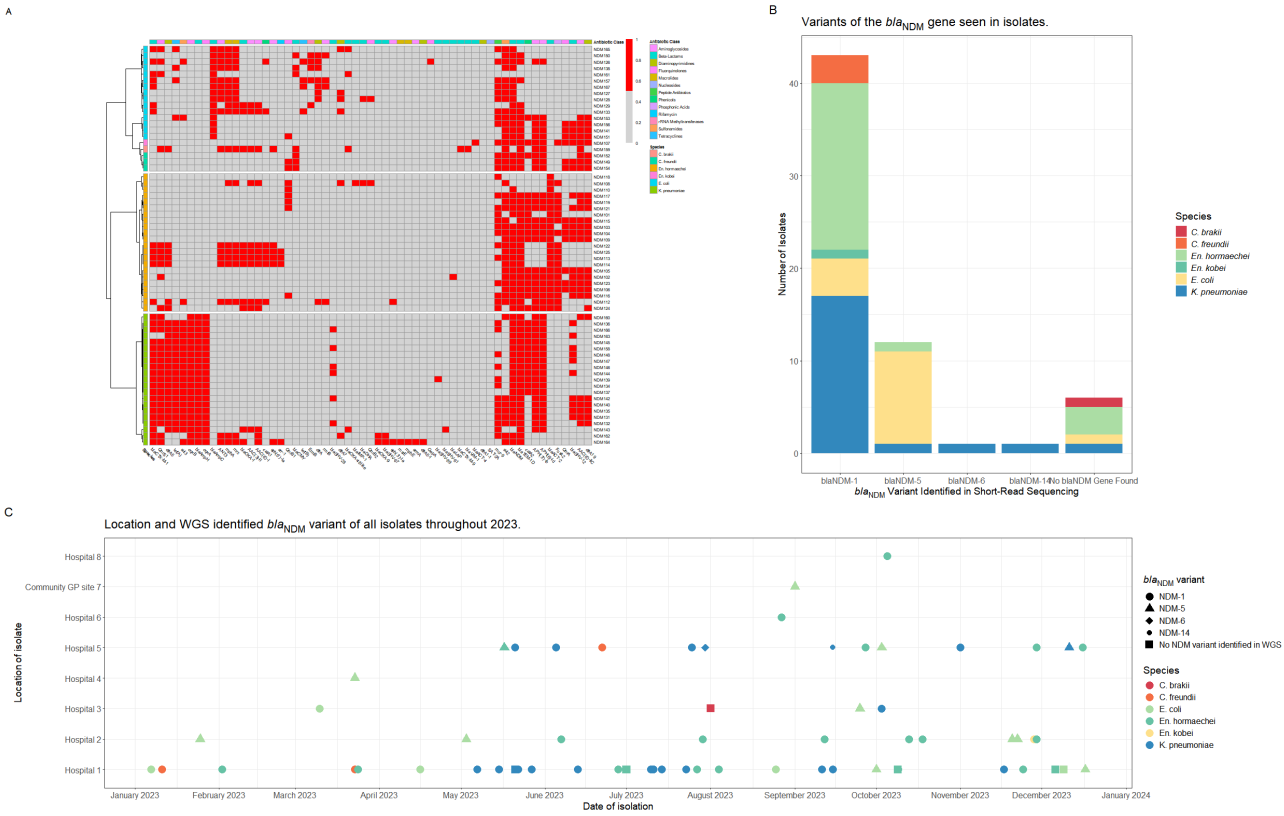

Figure 4: Characterisation of resistance gene profiles. **A:** Clustered heatmap of AMR genes. Isolates are arranged on the guidance tree in the y-axis using hierarchical clustering from their estimated genome distances derived from Mash, and the columns represent resistance genes which have been arranged by clustering; a red colour denotes the identified presence of this gene and grey denotes absence of this gene within that isolate. **B:** *bla*$_{NDM}$ gene variants stratified by species, x-axis shows all variants of the *bla*$_{NDM}$ gene identified using WGS data; y-axis shows absolute numbers over the whole study period (2023) **C**: Timeline of all the isolation of all isolates stratified by their WGS identified *bla*$_{NDM}$ variant and species. The y-axis represents the location that the patient was within when the sample was taken, shown throughout 2023, from January to December. Each point denotes a singular isolate with shapes indicating the *bla*$_{NDM}$ variant identified using WGS, colours indicate species, as shown in the legend.

Genomic characterisation of plasmid replicons.

As displayed in Figure 5, the median number of plasmid replicons seen per isolate was 4 replicons (IQR, 3-5). Using short-read sequencing data, the most common plasmid replicons were the IncHI2 and IncHI2A replicons; these always occurred together in the same 73% of isolates (n=46) (Figure 5; Table S4). The IncHI2 and IncHI2A plasmid replicons were observed the most commonly in *En. hormaechei* and *K. pneumoniae*.

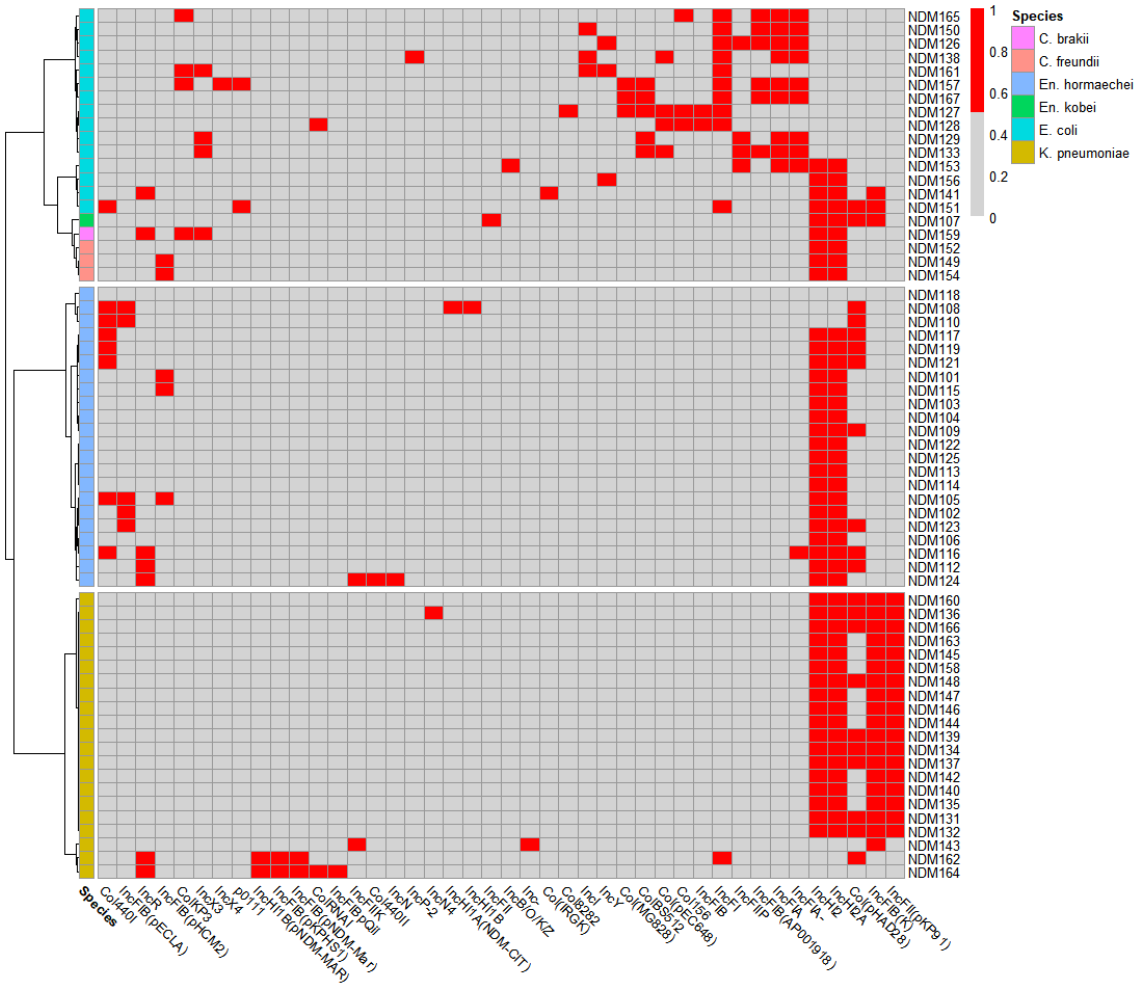

Figure 5: Clustered heatmap of plasmid replicons observed in all isolates analysed, using short-read sequencing data, isolates are arranged on the guidance tree in the y-axis using hierarchical clustering from their estimated genome distances derived from Mash. The columns on the x-axis represent plasmid replicons, that have been identified in at least one of the isolates; their order is organised by clustering. A red colour within the grid indicates that the replicon is present in that isolate while the grey colour indicates absence

For three isolates (NDM108, NDM159, NDM161), a *bla*$_{NDM}$ gene was not identified in assemblies using long-reads, consistent with short-read data. All other isolates selected for long-read sequencing (n=21) encoded the *bla*$_{NDM-1}$ (n=17) or the *bla*$_{NDM-5}$ (n=4) variants. Figure 6 presents the species and the location of the respective *bla*$_{NDM}$ gene in each isolate where long-read sequencing was completed.

The presence and location of the *bla*$_{NDM}$ gene in long-read sequenced isolates

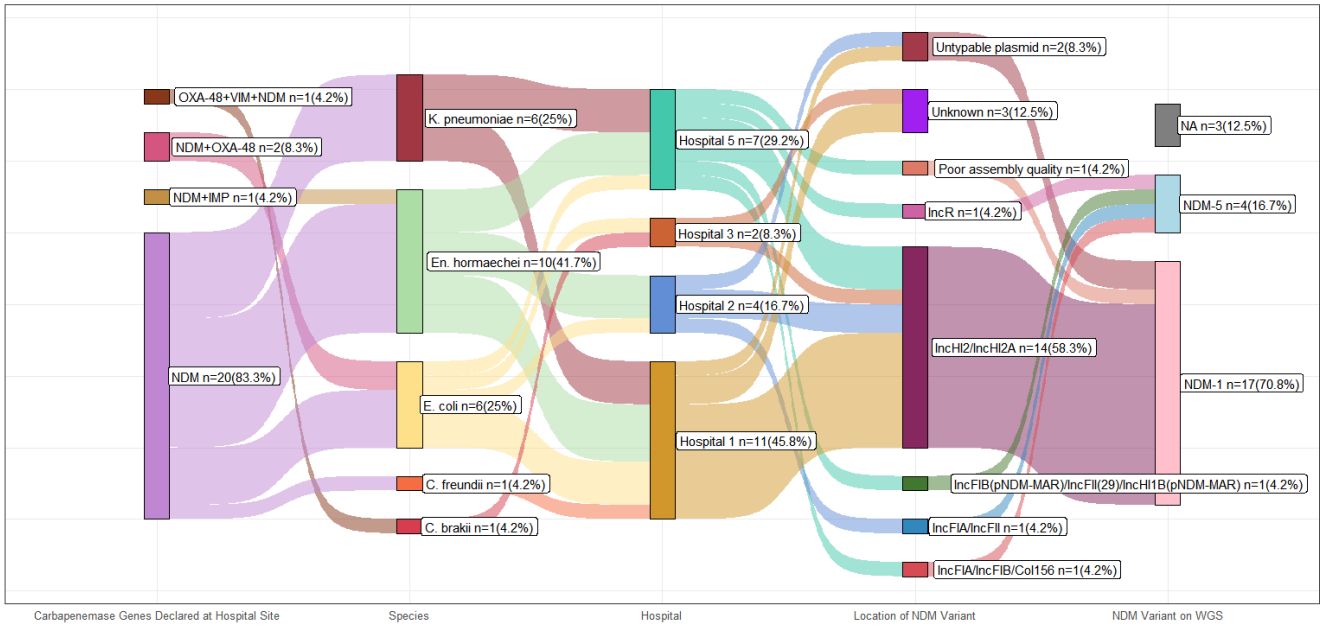

Figure 6: Sankey diagram characterising the isolates selected for further analysis through long-read sequencing, including the hospital sampled from and the resolved location of *bla*$_{NDM-1}$. In three of the isolates long-read sequenced, a *bla*$_{NDM}$ gene could not be identified, consistent with their results using short-read sequencing data.

Plasmid sequences indicate co-circulation of highly similar IncHI/IncHIA plasmids encoding for *bla*$_{NDM-1}$.

The majority of the *bla*$_{NDM-1}$ containing long-read sequenced isolates (82%, n=14) harbour the gene on a IncHI2/IncHI2A plasmid (Figure 6). The *bla*$_{NDM-1}$ containing IncHI2/IncHI2A multi-replicon plasmids had a median size of 377,727bp (IQR, 338,352-384,512). This plasmid additionally harboured multiple other AMR genes, including *bla*$_{TEM-1}$ closely upstream from *bla*$_{NDM-1}$ and *mcr-9* and *bla*$_{SHV-12}$. To observe changes in the IncHI2/IncHI2A multi-replicon plasmid throughout 2023, the first NDM isolate (NDM151) of 2023 which contained the IncHI2/IncHI2a plasmid encoding *bla*$_{NDM-1}$ was used as a reference against all other IncHI2/IncHI2A-*bla*$_{NDM-1}$ plasmids. We also included a IncHI2/IncHI2A plasmid where *bla*$_{NDM}$ has been lost before sequencing but after isolation from the patient (NDM159) (Figure 7). All genomes containing the *bla*$_{NDM-1}$ variant and had long-read sequencing assisted assemblies contained a single copy of *bla*$_{NDM-1}$ gene, apart from NDM121 (a ST161 *En. hormaechei*) which had two copies of *bla*$_{NDM-1}$ on a 392.1 kbp IncHI2/IncHI2A multi-replicon plasmid deemed circularised by Unicycler. The region surrounding the *bla*$_{NDM-1}$ gene displayed in Figure 8 was identical between both isolates apart from the *ble* gene which had 71% similarity.

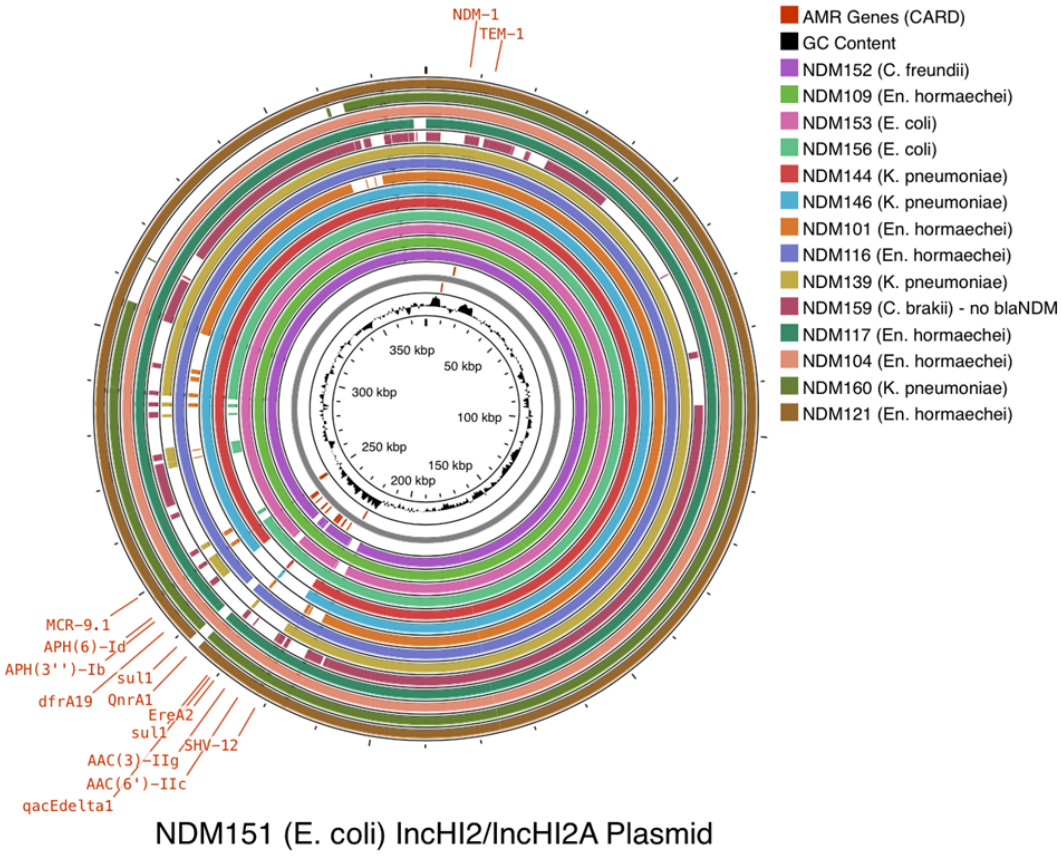

Figure 7: Comparison of IncHI2/IncHI2A contigs from all species where the *bla*$_{NDM-1}$ gene was found, across 2023, with isolates sampled in January 2023 on the inner most rings moving outwards throughout 2023 to isolates sampled in December 2023 in the final rings, AMR genes were annotated on the reference (NDM151) genome which all other plasmids were compared against in a blastn search within Proksee.

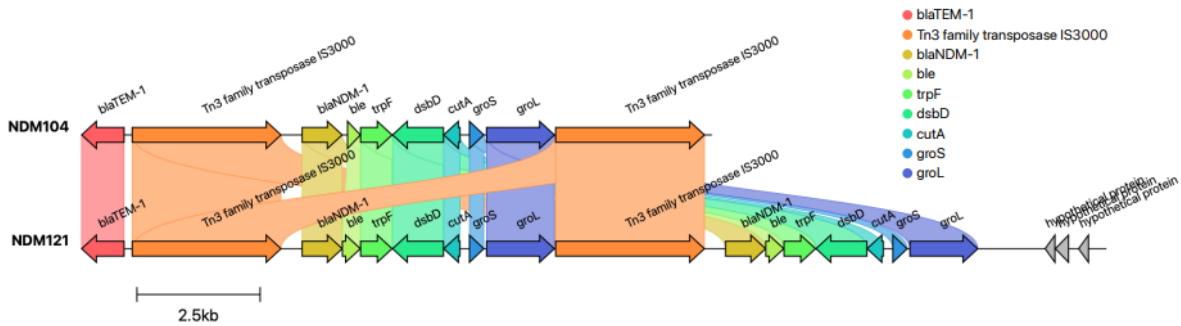

Figure 8: Visualisation of genomic context of our example of two copies of *bla*$_{NDM-1}$ on the IncHI2/IncHI2A plasmid in NDM121 (*En. hormaechei*) compared to the isolate NDM104 (*En. hormaechei)* with a single copy of *bla*$_{NDM-1}$ on the IncHI2/IncHI2A plasmid. In NDM121, the regions 5000 bp upstream and 15,000 bp downstream from the *bla*$_{NDM-1}$ gene was selected to capture the complete flanking region of both copies of *bla*$_{NDM-1}$. Coding sequences were annotated according to Prokka.

Two untypable circular elements containing *bla*$_{NDM-1}$ in isolates NDM103 (8,569 bp) and NDM147 (8,588 bp) were identified and compared in BLASTn using the megablast algorithm

with a 100% identity and 100% query cover. Both elements are shown in the circular element comparison of Figure 9. In a search against the core nucleotide BLAST database of assemblies, the untypable small element of NDM103 showed 99.4% identify and 100% query cover to a 304.6 kbp *bla*$_{NDM-1}$ encoding *K. pneumoniae* plasmid pKJNM8C2.1 (accession CP030858.1) collected from a neonate in India in 2017, and the untypable small element of NDM147 had 100% identity and 100% query cover to the 62.7kbp *bla*$_{NDM-1}$ encoding *E. cloacae* plasmid pNDM1-CBG collected from a sputum sample in China in 2017 (accession CP046118.1). Both isolates had the IncHI2/IncHI2A plasmid replicons identified within the genome on separate contigs. In NDM103 the IncHI2/IncHI2A replicons were identified to be located on the 5,163,405 bp chromosome. In NDM147 the IncHI2/IncHI2A replicons were in a 353.8 kbp plasmid missing the region surrounding the *bla*$_{NDM-1,}$ as shown in Figure 10, when compared to another *K. pneumoniae* ST101 isolate with the IncHI2/IncHI2A plasmid encoding the *bla*$_{NDM-1}$ gene.

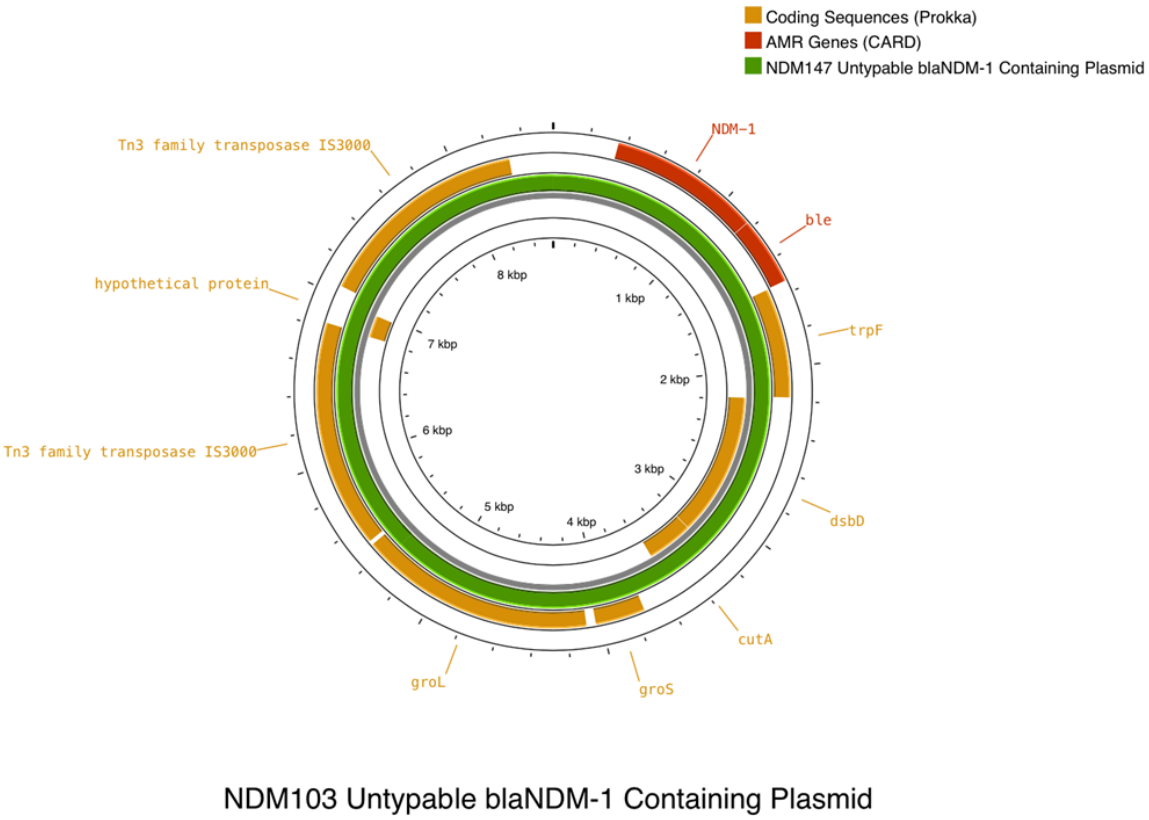

Figure 9: Comparison of two isolates (NDM103, *En. hormaechei*; NDM147, *K. pneumoniae*) with a smaller circularised element encoding for *bla*$_{NDM-}$1. AMR genes were annotated on the reference (NDM103) genome using the CARD AMR tool within Proksee and the coding sequences were annotated using the Prokka tool within Proksee, which the untypable plasmid from NDM147 was then compared against in a blastn search within Proksee.

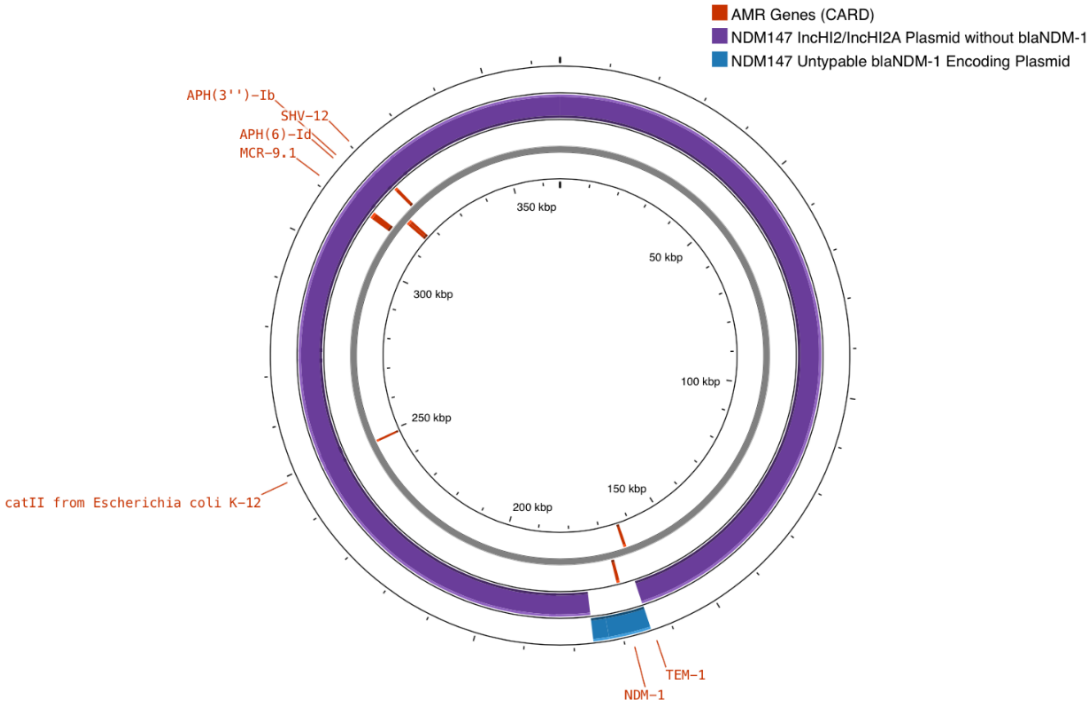

Figure 10: Comparison of the IncHI2/IncHI2A plasmid and separate untypable *bla*$_{NDM-1}$ encoding element of NDM147 (ST101, *K. pneumoniae*) to the *bla*$_{NDM-1}$ encoding IncHI2/IncHI2A plasmid of NDM146 (ST101, *K. pneumoniae*). AMR genes were annotated on the reference (NDM146) genome using the CARD AMR tool within proksee, the IncHI2/IncHI2A plasmid and *bla*$_{NDM-1}$ encoding element from NDM147 were then compared against in a blastn search within Proksee.

bla$_{NDM-5}$ is encoded on different plasmid replicons with an identical gene cassette.

All isolates with *bla*$_{NDM-5}$ had only a single copy of the gene encoded on a plasmid. The plasmids encoding for the *bla*$_{NDM-5}$ gene varied in the isolates long-read sequenced, the gene was identified on a IncR plasmid (NDM112); an IncFIA/IncFII multi-replicon plasmid (NDM150); an IncFIB(pNDM-Mar)/IncFII(29)/IncHI1B(pNDM-MAR) multi-replicon plasmid (NDM162) and a IncFIA/IncFIB/Col156 multi-replicon plasmid (NDM165) (Table S5). Based on assembler outputs, all contigs with *bla*$_{NDM-5}$ were circularised. A gene cluster comparison of the genes surrounding *bla*$_{NDM-5}$ in all isolates is visualised in Figure 11, emphasizing that despite the variable plasmid background, there is a conserved gene cassette of *ble*, *trpF* and *dsbD* that seems to translocate together with *bla*$_{NDM-5}$.

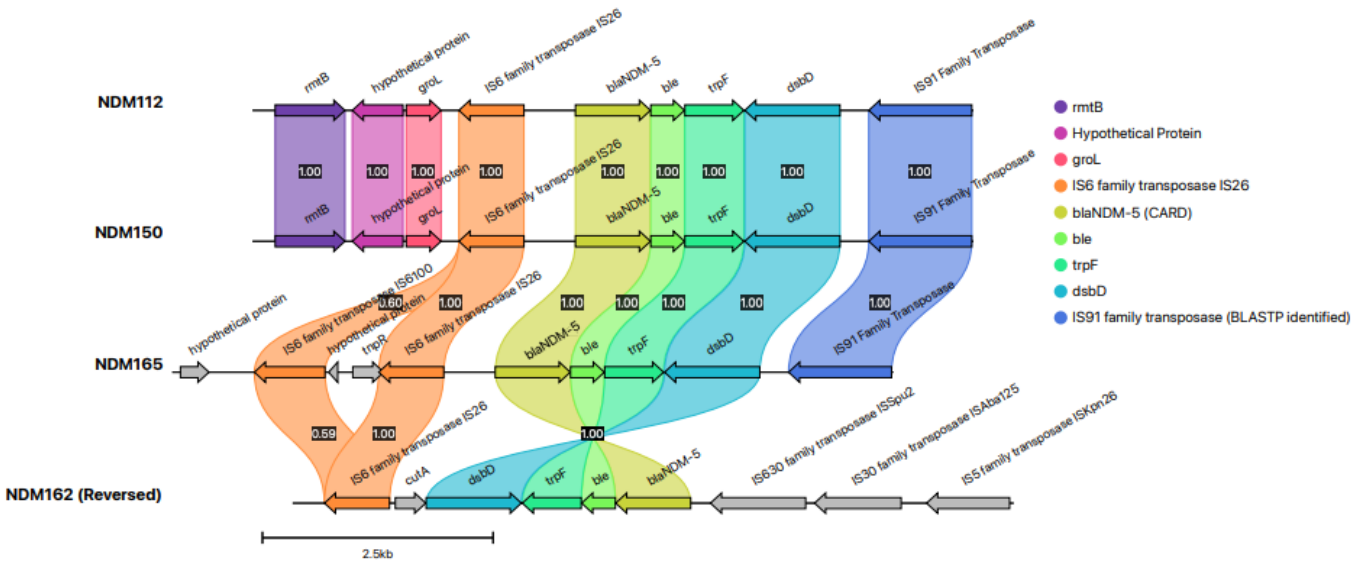

Figure 11: Comparison of the genomic context including flanking region 3500 bp upstream and 3500 bp downstream of the *bla*$_{NDM-5}$ gene in NDM112 (*En. hormaechei*), NDM150 (*E. coli*), NDM165 (*E. coli*), NDM162 *(K. pneumoniae*). Upstream and downstream of NDM162 has been reversed for ease of viewing. Numbers between the same gene in different isolate represents the similarity on a score of 0.00-1.00, where 0.00 is no similarity and 1.00 is identical. Coding sequences were annotated according to Prokka, apart from where it is stated that BLASTP was used due to no identification using Prokka.

Discussion

In 2023, it was noted that in seven Merseyside hospitals and a GP centre there was an accumulation of NDM-positive isolates. Subsequent combination of clinical and genome epidemiology revealed a combination of small clonal outbreaks, most notably eighteen closely related K. pneumoniae ST101 isolates, and transmission of plasmids as well as smaller mobile elements, though we cannot distinguish within-hospital evolution of plasmids from circulation of different variants in the community.

Using WGS, including the generation of several reference genomes for our study, we have been able to identity a successful multi-replicon IncHI2/IncHI2A plasmid encoding for the *bla*$_{NDM-1}$ gene, within these seven Merseyside hospitals and a GP practice community isolate, causing the majority of observed cases. The phylogenies and STs identified in all species emphasise the diversity of the isolates involved, further indicating the importance of MGEs in the spread of the *bla*$_{NDM}$ gene.

IncHI2/IncHI2A plasmids were observed in clinically retrieved isolates from six hospitals across England in 2021 – 2023, encoding for multiple carbapenemase genes (*bla*$_{NDM-1}$, *bla*$_{KPC-2}$, *bla*$_{IMP-1}$, and *bla*$_{IMP-4}$ genes) (49). Here we observed the *bla*$_{NDM-1}$ gene alongside multiple other AMR genes located on the IncHI2/IncHI2A plasmid. In one isolate, NDM121, two copies of the *bla*$_{NDM-1}$ gene were seen in a tandem repeat on the IncHI2/IncHI2A

plasmid. This work highlights the success of *bla*$_{NDM-1}$ associated with a single IncHI2/IncHI2A plasmid backbone in multiple bacterial species after being identified multiple Merseyside hospitals. However, further efforts are required to better understand the successful persistence and spread of this plasmid to continue to be identified in both clinical and screening samples across 2023. Due to a lack of environmental isolates collected throughout 2023 the prevalence of this plasmid within the hospital environment is unknown. The continued prevalence of this plasmid within patient samples over 2023, indicates that the efficacy of current IPC measures, should be evaluated to ensure that persistence of this plasmid in healthcare settings cannot occur.

In contrast, in a subset of 4/12 of the isolates where the *bla*$_{NDM-5}$ gene present, it was identified that the *bla*$_{NDM-5}$ gene was encoded by a diverse variety of smaller IncF and IncR plasmids. *bla*$_{NDM-5}$ isolates were derived from three species in five hospitals and a community GP sample: with all samples being identified between January and December 2023. Thus, it is far more likely to represent community-acquired rather than within-hospital transmission. Nearly identical flanking regions surrounding the *bla*$_{NDM-5}$ including transposable elements were noted, indicating a conserved transposable unit. Across England, in hospital settings the *bla*$_{NDM-5}$ and *rmtB* genes have been seen on IncFIA/IncFII multi-replicon plasmids in *E. coli* isolates, as in our study (50). However, the flanking region surrounding the *bla*$_{NDM-5}$ gene differed, with an IS30 transposon seen immediately upstream from the *bla*$_{NDM-5}$ gene rather than the IS26 transposon observed in NDM150 (50). Emphasising the diversity of *bla*$_{NDM-5}$ encoding plasmids circulating within Merseyside and England in recent years.

In two isolates, the final genome assemblies indicate that the *bla*$_{NDM-1}$ gene is encoded on a small circular element that could not be classified following incompatibility method plasmid typing. The Tn3 family of transposons, located upstream and downstream of the *bla*$_{NDM-1}$ gene in these isolates, has been noted to move in and out of circular DNA such as plasmids or the chromosome (51). Further study is required to better understand the element the *bla*$_{NDM-1}$ gene is encoded on, how this could affect the transmission of the *bla*$_{NDM-1}$ gene in these cases, or if the isolates sequenced here are beginning to lose the *bla*$_{NDM-1}$ gene. It has previously been observed that *bla*$_{NDM}$ can get lost during passage; particularly without antibiotic selection (52). Before long-read sequencing, these isolates were not cultured under antibiotic pressure, and this could have caused the results seen here. Additionally, this could have occurred, in the six isolates who were identified to have no *bla*$_{NDM}$ gene after being kept in storage before the study began.

Overall, this work highlights how crucial long-read sequencing data was to fully resolve the location of *bla*$_{NDM}$ genes in a selection of our isolates. Without this data, the identification of a

multi-species *bla*$_{NDM-1}$ encoding plasmid across multiple hospitals and potential transposition within *bla*$_{NDM-5}$ encoding isolates would have been impossible. The use of long-read sequencing within hospital has previously been used to identify the spread of plasmids of concern in clinical isolates, allowing IPC interventions to occur quickly, however the issues of cost and time required to complete this analysis have been highlighted too (53, 54). Nonetheless the benefit of this technology combined with clinical data (while retrospective in nature) is further highlighted here and should be recommended to be implemented, where possible, when carbapenemase genes of concern such as NDM are first identified using current hospital molecular identification techniques.

In recent years, *K. pneumoniae* ST101 and ST147 have emerged as a potential high-risk CPE clone in European countries and have been identified here (55, 56). All *K. pneumoniae* isolates within this study were from screening samples across three hospitals. In a singular ST147 isolate that contained the *bla*$_{NDM-14}$ variant, encoded only a partial *rmpA2* gene, a regulator of hypermucoidity which can indicate hypervirulence, in line with all samples being negative in the string test (57).

Within this collection of NDM producing *K. pneumoniae* ST101 isolates, two clonal sub-lineages represented 15 of the ST101 isolates, using a sublineage cutoff of five SNPs which is a community standard to define likely recent transmission events (58). Clonal outbreaks linked to within-hospital transmission is supported by being all isolates from sublineage 1 of ST101 being isolated from the same hospital (Hospital 1) in the span of three months. However, for sublineage 2 of ST101 isolates were identified within three hospitals across six months; here further patient data is required to better understand the likelihood of patient movement between Merseyside hospitals within 2023. Increased IPC practices at high-risk sites can break hospital transmission chains, which can be derived from shared equipment as recently shown in a similar setting for *K. pneumoniae* (59).

This work had some limitations. While it is known that these isolates were taken from patients or the environment from one of the study hospitals, further information regarding the patient's stay regarding the duration and locations visited within the hospital site was unavailable for this study. This patient data would allow the identification of patterns of species, plasmids or *bla*$_{NDM}$ variants within a certain location and could begin to allow for inference regarding potential hospital verses community transmission occurring within Merseyside in 2023. Long read sequencing to fully resolve plasmids was carried out on a selection of isolates rather than the entire isolate set, which would have provided a more complete analysis of the *bla*$_{NDM}$ associated plasmids and insertion sequences.

Conclusion

There are multiple variants of the *bla*$_{NDM}$ gene circulating within Merseyside, primarily *bla*$_{NDM-1}$ on a IncHI2/IncHI2A plasmid. This is occurring within multiple species and was identified in both screening and clinical samples from patients. The *bla*$_{NDM-1}$ encoding IncHI2/IncHI2A plasmid was driven majorly by a single ST in *K. pneumoniae*, but a variety of different STs in *E. coli* and in *En. hormaechei*, highlighting the importance of MGEs for the *bla*$_{NDM}$ gene to establish a presence in different genomic contexts. The local emergence of *bla*$_{NDM-1}$, within 2023, emphasises the need for new enhanced surveillance and potential infection prevention measures to begin to be adapted for use within healthcare settings to better equip hospital staff with the knowledge required to effectively tackle outbreaks involving this gene.

Competing interest statement

None to declare

Author contributions following CRediT taxonomy

Conceptualization: CB, TE, EH. Data curation: CD, DL, JML. Formal analysis: CD, AJF. Funding acquisition: TE, EH. Investigation: CD, DL, VO. Methodology: AJF, FEG, DC, JML, TE, EH. Project administration: TE, EH. Software: CD, EH. Resources: CB, DL, VO, TN, JC, TE. Supervision: AJF, FEG, TE, EH. Validation: DL, VO. Visualization: CD. Writing – original draft: CD, JML, TE, EH. Writing – review & editing: CD, CB, DL, VO, AJF, FEG, DC, JML, TE, EH. All authors read and approved the final version of the manuscript.

Funding statement

EH acknowledges funding from Wellcome (217303/Z/19/Z) and the BBSRC (BB/V011278/1, BB/V011278/2). TE acknowledges funding from an Academy of Medical Sciences Springboard award (REF: SBF009\1181).

References

1. Murray, C.J., et al., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The Lancet, 2022. **399**(10325): p. 629-655.
2. Moellering Jr, R.C., G.M. Eliopoulos, and D.E. Sentochnik, *The carbapenems: new broad spectrum β-lactam antibiotics.* Journal of Antimicrobial Chemotherapy, 1989. **24**(suppl_A): p. 1-7.
3. Fröhlich, C., et al., *Evolution of β-lactamase-mediated cefiderocol resistance.* Journal of Antimicrobial Chemotherapy, 2022. **77**(9): p. 2429-2436.

4. Schwaber, M.J., et al., *High levels of antimicrobial coresistance among extended-spectrum-β-lactamase-producing Enterobacteriaceae.* Antimicrobial Agents and Chemotherapy, 2005. **49**(5): p. 2137-2139.
5. WHO, *WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance*. 2024.
6. Seo, H., et al., *Comparison of clinical outcomes of patients infected with KPC-and NDM-producing Enterobacterales: a retrospective cohort study.* Clinical Microbiology and Infection, 2021. **27**(8): p. 1167. e1-1167. e8.
7. Lin, Q., et al., *Clinical and microbiological characterization of carbapenem-resistant enterobacteriales: a prospective cohort study.* Frontiers in Pharmacology, 2021. **12**: p. 716324.
8. UKHSA. *Carbapenemase-producing Gram-negative organisms in England since October 2020: quarterly update, Q3 2024*.
9. UKHSA. *Quarterly laboratory surveillance of acquired carbapenemase-producing Gram-negative organisms in England: October 2020 to December 2024 web appendix*. 2025; Available from: https://assets.publishing.service.gov.uk/media/67ea87649110281dc1056257/CPO_quarterly_data_2024_Q4.xlsx.
10. Walsh, T.R., et al., *Metallo-β-lactamases: the quiet before the storm?* Clinical Microbiology Reviews, 2005. **18**(2): p. 306-325.
11. Yong, D., et al., *Characterization of a new metallo-β-lactamase gene, bla NDM-1, and a novel erythromycin esterase gene carried on a unique genetic structure in Klebsiella pneumoniae sequence type 14 from India.* Antimicrobial Agents and Chemotherapy, 2009. **53**(12): p. 5046-5054.
12. Lawrie, D., *Email to Caitlin Duggan*, C. Duggan, Editor. 2024.
13. UKHSA. *Framework of actions to contain carbapenemase-producing Enterobacterales*. 2022; Available from: https://assets.publishing.service.gov.uk/media/63346c44d3bf7f34f1bc882d/Framework_of_actions_to_contain_CPE.pdf.
14. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-2120.
15. Andrews, S. *FastQC: a quality control tool for high throughput sequence data*. 2015; Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
16. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-3048.
17. Parks, D.H., et al., *CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.* Genome Res, 2015. **25**(7): p. 1043-55.
18. Simão, F.A., et al., *BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.* Bioinformatics, 2015. **31**(19): p. 3210-3212.
19. Carver, T.J., et al., *ACT: the Artemis comparison tool.* Bioinformatics, 2005. **21**(16): p. 3422-3423.
20. Lumpe, J., et al., *GAMBIT (Genomic Approximation Method for Bacterial Identification and Tracking): A methodology to rapidly leverage whole genome sequencing of bacterial isolates for clinical identification.* PloS One, 2023. **18**(2): p. e0277575.
21. Bouras, G., et al., *Hybracter: enabling scalable, automated, complete and accurate bacterial genome assemblies.* Microbial Genomics, 2024. **10**(5): p. 001244.
22. Kolmogorov, M., et al., *Assembly of long, error-prone reads using repeat graphs.* Nature Biotechnology, 2019. **37**(5): p. 540-546.
23. Wick, R.R., et al., *Unicycler: resolving bacterial genome assemblies from short and long sequencing reads.* PLoS Computational Biology, 2017. **13**(6): p. e1005595.
24. Wick, R.R. and K.E. Holt, *Polypolish: short-read polishing of long-read bacterial genome assemblies.* PLoS Computational Biology, 2022. **18**(1): p. e1009802.

25. Bouras, G., et al., *How low can you go? Short-read polishing of Oxford Nanopore bacterial genome assemblies.* bioRxiv, 2024: p. 2024.03. 07.584013.
26. Bharat, A., et al., *Correlation between phenotypic and in silico detection of antimicrobial resistance in Salmonella enterica in Canada using Staramr.* Microorganisms, 2022. **10**(2): p. 292.
27. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-2069.
28. Ondov, B.D., et al., *Mash: fast genome and metagenome distance estimation using MinHash.* Genome Biology, 2016. **17**: p. 1-14.
29. Hunt, M., et al., *ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads.* Microbial Genomics, 2017. **3**(10): p. e000131.
30. Wirth, T., et al., *Sex and virulence in Escherichia coli: an evolutionary perspective.* Molecular microbiology, 2006. **60**(5): p. 1136-1151.
31. Tonkin-Hill, G., et al., *Producing polished prokaryotic pangenomes with the Panaroo pipeline.* Genome Biology, 2020. **21**: p. 1-21.
32. Page, A.J., et al., *SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments.* Microbial Genomics, 2016. **2**(4): p. e000056.
33. Nguyen, L.-T., et al., *IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.* Molecular Biology and Evolution, 2015. **32**(1): p. 268-274.
34. Wailan, A.M., et al., *rPinecone: Define sub-lineages of a clonal expansion via a phylogenetic tree.* Microbial Genomics, 2019.
35. Croucher, N.J., et al., *Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins.* Nucleic Acids Research, 2015. **43**(3): p. e15-e15.
36. Kozlov, A.M., et al., *RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference.* Bioinformatics, 2019. **35**(21): p. 4453-4455.
37. Yu, G., et al., *ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data.* Methods in Ecology and Evolution, 2017. **8**(1): p. 28-36.
38. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2016, New York: Springer-Verlag.
39. Pedersen, T.L., *patchwork: The Composer of Plots*. 2025. p. R package version 1.3.0.9000, https://github.com/thomasp85/patchwork.
40. Inouye, M., et al., *SRST2: Rapid genomic surveillance for public health and hospital microbiology labs.* Genome medicine, 2014. **6**: p. 1-16.
41. Lam, M.M.C., et al., *A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex.* Nature Communications, 2021. **12**(1): p. 4188.
42. Graf, F.E., et al., *Molecular mechanisms of re-emerging chloramphenicol susceptibility in extended-spectrum beta-lactamase-producing Enterobacterales.* Nature Communications, 2024. **15**(1): p. 9019.
43. Carattoli, A., et al., *In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing.* Antimicrobial agents and chemotherapy, 2014. **58**(7): p. 3895-3903.
44. Wick, R.R., et al., *Bandage: interactive visualization of de novo genome assemblies.* Bioinformatics, 2015. **31**(20): p. 3350-3352.
45. Grant, J.R., et al., *Proksee: in-depth characterization and visualization of bacterial genomes.* Nucleic Acids Research, 2023. **51**(W1): p. W484-W492.
46. Carver, T., et al., *Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data.* Bioinformatics, 2012. **28**(4): p. 464-469.
47. Gilchrist, C.L. and Y.-H. Chooi, *Clinker & clustermap. js: Automatic generation of gene cluster comparison figures.* Bioinformatics, 2021. **37**(16): p. 2473-2475.

48. Shon, A.S., R.P. Bajwa, and T.A. Russo, *Hypervirulent (hypermucoviscous) Klebsiella pneumoniae: a new and dangerous breed.* Virulence, 2013. **4**(2): p. 107-118.
49. Turton, J.F. and J.A. Turton, *Long read sequencing provides an insight into plasmids found among carbapenemase producing Enterobacterales from hospitals in the United Kingdom during 2021 to 2023.* bioRxiv, 2024: p. 2024.03. 19.585710.
50. Turton, J.F., et al., *Wide distribution of Escherichia coli carrying IncF plasmids containing bla NDM-5 and rmtB resistance genes from hospitalized patients in England.* Journal of Medical Microbiology, 2022. **71**(8): p. 001569.
51. Tang, Y., et al., *Transposition with Tn 3-family elements occurs through interaction with the host β-sliding clamp processivity factor.* Nucleic Acids Research, 2024: p. gkae674.
52. Li, X., et al., *Genetic characterization and passage instability of a hybrid plasmid co-harboring bla IMP-4 and bla NDM-1 reveal the contribution of insertion sequences during plasmid formation and evolution.* Microbiology spectrum, 2021. **9**(3): p. e01577-21.
53. Wu, C.-T., et al., *Rapid whole genome characterization of antimicrobial-resistant pathogens using long-read sequencing to identify potential healthcare transmission.* Infection Control & Hospital Epidemiology, 2025. **46**(2): p. 129-135.
54. Kwong, J.C., et al., *Translating genomics into practice for real-time surveillance and response to carbapenemase-producing Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak.* PeerJ, 2018. **6**: p. e4210.
55. Rubic, Z., et al., *Molecular characterization of colistin resistance genes in a high-risk ST101/KPC-2 clone of Klebsiella pneumoniae in a University Hospital of Split, Croatia.* International Microbiology 2023. **26**(3): p. 631-637.
56. Peirano, G., et al., *Emerging antimicrobial-resistant high-risk Klebsiella pneumoniae clones ST307 and ST147.* Antimicrobial Agents and Chemotherapy, 2020. **64**(10): p. 10.1128/aac. 01148-20.
57. Russo, T.A., et al., *Identification of biomarkers for differentiation of hypervirulent Klebsiella pneumoniae from classical K. pneumoniae.* Journal of clinical microbiology, 2018. **56**(9): p. 10.1128/jcm. 00776-18.
58. Heinz, E., et al., *Longitudinal analysis within one hospital in sub-Saharan Africa over 20 years reveals repeated replacements of dominant clones of Klebsiella pneumoniae and stresses the importance to include temporal patterns for vaccine design considerations.* Genome Medicine, 2024. **16**(1): p. 67.
59. Fraser, A.J., et al., *Genomic insights into an outbreak of SHV-98, TEM-1 and DHA-1 producing Klebsiella pneumoniae in a paediatric cardiac unit.* medRxiv, 2025: p. 2025.02. 14.25322074.