# Comprehensive analysis of horizontal gene transfer among multidrug-resistant bacterial pathogens in a single hospital

Daniel R. Evans[1,2], Marissa P. Griffith[3], Mustapha M. Mustapha[3], Jane W. Marsh[3], Alexander J. Sundermann[3], Vaughn S. Cooper[4], Lee H. Harrison[3], Daria Van Tyne[1]

[1]Division of Infectious Diseases, University of Pittsburgh School of Medicine
[2]Department of Infectious Diseases and Microbiology, University of Pittsburgh Graduate School of Public Health
[3]Microbial Genomic Epidemiology Laboratory, Infectious Diseases Epidemiology Research Unit, University of Pittsburgh School of Medicine and Graduate School of Public Health
[4]Department of Microbiology and Molecular Genetics, and Center for Evolutionary Biology and Medicine, University of Pittsburgh

## ABSTRACT

Multidrug-resistant bacterial pathogens pose a serious public health threat, especially in hospital settings. Horizontal gene transfer (HGT) of mobile genetic elements (MGEs) contributes to this threat by facilitating the rapid spread of genes conferring antibiotic resistance, enhanced virulence, and environmental persistence between nosocomial pathogens. Despite recent advances in microbial genomics, studies of HGT in hospital settings remain limited in scope. The objective of this study was to identify and track the movement of MGEs within a single hospital system using unbiased methods. We screened the genomes of 2,173 bacterial isolates from healthcare-associated infections collected over an 18-month time period to identify nucleotide regions that were identical in the genomes of bacteria belonging to distinct genera. These putative MGEs were found in 196 isolates belonging to 11 different genera; they grouped into 51 clusters of related elements, and they were most often shared between related genera.

To resolve the genomic locations of the most prevalent MGEs, we performed long-read sequencing on a subset of representative isolates and generated highly contiguous, hybrid-assembled genomes. Many of these genomes contained plasmids and chromosomal elements encoding one or more of the MGEs we identified, which were often arranged in a mosaic fashion. We then tracked the appearance of ten MGE-bearing plasmids in all 2,173 genomes, and found evidence supporting the transfer of plasmids between patients independent from bacterial transmission. Finally, we identified two instances of likely plasmid transfer across genera within individual patients. In one instance, the plasmid appeared to have subsequently transferred to a second patient. By surveying a large number of bacterial genomes sampled from infections at a single hospital in a systematic and unbiased manner, we were able to track the independent transfer of MGEs over time. This work expands our understanding of HGT in healthcare settings, and can inform efforts to limit the spread of drug-resistant pathogens in hospitals.

## INTRODUCTION

Horizontal gene transfer (HGT) is a driving force behind the multidrug-resistance and heightened virulence of healthcare-associated bacterial infections[1]. Genes conferring antibiotic resistance, heightened virulence, and environmental persistence are often encoded on mobile genetic elements (MGEs), which can be readily shared between bacterial pathogens via HGT[2]. While rates of HGT are not well quantified in clinical settings, prior studies have shown that MGEs can mediate and/or exacerbate nosocomial outbreaks[3–6]. Recent studies have also demonstrated that multidrug-resistant healthcare-associated bacteria share MGEs across large phylogenetic distances[7–9]. Understanding the dynamics of MGE transfer in clinical settings can uncover important epidemiologic links that are not currently identified by traditional infection control methodologies[1,10,11].

Methods to identify and track the movement of MGEs among bacterial populations on short timescales are limited. Bacterial whole-genome sequencing has transformed infectious disease epidemiology over the last decade[12], providing powerful new tools to identify and intervene against outbreaks[13]. Despite these advances, efforts to track MGE movement have focused almost exclusively on drug resistance and virulence genes[5,7,11,14], often ignoring the broader genomic context of the mobile elements themselves. Many studies rely on the identification of plasmid replicons, transposases, and other "marker genes",[15] an approach that oversimplifies the diversity of MGEs and may lead to incomplete or erroneous conclusions about their epidemiology. While querying databases containing curated MGE-associated sequences is useful for the rapid screening of clinical isolates for known MGEs, it will not capture novel MGEs. Additionally, whole-genome sequencing using short-read technologies generates genome assemblies that usually do not resolve MGE sequences, due to the abundance of repetitive elements that MGEs often contain[16]. Advances in long-read sequencing can mitigate this problem; the combination of short- and long-read sequence data can allow the genomic context of chromosomal and extrachromosomal MGEs to be precisely visualized[7,17,18]. Finally, studying the epidemiology of MGEs in clinical settings requires detailed individual-level patient clinical data, without which HGT occurrence in the hospital cannot be identified[18].

Here we performed an alignment-based screen for MGEs in a large and diverse collection of bacterial genomes sampled from infections within a single hospital over an 18-month time period. We identified complete and fragmented MGEs that were identical in nucleotide sequence and occurred in the genomes of bacteria belonging to different genera. Because they are identical, we suspect that these MGEs have recently transferred between bacteria within the hospital setting. Further analysis using long-read sequencing and referenced-based resolution of distinct MGEs enabled us to precisely characterize MGE architecture and cargo, and to track MGE occurrence over time. Cross-referencing our results with available patient metadata

allowed us to follow these elements as they emerged and were maintained among nosocomial bacterial populations.

## RESULTS

### Identification of MGEs shared across bacterial genera in a single hospital

Our experimental workflow is depicted in Fig. 1A. To identify genetic material being shared between distantly related bacteria in the hospital setting, we screened a dataset containing 2,173 whole-genome sequences of clinical isolates of high-priority Gram-positive and Gram-negative bacteria collected from a single hospital over an 18-month period as part of the Enhanced Detection System for Hospital-Acquired Transmission (EDS-HAT) project at the University of Pittsburgh[19] (Methods). To have maximal contrast in our identification of MGEs, we focused on identical sequences found in the genomes of bacteria belonging to different genera. We performed an all-by-all alignment of the 2,173 genomes in the dataset using nucmer[20], and filtered the results to retain alignments of at least 5kb that shared 100% identity between bacteria of different genera. The resulting sequences were extracted and clustered using Cytoscape (Fig. 1B). This approach identified putative MGE sequences in 196 genomes belonging to 11 genera, which could be grouped into 51 clusters of related MGEs. These MGE clusters ranged in size from two to 52 genomes, and comprised two, three, or four different genera (Fig. 1B). MGE sequences were found predominantly among Gram-negative *Enterobacteriaceae*, particularly *Klebsiella spp.*, *Escherichia coli*, and *Citrobacter spp.* (Fig. 1C). Annotation of clustered sequences confirmed that more than 80% of the MGE clusters encoded one or more genes involved in DNA mobilization, plasmid replication, or another mobile function presumably involved in HGT (Fig. 1D). Somewhat surprisingly, only about one-quarter of the MGE clusters contained antimicrobial resistance genes, including genes encoding resistance to aminoglycosides, antifolates, beta-lactams, macrolides, quinolones, sulphonamides, and tetracyclines (Fig. 1D, 1E). Finally, 8/51 MGE clusters encoded genes and operons whose

105 products were predicted to interact with metals, including arsenic, copper, mercury, nickel, and
106 silver (Fig. 1D). Collectively, these results indicate that our unbiased, alignment-based method
107 successfully identified putative MGEs, particularly in pathogens known to engage in HGT[2,21].
108
109 To assess the phylogenetic distribution of the putative MGEs we identified, we constructed a
110 core gene phylogeny of the 196 genomes encoding one or more MGE clusters using the
111 Genome Taxonomy Database Tool Kit (GTDBTK)[22] (Fig. 2). MGE clusters were often found
112 among bacteria in related genera, in particular the *Enterobacteriaceae.* We did not observe any
113 MGEs that were present in both Gram-positive and Gram-negative isolate genomes, but we did
114 find MGEs in the genomes of distantly related bacteria. For example, we identified an MGE
115 carrying three aminoglycoside resistance genes that was identical in sequence between a
116 vancomycin resistance-encoding plasmid carried by *E. faecium* and the *C. difficile* chromosome
117 (MGE cluster C9, Fig. 3A). The *C. difficile* strain carrying this element was previously found to
118 also harbor an *npmA* aminoglycoside resistance gene[23]. We also found portions of an
119 integrative conjugative element that were identical between two *P. aeruginosa* isolates and a S.
120 *marcescens* isolate (MGE cluster C30, Fig. 3B). Identical regions of this element included
121 formaldehyde resistance genes and Uvr endonucleases. Finally, we detected complete and
122 identical Tn7 transposons in the genomes of *A. baumannii, E. coli*, and *P. mirabilis* isolates
123 (MGE cluster C17, Fig. 3C). The Tn7 sequence we detected was also identical to the Tn7
124 sequence of pR721, an *E. coli* plasmid that was first described in 1990 and was sequenced in
125 2014[24]. Taken together, these results indicate that while many of the sequences we identified
126 were from MGEs shared between related bacterial genera, our approach also identified partial
127 or complete MGEs that were identical in the genomes of distantly related pathogens.
128
129 **MGEs often reside on larger elements in different combinations and contexts**

To further investigate the genomic context of the MGEs identified, we selected representative isolates from the largest MGE clusters for long-read sequencing using Oxford Nanopore technology. Hybrid assembly using short Illumina reads and long Nanopore reads generated highly contiguous chromosomal and plasmid sequences, which allowed us to resolve larger elements carrying one or more of the most prevalent MGE clusters (Table 1). We found that several of the smaller and more prevalent MGEs were carried on a variety of different plasmid and chromosomal elements, which we designated as "MGE lineages" (Table 1, Fig. 4A). These smaller MGEs co-occurred in different orders, orientations, and combinations on the larger elements. This kind of "nesting" of MGEs within larger mobile elements has been previously observed[6], and our findings further support the mosaic, mix-and-match nature of the smaller MGEs we identified. We also confirmed that these MGEs were truly mobile, since they appeared to be able to move independently between multiple distinct larger mobile elements. A closer examination of the three largest MGE clusters (C1, C2, C3) showed that C1 sequences did not all share a common "core" nucleotide sequence, but rather could be aligned in a pairwise fashion to generate a contiguous sequence (Fig. 4B). MGE clusters C2 and C3, on the other hand, did contain "core" sequences that were present in all genomes carrying the MGE (Fig. 4C, 4D).

**Plasmids carrying MGE clusters are found in multiple sequence types, species, and genera circulating in the same hospital**

More than half (104/196) of the MGE-carrying genomes in our dataset contained one or more of the five most prevalent MGEs we identified (C1-C5, Fig. 1B). All five MGEs were small (usually less than 10kb), and were predicted to be carried on plasmids shared between *Enterobacteriaceae*. We set out to resolve the genomic context of each of these five MGEs in all isolates containing them. We used an iterative approach involving long-read sequencing and hybrid assembly of representative isolates to generate reference sequences of MGE-containing

elements (chromosomal or plasmid), followed by mapping of contigs from Illumina-only assemblies to these reference sequences to assess their coverage in every genome (Methods). This approach allowed us to query the presence of plasmids and chromosomal elements from genomes sequenced with Ilumina technology alone, without requiring long-read sequencing of all isolates or relying on external reference sequences. We found that 11 of the 104 isolates (all *E. coli*) carried cluster C1 and C3 MGEs on their chromosome, while the remaining 93 isolates carried clustered MGEs on 17 distinct plasmids. Seven of these plasmids were present in only one isolate in the dataset, but 10 plasmids appeared to be shared between more than one isolate (Table 1, Fig. 5). We also conducted the same reference-based coverage analysis for all 2,173 genomes in the original dataset, and identified an additional 16 isolates with >90% coverage of an MGE lineage.

While all of the MGEs we originally identified were present in the genomes of bacteria belonging to different genera, the plasmids that we resolved were variable in how widely they were shared. For example, some plasmids were only found among isolates belonging to a single species and multilocus sequence type (ST), suggesting that they were likely transmitted between patients along with the bacteria that were carrying them (Fig. 5A). These included a *blaKPC-3* carbapenemase-encoding plasmid (pKLP00149_2) found in *K. pneumoniae* isolates belonging to ST258, a multidrug-resistant and highly virulent hospital-adapted bacterial lineage that has recently undergone clonal expansion in our hospital[18]. We also found a *blaOXA-1* extended spectrum beta-lactamase-encoding plasmid in *E. coli* isolates belonging to ST131, another multidrug-resistant and hypervirulent bacterial lineage[25]. In addition to plasmids that occurred in bacteria belonging to the same ST, we also identified plasmids that were present in isolates belonging to different STs of the same species, or in different species of the same genus (Fig. 5B). All isolates in this case were *K. pneumoniae* or *K. oxytoca*, suggesting widespread sharing of plasmids between distinct *Klebsiella* species and STs. The plasmids often carried antibiotic

resistance genes, and many also carried metal interaction genes (Table 1). Finally, we identified three different plasmids that were shared between different bacterial genera all belonging to the *Enterobacteriaceae* (Fig. 5C). One small plasmid (pKLP00155_6) carrying the colicin bacterial toxin was found in 26 isolates belonging to 10 different STs and four different genera. Taken together, these results indicate that some plasmids carrying putative MGEs were likely inherited vertically as bacteria were transmitted between patients in the hospital, while others appear to have transferred independently of bacterial transmission.

**Likely HGT across genera within individual patients**

By cross-referencing the isolates containing MGE sequences with de-identified patient data, we found two instances where identical MGEs were found in pairs of isolates of different genera that were collected from the same patient, on the same date, and from the same sample source. To resolve the complete MGE profiles of these cases, we performed long-read sequencing and hybrid assembly on all genomes involved (Fig. 6). A *K. pneumoniae* ST405 isolate (KLP00215) and an *E. coli* ST69 isolate (EC00678) collected from a tissue infection from Patient A each harbored a 113.6kb IncFIB(pQil)/IncFII(K) plasmid carrying *blaKPC*, *blaTEM*, and *blaOXA* enzymes, as well as a mercury detoxification operon (Fig. 6A, B). In addition, an isolate from a second patient (Patient B, EC00701, *E. coli* ST131), also encoded a nearly identical plasmid. A systematic chart review for Patients A and B revealed that they occupied adjacent hospital rooms for four days during a time period after Patient A's isolates were collected but before Patient B's isolate was collected. During this time the two patients would have shared the same healthcare staff, who might have transferred bacteria between them.

In the second case of putative within-patient HGT, a *K. pneumoniae* ST231 isolate (KLP00187) and a *C. braakii* ST356 isolate (CB00017) were both collected from the same urine sample of Patient C (Fig. 6B). Both isolates carried nearly identical 196.8kb IncFIB(K)/IncFII(K) plasmids

conferring resistance to aminoglycosides, beta-lactams, chloramphenicol, fluoroquinolones, sulfonamides, tetracyclines, and trimethoprim, as well as operons encoding copper and arsenic resistance. In addition, isolates from two subsequent patients (Patient D and Patient E) also carried plasmids belonging to the same lineage as the plasmid shared between KLP00187 and CB00017. Alignment of the sequences of all four plasmids showed that the plasmids isolated from Patient C were nearly identical, while the plasmids from Patients D and E had small differences in their gene content and organization (Fig. 6B). Systematic chart review did not identify any strong epidemiologic links between the three patients, suggesting that this plasmid was not passed directly between these patients and might instead have transferred via additional bacterial isolates or populations that were not sampled.

## DISCUSSION

In this study, we identified MGEs in a large dataset of whole-genome sequences of clinical bacterial isolates collected over an 18-month period from a single hospital. We identified, clustered, and characterized identical sequences found in multiple distinct genera, and in the process uncovered both expected and unexpected cases of MGE occurrence. We confirmed that some of the most common MGEs identified were fragments of larger mobile elements. We performed long-read sequencing to resolve these larger elements, which were almost always plasmids. When we traced the presence of various plasmid lineages over time, we found some that were likely transmitted vertically along with the bacteria carrying them, and others that appeared to be transferred horizontally between unrelated bacteria.

Our study adds to the body of knowledge of HGT in hospital settings in new and important ways. We analyzed a large dataset of clinical isolates collected from a single health system, and used a systematic and unbiased approach to identify MGEs regardless of their type or gene content. While prior studies have used genomic epidemiology to study how HGT contributes to

234 the transmission, persistence, and virulence of bacterial pathogens[4,5,19,20], the technical
235 challenges of resolving MGEs from whole-genome sequencing data have limited the scope of
236 these findings[16]. Other studies have deliberately tracked HGT in healthcare settings by focusing
237 either on mobile genes of interest, such as those encoding drug resistance[7,9,14], or on specific
238 classes of MGEs[28]. Both of these approaches can generate biased interpretations of the driving
239 forces behind HGT in clinical settings. For this reason we selected a pairwise alignment-based
240 approach, whereby we only looked for identical sequences in the genomes of very distantly
241 related bacteria. In doing so, we did not limit ourselves to only looking for "known" MGEs, and
242 thus obtained a more accurate and comprehensive overview of the dynamics of HGT between
243 bacterial genera in our hospital.
244
245 What might cause horizontally-transferred nucleotide sequences to be found at very high
246 identity within phylogenetically distinct bacteria? We predicted that there might be two possible
247 causes: Either the sequences we identified represent MGEs that recently underwent HGT and
248 have not had time to diverge from one another, or they represent genetic elements that are
249 highly intolerant to mutation. We suspect that our dataset contains both cases. In the two
250 instances of likely within-patient HGT, both plasmids isolated from the same patient were nearly
251 identical to one another, suggesting that they were indeed transferred shortly before the
252 bacteria were isolated. In both cases we also observed similar plasmids in the genomes of
253 isolates from other patients, but we identified a likely route of transfer between patients only in
254 the case where the subsequent plasmid was also nearly identical. This finding further supports
255 the idea that high plasmid identity is evidence of recent transfer. On the other hand, the Tn7
256 transposon sequence we uncovered that was identical in bacterial isolates from three different
257 genera was also identical to over two dozen publicly available genome sequences queried
258 through a standard NCBI BLAST search. The insertion of the Tn7 transposon downstream of

*glmS* in all of our isolates suggests TnsD-mediated transposition[29], but the reason why the entire transposon sequence is so highly conserved is unclear.

The vast majority of MGE sequences identified through our approach contained signatures of mobile elements, and our follow-up work demonstrated that they could very likely move independently and assemble mosaically on larger mobile elements, such as plasmids, integrative conjugative elements, and other genomic islands. Antibiotic resistance genes were present in fewer MGE clusters than we anticipated, given how many resistance genes are known to be MGE-associated. Our follow-up analysis showed, however, that resistance genes were indeed highly prevalent among the larger MGEs that we resolved. This suggests that resistance genes often reside on smaller and more variable elements, which would have been filtered out by the parameters of our initial screen. A recent study of clinical *K. pneumoniae* genomes showed that while antibiotic resistance genes were largely maintained at the population level, they were variably present on different MGEs that fluctuated in their prevalence over time[24]. Finally, we were somewhat surprised by the large number of metal-interacting genes and operons within the MGEs that we identified. While metal-interacting genes and operons have been hypothesized to confer disinfectant tolerance and increased virulence[30,31], precisely how these elements might increase bacterial survival in the hospital environment and/or contribute to infection requires further study.

Identification of risk factors and common exposures for HGT has previously been proposed[1,14,18,32], but the results of prior efforts have been limited because large genomic datasets from single health systems with corresponding epidemiologic data have not been widely available[33]. The use of routine whole-genome sequencing for outbreak surveillance in our hospital has allowed us to begin to study how the transmission of MGEs might be similar or different from bacterial transmission. In addition to finding evidence of vertical transfer of

plasmids accompanying bacterial transmission, we also identified several cases in which the same MGE lineage was identified in two or more isolates of different sequence types, species, or genera. In some cases, these isolates were collected within days or weeks of one another. This finding underscores how rapidly MGEs can move between bacterial populations, particularly in hospitalized patients[1,21], and highlights the importance of pairing genome sequencing with epidemiologic data to uncover routes of MGE transmission.

There were several limitations to our study. First, the dataset that we used only contained genomes of isolates from clinical infections from a pre-selected list of species, and did not include environmental samples or isolates from patient colonization. Second, our method to screen for putative MGE sequences based on cross-genus alignment was based on somewhat arbitrary cutoffs, and we largely ignored MGEs that only transferred between bacteria within a single genus. Additionally, the cross-genus parameter we employed may have artificially enriched the number of MGEs we identified among *Enterobacteriaceae*, which are known to readily undergo HGT with one another[7]. Third, we assigned MGE lineages relative to single reference sequences and based on our analysis on reference sequence coverage; subsequent MGEs that either gained additional sequence or rearranged their contents would still be assigned to the same lineage, even though they may have diverged substantially from the reference MGE[6]. Finally, this study was based exclusively on comparative genome analyses, and the MGEs we resolved from clinical isolate genomes were not queried for their capacity to undergo HGT *in vitro*.

In conclusion, we have shown how bacterial whole genome sequence data, which is increasingly being generated in clinical settings, can be leveraged to study the dynamics of HGT between drug-resistant bacterial pathogens within a single hospital. Our future work will include further characterization of the MGEs we resolved, assessment of MGE sharing across closer

311 genetic distances, and incorporation of additional epidemiologic information to identify shared
312 exposures and possible routes for MGE transfer independent from bacterial transmission.
313 Ultimately we aim to develop this analysis into a reliable method that can generate actionable
314 information and enhance traditional approaches to prevent and control multidrug-resistant
315 bacterial infections.
316
317 **METHODS**
318 **Isolate collection**
319 Isolates were collected through the Enhanced Detection System for Hospital-Acquired
320 Transmission (EDS-HAT) project at the University of Pittsburgh[19]. Eligibility of bacterial isolates
321 for genome sequencing under EDS-HAT required positive clinical culture for high-priority and
322 multidrug-resistant bacterial pathogens with either of the following criteria: >3 hospital days after
323 admission, and/or any procedure or prior inpatient stay in the 30 days prior to isolate collection.
324 Pathogens collected included: *Acinetobacter spp., Burkholderia spp., Citrobacter spp.,*
325 *Clostridioides difficile,* vancomycin-resistant *Enterococcus spp.,* extended-spectrum beta-
326 lactamase (ESBL)-producing *E. coli*, ESBL-producing *Klebsiella spp., Proteus spp., Providencia*
327 *spp., Pseudomonas spp., Serratia spp., Stenotrophomonas spp.,* and methicillin-resistant *S.*
328 *aureus*. Eligible isolates were identified using TheraDoc software (Version 4.6, Premier, Inc.,
329 Charlotte, NC). The EDS-HAT project involves no contact with human subjects; the project was
330 approved by the University of Pittsburgh Institutional Review Board and was classified as being
331 exempt from informed consent. De-identified patient IDs and culture dates were utilized in
332 downstream analysis.
333
334 **Whole genome sequencing and analysis**
335 Genomic DNA was extracted from pure overnight cultures of single bacterial colonies using a
336 Qiagen DNeasy Tissue Kit according to manufacturer's instructions (Qiagen, Germantown, MD).

Illumina library construction and sequencing were conducted using the Illumina Nextera DNA Sample Prep Kit with 150bp paired-end reads, and libraries were sequenced on the NextSeq sequencing platform (Illumina, San Diego, CA). Selected isolates were also sequenced with long-read technology on a MinION device (Oxford Nanopore Technologies, Oxford, United Kingdom). Long-read sequencing libraries were prepared and multiplexed using a rapid multiplex barcoding kit (catalog RBK-004), and were sequenced on R9.4.1 flow cells. Base-calling on raw reads was performed using Albacore v2.3.3 or Guppy v2.3.1 (Oxford Nanopore Technologies, Oxford, United Kingdom).

Illumina sequencing data were processed with Trim Galore v0.6.1 to remove sequencing adaptors, low-quality bases, and poor-quality reads. Bacterial species were assigned by k-mer clustering with Kraken v1.0[34] and RefSeq[35] databases. Genomes were assembled with SPAdes v3.11[36], and assembly quality was verified using QUAST[37]. Genomes were annotated with Prokka v1.13[38]. Multi-locus sequence types (STs) were assigned using PubMLST typing schemes with mlst v2.16.1[39,40], and ribosomal sequence types (rMLSTs) for isolates not assigned an ST were approximated by alignment to rMLST reference sequences. Long-read sequence data was combined with Illumina data for the same isolate, and hybrid assembly was conducted using Unicycler v0.4.7 or v0.4.8-beta[41].

**Identification and phylogenetic analysis of putative MGEs**

Illumina genome assemblies were screened all-by-all against one another to identify alignments of at least 5,000bp at 100% identity using nucmer v4.0.0beta2[20]. The nucmer output was filtered to only include alignments between bacterial isolates of different genera. Nucleotide sequences from the resulting alignments were then extracted and compared against one another by all-by-all BLASTn v2.7.1[42]. Results were filtered to only include nucleotide sequences having 100% identity over at least 5,000bp to at least one sequence from another genus. The resulting

363 comparisons were clustered and visualized using Cytoscape v3.7.1[43]. A phylogeny of MGE-
364 encoding genomes was constructed using the Genome Taxonomy Database Tool Kit
365 (GTDBTK)[22]. Briefly, translated amino acid sequences of 120 ubiquitous bacterial genes were
366 generated, concatenated, and aligned using GTDBTK's *identify* pipeline. The resulting multiple
367 sequence alignment was masked for gaps and uncertainties, then a phylogenetic tree was
368 generated using RAxML v8.0.26 with the PROTGAMMA substitution model[44] and 1000
369 bootstraps.
370
371 **Characterization of MGE fragments and assignment of chromosomal element and**
372 **plasmid lineages**
373 The longest nucleotide sequence in each MGE cluster was considered representative of that
374 cluster, and was annotated with Prokka v1.13. Representative sequences were compared to
375 publicly available genomes by BLASTn v2.7.1 against the NCBI Nucleotide database. Antibiotic
376 resistance genes were identified by a BLASTn-based search against the CARD v3.0.1[29] and
377 ResFinder v3.2[46] databases, and plasmid replicons were identified by a BLASTn-based search
378 against the PlasmidFinder database v2.0.2[47]. Additional features of each MGE cluster were
379 identified by consulting annotations assigned by Prokka. MGEs were aligned to one another
380 using Geneious v11.1.5 (Biomatters Ltd., Aukland, New Zealand) and EasyFig v2.2.2[48].
381
382 To resolve larger mobile elements encodings MGEs C1-C5, we first selected the earliest isolate
383 containing each MGE for long-read sequencing and hybrid assembly. The closed, MGE-
384 encoding element (plasmid or chromosomal) from this earliest isolate was used as a reference
385 for mapping contigs from Illumina assemblies from all other isolates using BLASTn. Briefly,
386 contigs from Illumina-only assemblies were aligned to each reference MGE-encoding element,
387 and isolates having at least 90% coverage of the reference element were assigned to that
388 element's "lineage." Among isolates having less than 90% coverage, a representative was again

selected for long-read sequencing and hybrid assembly, and the process was repeated until all 104 isolates had been assigned to a lineage. Lineages were named based on the MGE-containing element type (c = chromosomal, p = plasmid), the reference isolate, and the hybrid assembly contig number, denoted with an underscore at the end of the name. MGE cluster-containing plasmids resolved through hybrid assembly were also used as reference sequences to query their presence in the entire 2,173 genome data set using the same BLASTn coverage-based analysis as above. When isolate genomes showed high coverage of multiple reference plasmids, the longest plasmid having at least 90% coverage was recorded.

**Systematic chart review to assess epidemiologic links between patients with the same plasmids**

Patients whose isolates carried the two plasmids found to putatively transfer within individual patients were reviewed using a systematic approach modified from previously published methodologies examining patient locations and procedures for potential similarities[49,50]. Patients were considered infected/colonized with the recovered plasmid on the day of the patients' culture and all subsequent days. Potential transfer events were considered significant for locations if an uninfected/uncolonized patient was housed on the same unit location or service line location (units with shared staff) at the same time or different time as a patient infected/colonized with the plasmid, using a 60-day window prior to the newly infected/colonized patient's culture date. Additionally, procedures (e.g. operation room procedures, bedside invasive procedures) were evaluated for commonalities among all patients 60 days prior to infection/colonization, as well as potential procedures contaminated by prior infected/colonized patients that could have transferred to newly infected/colonized patients, again using a 60-day window prior to the culture date. Procedures were deemed significant if >1 patient had a similar procedure, or if there was a shared procedure within the 60-day window.

**Acknowledgements**

We gratefully acknowledge Chinelo Ezeonwuka, Kathleen Shutt, Daniel Snyder, Jieshi Chen, and Hayley Nordstrom for their helpful contributions to this study. This work was supported by a grant from the Competitive Medical Research Fund of the UPMC Health System to D.V.T., by NIAID grants R21AI109459 and R01AI127472 to L.H.H., and by the Department of Medicine at the University of Pittsburgh School of Medicine. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

**REFERENCES**

1. Lerminiaux, N. A. & Cameron, A. D. S. Horizontal transfer of antibiotic resistance genes in clinical environments. *Can. J. Microbiol.* **65**, 34–44 (2019).
2. Juhas, M. Horizontal gene transfer in human pathogens. *Crit. Rev. Microbiol.* **41**, 101–108 (2015).
3. Jamrozy, D. *et al.* Evolution of mobile genetic element composition in an epidemic methicillin-resistant Staphylococcus aureus: temporal changes correlated with frequent loss and gain events. *BMC Genomics* **18**, 684 (2017).
4. Bosch, T. *et al.* Outbreak of NDM-1-Producing Klebsiella pneumoniae in a Dutch Hospital, with Interspecies Transfer of the Resistance Plasmid and Unexpected Occurrence in Unrelated Health Care Centers. *J. Clin. Microbiol.* **55**, 2380–2390 (2017).
5. Martin, J. *et al.* Covert dissemination of carbapenemase-producing Klebsiella pneumoniae (KPC) in a successfully controlled outbreak: long- and short-read whole-genome sequencing demonstrate multiple genetic modes of transmission. *J. Antimicrob. Chemother.* **72**, 3025–3034 (2017).
6. Sheppard, A. E. *et al.* Nested Russian Doll-Like Genetic Mobility Drives Rapid Dissemination of the Carbapenem Resistance Gene blaKPC. *Antimicrob. Agents Chemother.* **60**, 3767–3778 (2016).

7. Cerqueira, G. C. *et al.* Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1135–1140 (2017).
8. 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* **6**, e4210 (2018).
9. Hazen, T. H. *et al.* Diversity among blaKPC-containing plasmids in Escherichia coli and other bacterial species isolated from the same patients. *Sci. Rep.* **8**, 10291 (2018).
10. Schmithausen, R. M. *et al.* The washing machine as a reservoir for transmission of extended spectrum beta-lactamase (CTX-M-15)-producing Klebsiella oxytoca ST201 in newborns. *Appl. Environ. Microbiol.* (2019) doi:10.1128/AEM.01435-19.
11. Stadler, T. *et al.* Transmission of ESBL-producing Enterobacteriaceae and their mobile genetic elements-identification of sources by whole genome sequencing: study protocol for an observational study in Switzerland. *BMJ Open* **8**, e021823 (2018).
12. Ladner, J. T., Grubaugh, N. D., Pybus, O. G. & Andersen, K. G. Precision epidemiology for infectious disease control. *Nat. Med.* **25**, 206–211 (2019).
13. Sundermann, A. J. *et al.* Automated data mining of the electronic health record for investigation of healthcare-associated outbreaks. *Infect. Control Hosp. Epidemiol.* **40**, 314–319 (2019).
14. Hardiman, C. A. *et al.* Horizontal Transfer of Carbapenemase-Encoding Plasmids and Comparison with Hospital Epidemiology Data. *Antimicrob. Agents Chemother.* **60**, 4910–4919 (2016).
15. Orlek, A. *et al.* Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. *Front. Microbiol.* **8**, 182 (2017).
16. Arredondo-Alonso, S., Willems, R. J., van Schaik, W. & Schürch, A. C. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb.*

*Genomics* **3**, e000128 (2017).

17. George, S. *et al.* Resolving plasmid structures in Enterobacteriaceae using the MinION nanopore sequencer: assessment of MinION and MinION/Illumina hybrid data assembly approaches. *Microb. Genomics* **3**, e000118 (2017).
18. Conlan, S. *et al.* Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Sci. Transl. Med.* **6**, 254ra126 (2014).
19. Sundermann, A. J. *et al.* Outbreak of Vancomycin-resistant Enterococcus faecium in Interventional Radiology: Detection Through Whole Genome Sequencing-Based Surveillance. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* (2019) doi:10.1093/cid/ciz666.
20. Marçais, G. *et al.* MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **14**, e1005944 (2018).
21. Huddleston, J. R. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infect. Drug Resist.* **7**, 167–176 (2014).
22. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
23. Marsh, J. W. *et al.* Clostridioides difficile: a potential source of NpmA in the clinical environment. *J. Antimicrob. Chemother.* **74**, 521–523 (2019).
24. Komano, T., Fujitani, S., Funayama, N., Kanno, A. & Sakuma, K. Physical and genetic analyses of IncI2 plasmid R721: evidence for the presence of shufflon. *Plasmid* **23**, 248–251 (1990).
25. Manges, A. R. *et al.* Global Extraintestinal Pathogenic Escherichia coli (ExPEC) Lineages. *Clin. Microbiol. Rev.* **32**, (2019).
26. Valenzuela, J. K. *et al.* Horizontal gene transfer in a polyclonal outbreak of carbapenem-resistant Acinetobacter baumannii. *J. Clin. Microbiol.* **45**, 453–460 (2007).
27. Schweizer, C. *et al.* Plasmid-Mediated Transmission of KPC-2 Carbapenemase in

493 Enterobacteriaceae in Critically Ill Patients. *Front. Microbiol.* **10**, 276 (2019).

494 28. Savinova, T. A., Bocharova, Y. A., Lazareva, A. V., Chebotar, I. V. & Mayanskiy, N. A.
495 [Integron diversity in blavim-2-carrying carbapenem-resistant clinical Pseudomonas
496 aeruginosa isolates.]. *Klin. Lab. Diagn.* **64**, 497–502 (2019).

497 29. Parks, A. R. & Peters, J. E. Tn7 elements: engendering diversity from chromosomes to
498 episomes. *Plasmid* **61**, 1–14 (2009).

499 30. McDonnell, G. & Russell, A. D. Antiseptics and disinfectants: activity, action, and resistance.
500 *Clin. Microbiol. Rev.* **12**, 147–179 (1999).

501 31. Chandrangsu, P., Rensing, C. & Helmann, J. D. Metal homeostasis and resistance in
502 bacteria. *Nat. Rev. Microbiol.* **15**, 338–350 (2017).

503 32. Pecora, N. D. *et al.* Genomically Informed Surveillance for Carbapenem-Resistant
504 Enterobacteriaceae in a Health Care System. *mBio* **6**, e01030 (2015).

505 33. Struelens, M. J. The epidemiology of antimicrobial resistance in hospital acquired infections:
506 problems and possible solutions. *BMJ* **317**, 652–654 (1998).

507 34. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using
508 exact alignments. *Genome Biol.* **15**, R46 (2014).

509 35. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI reference sequences (RefSeq): a curated
510 non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids*
511 *Res.* **35**, D61-65 (2007).

512 36. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to
513 single-cell sequencing. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* **19**, 455–477 (2012).

514 37. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for
515 genome assemblies. *Bioinforma. Oxf. Engl.* **29**, 1072–1075 (2013).

516 38. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinforma. Oxf. Engl.* **30**,
517 2068–2069 (2014).

518 39. Jolley, K. A. & Maiden, M. C. J. BIGSdb: Scalable analysis of bacterial genome variation at

519 the population level. *BMC Bioinformatics* **11**, 595 (2010).

520 40. Seemann, T. *mlst*. (Github).

521 41. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome
522 assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595
523 (2017).

524 42. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment
525 search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

526 43. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular
527 interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

528 44. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
529 phylogenies. *Bioinforma. Oxf. Engl.* **30**, 1312–1313 (2014).

530 45. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive
531 antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).

532 46. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J. Antimicrob.*
533 *Chemother.* **67**, 2640–2644 (2012).

534 47. Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and
535 plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).

536 48. Sullivan, M. J., Petty, N. K. & Beatson, S. A. Easyfig: a genome comparison visualizer.
537 *Bioinforma. Oxf. Engl.* **27**, 1009–1010 (2011).

538 49. Eyre, D. W. *et al.* Diverse sources of C. difficile infection identified on whole-genome
539 sequencing. *N. Engl. J. Med.* **369**, 1195–1205 (2013).

540 50. Ward, D. V. *et al.* Integration of genomic and clinical data augments surveillance of
541 healthcare-acquired infections. *Infect. Control Hosp. Epidemiol.* **40**, 649–655 (2019).

542 **Table 1.** MGE-encoding lineages and associated antibiotic resistance and metal interaction

543 gene contents.

| MGE Lineage[a] | Length (kb) | Replicons (PlasmidFinder) | Antibiotic Resistance (ResFinder) | Metal Interaction (Prokka) |
|---|---|---|---|---|
| cEC00609 | 39 | None | *aac(3)-IIa, aac(6')-Ib-cr, blaCTX-M-1, blaOXA-1, catB3, tet(A)* | None |
| pCB00017_2 | 196.8 | IncFIB(K), IncFII(K) | *aac(6')-Ib-cr, aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaOXA-1, blaTEM-1B, catB3, qnrB1, tet(A), sul2* | *copD* operon, *pcoE, silE, silP, ars* operon |
| pCB00028_2 | 383.1 | IncHI2, IncHI2A, repA | *aac(3)-IIa, aac(6')-Ib-cr, aadA1, aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaOXA-1, baTEM-1B, catA1, catB3, dfrA14, sul2, tet(A)* | *pcoE, merR, merB* |
| pEC00668_2 | 145.4 | IncFIA, IncFIB(AP001918), IncFII(pRSB107) | *aac(6)-Id, aph(3'')-Ib, dfrA14, blaTEM-1B, mph(A), sul2* | *efeU, merA, merC, merP, merR, merT* |
| pEC00690_2 | 106.8 | IncFIA, IncFII | *aac(6')-Ibcr, blaOXA-1, catB3, tet(A)* | *efeU* |
| pKLP00149_2 | 165.2 | FII(pBK30683) | *aac(6')-Ib, aac(6')-Ib-cr, aadA1, aph(3'')-Ib, aph(6)-Id, blaKPC-3, blaOXA-9, blaSHV-182, blaTEM-1A, dfrA14, sul2* | *csoR* |
| pKLP00155_6 | 9.5 | ColRNAI | None | None |
| pKLP00161_2 | 236.5 | IncFIB(K), IncFII(K) | *aac(6')-Ib-cr, aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaOXA-1, blaTEM-1B, dfrA14, qnrB1, sul2, tet(A)* | *copD* operon, *pcoC, pcoE, silE, silP, ars* operon |
| pKLP00177_3 | 170.8 | IncFIB(K) | *aac(3)-IIa, aac(6')-Ib-cr, aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaOXA-1, blaTEM-1B, catB3, dfrA14, qnrB1, sul2, tet(A)* | *copD* operon, *pcoC, pcoE, silE, silP, ars* operon |
| pKLP00182_3 | 15.8 | None | *aac(6')-Ib-cr, blaOXA-1, catB3, dfrA14, tet(A)* | None |
| pKLP00215_4 | 113.6 | IncFIB(pQil), IncFII(K) | *blaKPC-2, blaOXA-9, blaTEM-1A* | *merB, merR* |
| pKLP00218_2 | 164.7 | IncFIB(K), IncFII(K) | *aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaTEM-1B, dfrA14, sul2* | *copD* operon, *pcoC, pcoE, silE, silP, ars* operon |
| pKLP00221_2 | 242.3 | ColRNAI, IncFIB(K), IncFII(K) | *aac(6')-Ib, aada2, aph(3')-1a, blaKPC-2, blaOXA-9, blaTEM-1A, catA1, dfrA12, mph(A), sul1* | *copD* operon, *pcoC, pcoE, silE, silP, ars* operon |
| pKLO00017_2 | 226.2 | IncFIB(K), IncFII(K) | *aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaTEM-1B, dfrA14, sul2* | *copD* operon, *pcoC, pcoE, silE, silP, ars* operon |

544 [a]MGE lineage names include location (c = chromosome, p = plasmid), name of the reference

545 isolate sequenced, and assembly contig number (_2, _3, _4, _6).

## Figure Legends

**Figure 1.** Identical mobile genetic elements (MGEs) shared across bacterial genera in a single hospital. (A) Approach to identify and characterize MGEs. (B) MGEs identified by the approach shown in (A). 51 MGE clusters found in distinct genera visualized with Cytoscape. Nodes represent bacterial isolates and are color-coded by genus. Edges connect nodes from different genera sharing >5kb of sequence at 100% nucleotide sequence identity. Clusters examined more closely in subsequent figures are labeled. (C) Genus distribution of all 2,173 genomes in the dataset (left) and the 196 isolates with MGEs shared across genera (right). (D) Prevalence of annotated mobile element, antimicrobial resistance (AMR) and metal-interacting genes among 51 MGE clusters. (E) Summary of AMR genes identified in MGE clusters. Genes are grouped by antibiotic class and bubble sizes correspond to prevalence among the MGE clusters shown in (B). AMR gene names are listed inside each bubble.

**Figure 2.** Phylogenetic distribution of MGE clusters across 196 genomes. Black squares mark the presence of one or more MGE clusters in each genome, with each column corresponding to a different MGE cluster. The heat map to the right shows MGE cluster density (i.e. total number of cross-genus MGEs) in each bacterial genome. Clusters examined more closely in subsequent figures are labeled and shaded in gray.

**Figure 3.** Examples of MGE sharing across genera. (A) Genes shared between a vancomycin-resistant *E. faecium* (VRE) plasmid and a *C. difficile* chromosome (MGE cluster C9). The VanA operon, conferring vancomycin resistance, is marked with an orange bar. Shared drug resistance genes are colored magenta, and mobile element genes are colored blue. Gray shading marks a stretch of DNA sequence that is 100% identical between isolates. (B) Identical portions of an integrated conjugative element (MGE cluster C30) shared between an *S. marcescens* genome (SER00094) and two *P. aeruginosa* genomes (PSA00048 and PSA00656). Blue = *intS* integrase; green = formaldehyde resistance genes; gray = UvrABC system genes. Type IV secretion machinery is marked with an orange bar, and gray shading

572 marks sequences that are 100% identical between isolates. (C) Identical Tn7 transposons
573 shared between *A. baumannii*, *E. coli*, and *P. mirabilis* (MGE cluster C17). The Tn7 sequence of
574 the pR721 plasmid is shown at the top. The *tnsABCDE* transposon machinery is marked with an
575 orange bar, and the *glmS* gene, which flanks the Tn7 insertion site, is colored red. Shared drug
576 resistance genes are colored magenta, and an *xerH* tyrosine recombinase is colored blue. Gray
577 shading marks sequences that are 100% identical between isolates.

578 **Figure 4.** Mosaicism of MGE clusters present on diverse elements. (A) Circular plot of six
579 distinct sequence elements (black bars) that encode MGE clusters C1, C2, and C3. Lowercase
580 letters in sequence names indicate element type (c = chromosome, p = plasmid). Homologous
581 cluster sequences are connected to one another with colored links (purple = C1, orange = C2,
582 green = C3, gray = other). Inner circle depicts MGE genes involved in MGE mobilization (blue),
583 antibiotic resistance (red) and metal interaction (gray). (B-D) Alignments of sequences grouped
584 into MGE clusters C1 (B), C2 (C), and C3 (D) from the larger MGEs displayed in (A). ORFs are
585 colored by function (blue = mobilization, red = antibiotic resistance, green = other/hypothetical).
586 Antibiotic resistance genes are labeled above and dark gray blocks connect sequences that are
587 identical over at least 5kb.

588 **Figure 5:** Timelines of plasmid lineage occurrence among isolates of the same ST (A), same
589 genus (B), or different genera (C). Timelines show the culture date of isolates predicted to
590 encode plasmids belonging to the same lineage, based on coverage mapping to reference
591 plasmids listed to the left of each timeline. The MGE clusters carried by each plasmid are listed
592 in parentheses below the plasmid name. More information about each plasmid is provided in
593 Table 1. Shape and color of data points correspond to bacterial species and ST, respectively.

594 **Figure 6:** Cross-genus transfer of plasmids within and between patients. (A) Schematic diagram
595 showing *K. pneumoniae* and *E. coli* isolates bearing the same plasmid sampled from two
596 patients. (B) Nucleotide alignment of the plasmid presumably transferred within and between
597 the patients shown in (A). A 113.6kb IncFIB(pQil)/IncFII(K) carbapenemase-encoding plasmid

598 was resolved from two genomes of different bacterial isolates from the same clinical specimen
599 from Patient A. A nearly identical plasmid was also identified in an isolate from Patient B, who
600 occupied a hospital room adjacent to Patient A. (C) Alignment of a 196.8kb IncFIB(K)/IncFII(K)
601 multidrug-resistance plasmid resolved from two genomes of different bacterial isolates from the
602 same clinical specimen from Patient C. Similar plasmids were also found in isolates from two
603 additional patients (Patient D and Patient E), who had no identifiable epidemiologic links with
604 Patient C. ORFs are colored by function (blue = mobilization, red = antibiotic resistance, gray =
605 metal-interacting, green = other/hypothetical). Antibiotic resistance genes, metal-interacting
606 operons, and Type IV secretion components are labeled. Shading between sequences indicates
607 regions >5kb with >99.9% identity, and pairwise identities across the entire plasmid are noted to
608 the right.

Fig. 1

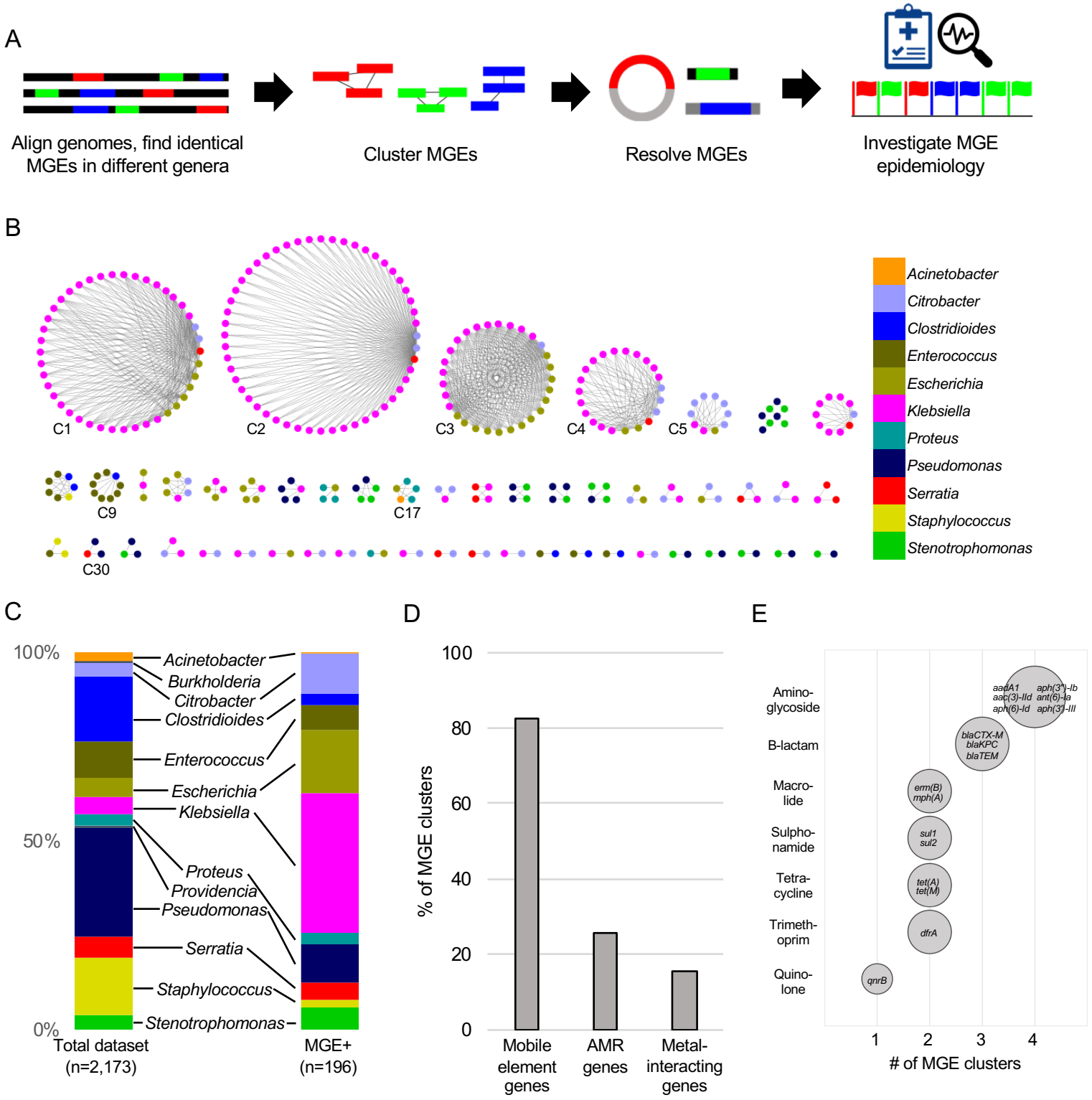


**Figure 1.** Identical mobile genetic elements (MGEs) shared across bacterial genera in a single hospital. (A) Approach to identify and characterize MGEs. (B) MGEs identified by the approach shown in (A). 51 MGE clusters found in distinct genera visualized with Cytoscape. Nodes represent bacterial isolates and are color-coded by genus. Edges connect nodes from different genera sharing >5kb of sequence at 100% nucleotide sequence identity. Clusters examined more closely in subsequent figures are labeled. (C) Genus distribution of all 2,173 genomes in the dataset (left) and the 196 isolates with MGEs shared across genera (right). (D) Prevalence of annotated mobile element, antimicrobial resistance (AMR) and metal-interacting genes among 51 MGE clusters. (E) Summary of AMR genes identified in MGE clusters. Genes are grouped by antibiotic class and bubble sizes correspond to prevalence among the MGE clusters shown in (B). AMR gene names are listed inside each bubble.

Fig. 2

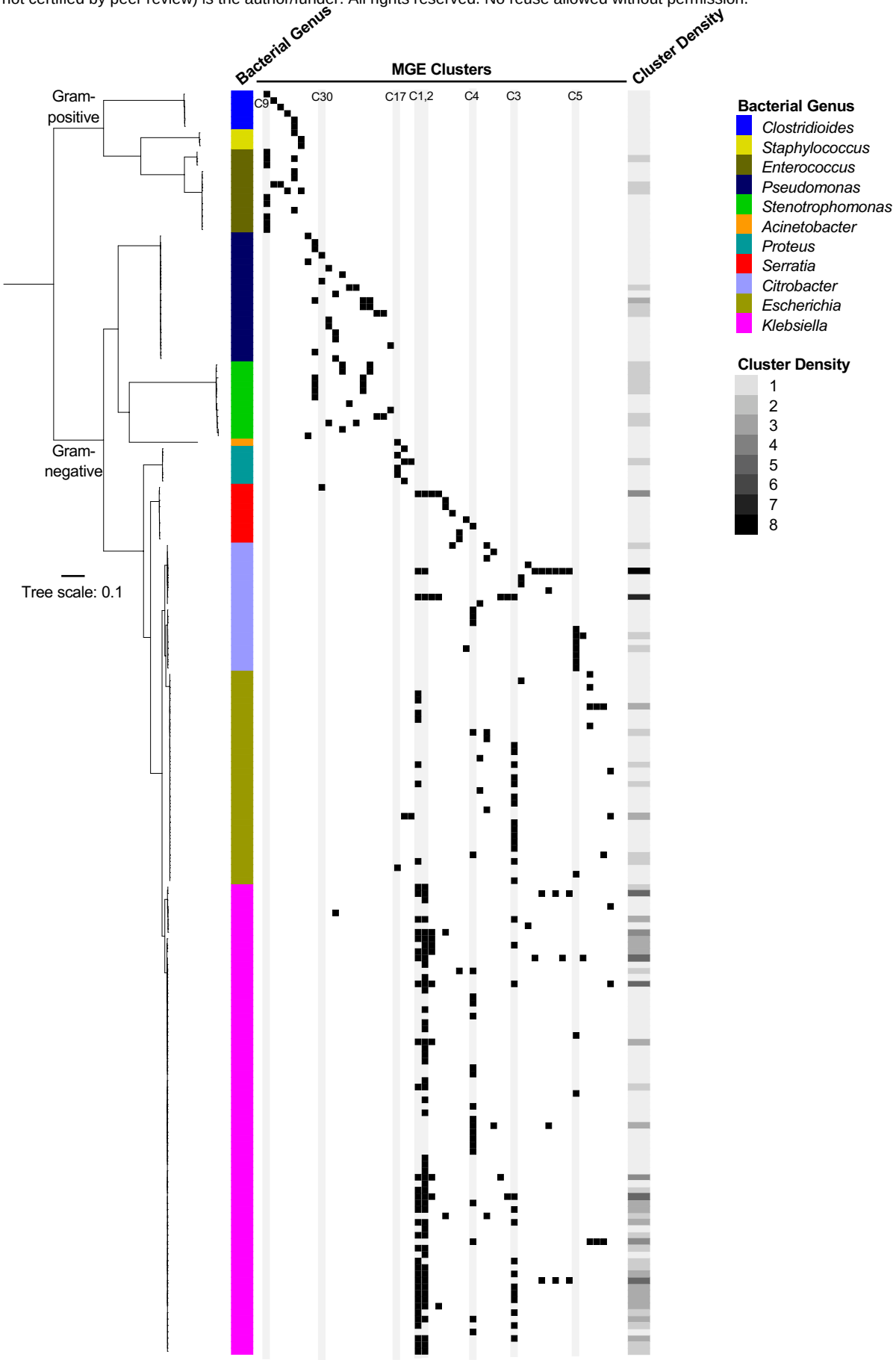


**Figure 2.** Phylogenetic distribution of MGE clusters across 196 genomes. Black squares mark the presence of one or more MGE clusters in each genome, with each column corresponding to a different MGE cluster. The heat map to the right shows MGE cluster density (i.e. total number of cross-genus MGEs) in each bacterial genome. Clusters examined more closely in subsequent figures are labeled and shaded in gray.

Fig. 3

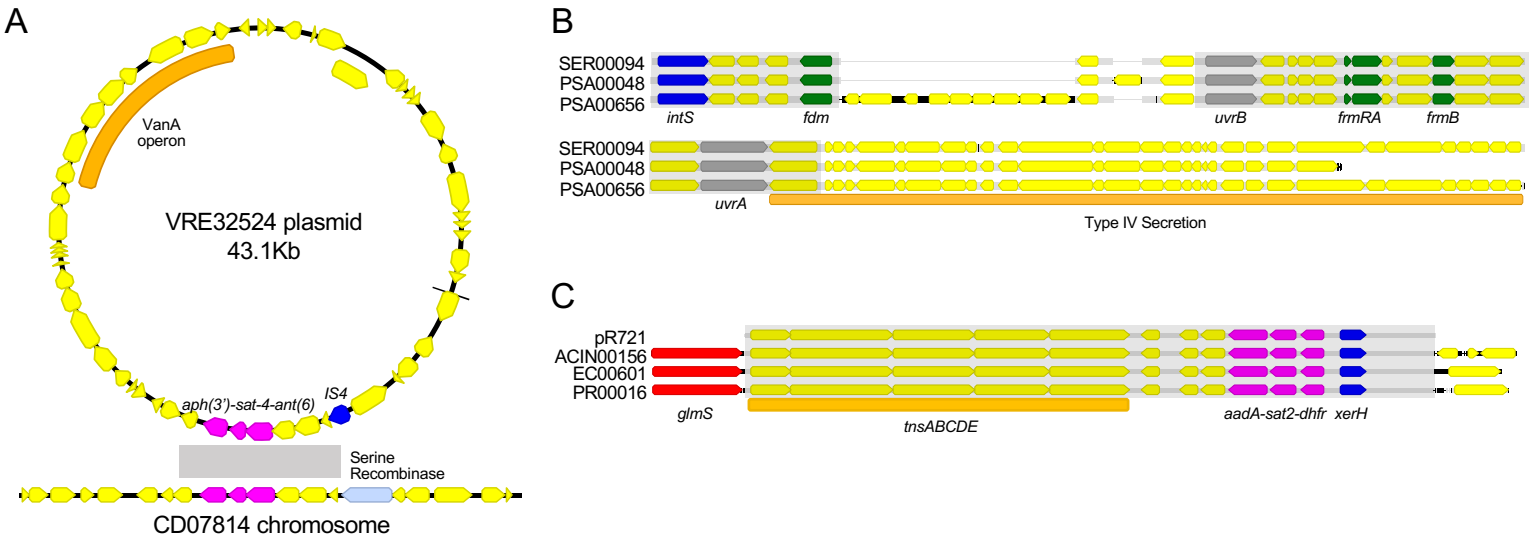


**Figure 3.** Examples of MGE sharing across genera. (A) Genes shared between a vancomycin-resistant *E. faecium* (VRE) plasmid and a *C. difficile* chromosome (MGE cluster C9). The VanA operon, conferring vancomycin resistance, is marked with an orange bar. Shared drug resistance genes are colored magenta, and mobile element genes are colored blue. Gray shading marks a stretch of DNA sequence that is 100% identical between isolates. (B) Identical portions of an integrated conjugative element (MGE cluster C30) shared between an *S. marcescens* genome (SER00094) and two *P. aeruginosa* genomes (PSA00048 and PSA00656). Blue = *intS* integrase; green = formaldehyde resistance genes; gray = UvrABC system genes. Type IV secretion machinery is marked with an orange bar, and gray shading marks sequences that are 100% identical between isolates. (C) Identical Tn7 transposons shared between *A. baumannii*, *E. coli*, and *P. mirabilis* (MGE cluster C17). The Tn7 sequence of the pR721 plasmid is shown at the top. The *tnsABCDE* transposon machinery is marked with an orange bar, and the *glmS* gene, which flanks the Tn7 insertion site, is colored red. Shared drug resistance genes are colored magenta, and an *xerH* tyrosine recombinase is colored blue. Gray shading marks sequences that are 100% identical between isolates.

Fig. 4

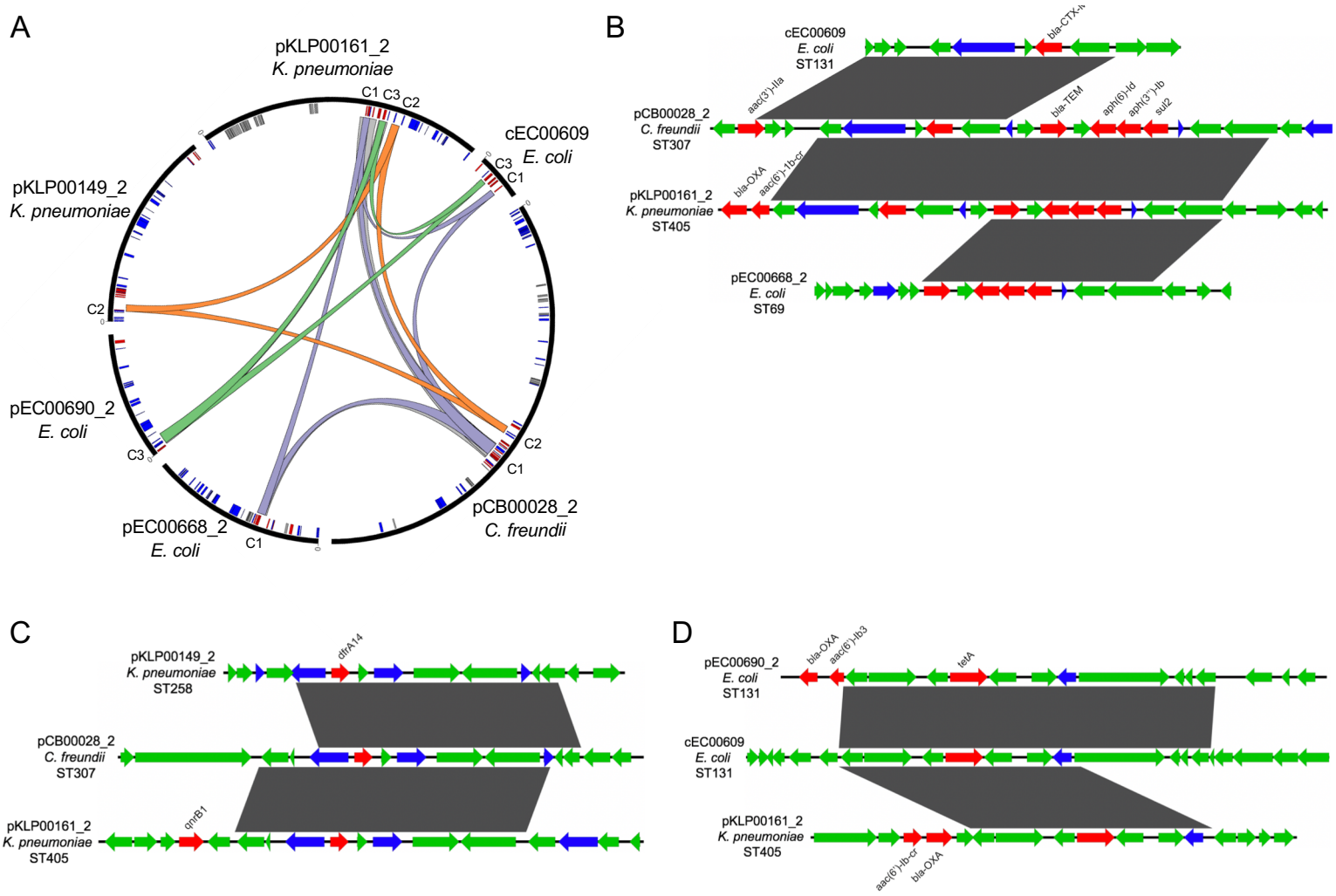


**Figure 4.** Mosaicism of MGE clusters present on diverse elements. (A) Circular plot of six distinct sequence elements (black bars) that encode MGE clusters C1, C2, and C3. Lowercase letters in sequence names indicate element type (c = chromosome, p = plasmid). Homologous cluster sequences are connected to one another with colored links (purple = C1, orange = C2, green = C3, gray = other). Inner circle depicts MGE genes involved in MGE mobilization (blue), antibiotic resistance (red) and metal interaction (gray). (B-D) Alignments of sequences grouped into MGE clusters C1 (B), C2 (C), and C3 (D) from the larger MGEs displayed in (A). ORFs are colored by function (blue = mobilization, red = antibiotic resistance, green = other/hypothetical). Antibiotic resistance genes are labeled above and dark gray blocks connect sequences that are identical over at least 5kb.

Fig. 5

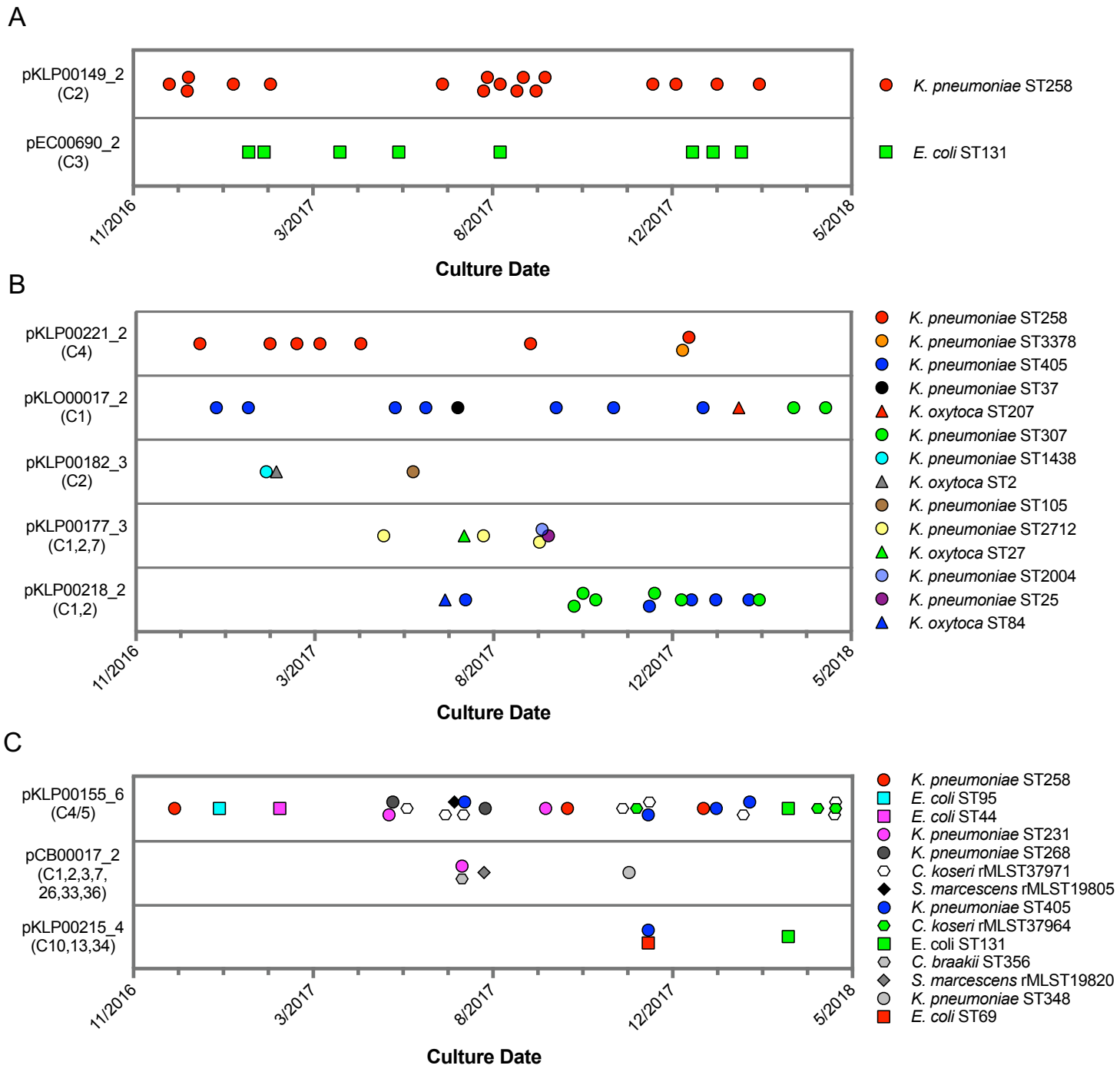


**Figure 5:** Timelines of plasmid lineage occurrence among isolates of the same ST (A), same genus (B), or different genera (C). Timelines show the culture date of isolates predicted to encode plasmids belonging to the same lineage, based on coverage mapping to reference plasmids listed to the left of each timeline. The MGE clusters carried by each plasmid are listed in parentheses below the plasmid name. More information about each plasmid is provided in Table 1. Shape and color of data points correspond to bacterial species and ST, respectively.

Fig. 6

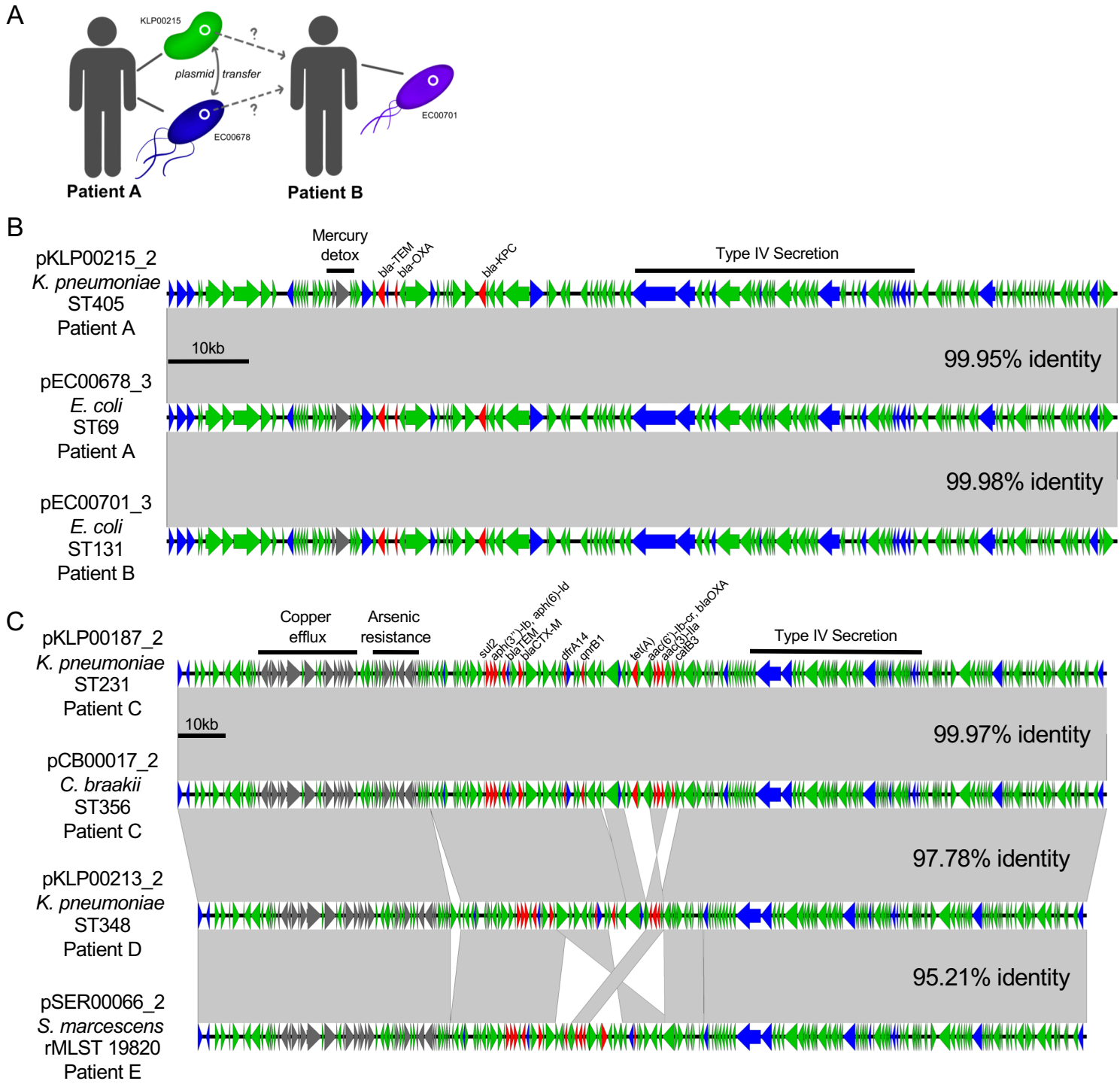


**Figure 6:** Cross-genus transfer of plasmids within and between patients. (A) Schematic diagram showing *K. pneumoniae* and *E. coli* isolates bearing the same plasmid sampled from two patients. (B) Nucleotide alignment of the plasmid presumably transferred within and between the patients shown in (A). A 113.6kb IncFIB(pQil)/IncFII(K) carbapenemase-encoding plasmid was resolved from two genomes of different bacterial isolates from the same clinical specimen from Patient A. A nearly identical plasmid was also identified in an isolate from Patient B, who occupied a hospital room adjacent to Patient A. (C) Alignment of a 196.8kb IncFIB(K)/IncFII(K) multidrug-resistance plasmid resolved from two genomes of different bacterial isolates from the same clinical specimen from Patient C. Similar plasmids were also found in isolates from two additional patients (Patient D and Patient E), who had no identifiable epidemiologic links with Patient C. ORFs are colored by function (blue = mobilization, red = antibiotic resistance, gray = metal-interacting, green = other/hypothetical). Antibiotic resistance genes, metal-interacting operons, and Type IV secretion components are labeled. Shading between sequences indicates regions >5kb with >99.9% identity, and pairwise identities across the entire plasmid are noted to the right.