Genomic dissection of endemic carbapenem resistance: metallo-beta-lactamase gene dissemination through clonal, plasmid and integron transfer pathways

Nenad Macesic¹,², Jane Hawkey¹, Ben Vezina¹, Jessica A. Wisniewski¹, Hugh Cottingham¹, Luke V. Blakeway¹, Taylor Harshey¹, Katherine Pragastis¹, Gnei Zweena Badoordeen¹, Amanda Dennison³, Denis W. Spelman¹,³, Adam W.J. Jenney¹,³, Anton Y. Peleg*¹,²,⁴

¹ Department of Infectious Diseases, The Alfred Hospital and Central Clinical School, Monash University, Melbourne, Australia
² Centre to Impact AMR, Monash University, Clayton, Australia
³ Microbiology Unit, Alfred Hospital, Melbourne, Australia.
⁴ Infection Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Australia

* Corresponding Author
Anton Peleg, Department of Infectious Diseases, The Alfred Hospital and Central Clinical School, Monash University, Level 2, Burnet Institute, 85 Commercial Rd, Melbourne, VIC 3000, Australia
anton.peleg@monash.edu

Keywords:
IMP-4; Enterobacter hormaechei; Klebsiella pneumoniae; Serratia marcescens; class I integron; long-read sequencing; multi-drug resistant Gram-negative bacteria
Abstract:

Infections caused by metallo-beta-lactamase-producing organisms (MBLs) are a global health threat. Our understanding of transmission dynamics and how MBLs establish endemicity remains limited. We analysed two decades of \textit{bla}\textsubscript{IMP-4} evolution in a hospital using sequence data from 270 clinical and environmental isolates (including 169 completed genomes) and identified extreme gene promiscuity across 7 Gram-negative genera, 68 bacterial strains and 7 distinct plasmid types. An initial multi-species outbreak of conserved IncC plasmids (95 genomes across 37 strains) allowed endemicity to be established through the ability of \textit{bla}\textsubscript{IMP-4} to disseminate in successful strain-genetic setting pairs we termed ‘propagators’, in particular \textit{Serratia marcescens} and \textit{Enterobacter hormaechei}. From this reservoir, \textit{bla}\textsubscript{IMP-4} persisted through diversification of genetic settings that resulted from transfer of \textit{bla}\textsubscript{IMP-4} plasmids between bacterial hosts and of the integron carrying \textit{bla}\textsubscript{IMP-4} between plasmids. Our findings provide a framework for understanding endemicity and spread of MBLs and may have broader applicability to other carbapenemase-producing organisms.
Introduction

Carbapenemase-producing organisms (CPOs) are now endemic in many regions. While there has been a significant focus on \textit{bla}\textsubscript{KPC} due to its spread through North America and Europe, metallo-beta-lactamases (MBLs) (e.g. \textit{bla}\textsubscript{NDM}, \textit{bla}\textsubscript{IMP} and \textit{bla}\textsubscript{VIM}) are endemic through much of Asia and Oceania, including Australia, where \textit{bla}\textsubscript{IMP} carbapenemases have dominated\textsuperscript{1-9}. Treatment options for infections caused by CPOs, particularly MBL-harbouring organisms, remain severely limited, highlighting the need to stop further spread of these extensively drug-resistant organisms. The mechanisms for carbapenemase spread differ according to carbapenemase type. Some carbapenemases spread through close associations with successful strains or lineages (e.g. \textit{bla}\textsubscript{KPC-23} and \textit{Klebsiella pneumoniae} clonal complex 258), while for others, spread is mediated through association with specific plasmids (e.g. \textit{bla}\textsubscript{OXA-48} and broad-host range IncL plasmids)\textsuperscript{10,11}. Notably, MBLs spread through both lineage-related clonal expansion and diverse plasmid types\textsuperscript{12,13}. While surveillance studies have captured some of these data, there have been few efforts to assess how these mechanisms of spread evolve over time. Understanding the transmission dynamics of carbapenem resistance genes will be crucial to inform future infection prevention efforts.

Previous work from our group and others have identified that the spread of MBLs, and particularly \textit{bla}\textsubscript{IMP-4}, is often driven by dissemination of a class 1 integron that has been able to insert into several genetic settings\textsuperscript{2,4,6,9,14,15}. The ability to study horizontal gene transfer has been significantly advanced by long-read sequencing, which enables high-quality \textit{de novo} assembly of bacterial genomes, including highly repetitive regions such as plasmids. Utilising long-read sequencing to generate completed, closed bacterial genomes provides a unique opportunity to study the complex, multi-level (bacterial strain, plasmid, gene) transmission dynamics that are likely occurring during MBL spread. When combined with short-read sequencing data, an unprecedented level of detail of the genetic context and likely mechanisms of an outbreak or endemicity is possible.

MBL-producing Gram-negative bacteria, dominated by \textit{bla}\textsubscript{IMP-4}, have been isolated in our institution (The Alfred Hospital) since 2002\textsuperscript{5-7,15}. After an initial outbreak period in 2004-2005, we experienced hyperendemicity, with a repeated outbreak period from 2017-2020. We aimed to assess the genetic settings of \textit{bla}\textsubscript{IMP-4}, its evolution over time, and the transmission pathways that resulted in repeated...
outbreaks and endemicity. We used long- and short-read whole genome sequencing to characterise the genetic settings of \textit{bla}\textsubscript{IMP-4} in bacterial chromosomes and plasmids from 277 clinical and environmental isolates from 2002-2020. This allowed us to track the spread of \textit{bla}\textsubscript{IMP-4} in 7 plasmid types and multiple chromosomal settings. In the 18-year period, we noted incredible plasticity of \textit{bla}\textsubscript{IMP-4} persistence, with vertical spread through the transmission of dominant strains and horizontal spread of both plasmids and the class 1 integron carrying \textit{bla}\textsubscript{IMP-4}. We also identified a persistent reservoir of \textit{bla}\textsubscript{IMP-4} in IncC plasmids in both clinical and environmental isolates. Our findings highlight the need for integration of long-read sequencing into CPO surveillance, as well as for multi-modal infection prevention approaches that address the diverse forms of CPO spread.

**Results**

\textit{bla}\textsubscript{IMP-4} found in diverse clinical and environmental isolates spanning two decades

We sequenced 277 \textit{bla}\textsubscript{IMP-4}-harbouring isolates from an institutional collection of carbapenem-resistant isolates systematically collected from 2002-2020, including 264 clinical isolates from 196 patients and 13 environmental isolates (Supp. Table 1). This included short-read (Illumina) data on all isolates and long-read (Oxford Nanopore) data on 172 isolates that best represented the strains across the study time periods. Seven isolates failed quality control and were excluded. In total, we analysed 270 isolates that were made up of 68 bacterial strains (defined as unique species/multi-locus sequence type [MLST] combinations) from 7 Gram-negative genera, highlighting the diversity of bacterial hosts for \textit{bla}\textsubscript{IMP-4} (Fig. 1A). The five most frequent strains accounted for 190/270 (70%) genomes and included \textit{Serratia marcescens} (52/270 isolates, 19%), \textit{Enterobacter hormaechei} ST190 (44/270 genomes, 17%), \textit{E. hormaechei} ST93 (36/270 genomes, 13%), \textit{Pseudomonas aeruginosa} ST111 (35/270 genomes, 13%) and \textit{E. hormaechei} ST114 (23/270 genomes, 9%) (Fig. 1A). In addition to \textit{bla}\textsubscript{IMP-4}, 8/270 (3%) genomes carried other carbapenemase genes (4 \textit{bla}\textsubscript{OXA-58}, 2 \textit{bla}\textsubscript{NDM-7}, 1 \textit{bla}\textsubscript{NDM-1}, 1 \textit{bla}\textsubscript{KPC-2}, 1 \textit{bla}\textsubscript{OXA-550}) and 121/270 (44%) carried \textit{mcr-9.1}, a novel determinant of colistin resistance\textsuperscript{14} (Supp. Table 1).

\textit{bla}\textsubscript{IMP-4} detected in multiple plasmid and chromosomal genetic settings across three distinct time periods

We first determined the genetic setting of \textit{bla}\textsubscript{IMP-4} using 169 completed, circularised genomes (three non-circularised genomes were excluded). \textit{bla}\textsubscript{IMP-4}-carrying plasmids were clustered using MOB-
typer\textsuperscript{16}, which uses a whole-sequence-based typing system to provide cluster codes for reconstruction and tracking of plasmids. Representative plasmids from each cluster were then used as references for mapping of the 99 genomes with short-read data only and the three non-circularised genomes (Supp. Table 1). Overall, 230 and 40 isolates carried \textit{bla}\textit{IMP}-4 on a plasmid or on the chromosome, respectively, with seven distinct plasmid types identified and chromosomal integration in multiple strains (Fig. 1B and Extended Data Table 1). For the majority of isolates (151/169, 89\%), \textit{bla}\textit{IMP}-4 was situated in a class 1 integron most commonly comprising the \textit{bla}\textit{IMP}-4-qacG-aacA4-catB3-qacE-sul1 cassette array. The bacterial host strain-plasmid relationships evolved over the course of the study, with three distinct time periods (Fig. 1B).

\textbf{Outbreak initiation and establishment of \textit{bla}\textit{IMP}-4 endemicity}

\textit{bla}\textit{IMP}-4 was first noted in a clinical \textit{S. marcescens} isolate in 2002, with \textit{bla}\textit{IMP}-4 being carried on an IncC plasmid (Extended Data Fig. 1). It took approximately two years before further \textit{bla}\textit{IMP}-4-carrying Gram-negative bacteria were identified, and these were dominated by an IncC genetic setting or chromosomal \textit{bla}\textit{IMP}-4 in \textit{P. aeruginosa} (Fig. 1B). Notably, the first \textit{S. marcescens} lineage (here called lineage 1) with IncC-carrying \textit{bla}\textit{IMP}-4 became a successful lineage across the entire study period (Fig. 1B). \textit{S. marcescens} lineage 1 included 47 genomes from both clinical and environmental (intensive care unit [ICU] sinks) isolates. Based on their genetic relatedness (median pairwise single nucleotide variant [SNV] distance of 8 [IQR 4–11]), vertical transmission was most likely (Fig. 2A and Extended Data Table 2). The IncC plasmid was promiscuous and was rapidly noted in diverse bacterial hosts, with expansion from \textit{S. marcescens} into 13 other bacterial strains in 2004 alone. Ultimately, this plasmid was noted in 95 genomes across 37 strains (Fig. 2B), with \textit{S. marcescens} genomes accounting for the majority (52/95 [54.7\%]). Despite the diversity of IncC-carrying bacterial hosts over the 18 years, IncC plasmids were highly conserved, with only 11 SNVs across the plasmid backbone (median pairwise distance of 0 SNVs) (Extended Data Table 3) and almost all (70/71, 98.5\%) IncC plasmids belonged to the same MOB-typer primary/secondary clusters (AA860 AJ266, Extended Data Table 1). We noted seven IncC mosaic plasmids with additional replicon types that were excluded from further clustering analyses (Extended Data Table 4).
We used Mashtree to further characterise relatedness of the 64 non-mosaic IncC plasmids, which highlighted the similarity of plasmids across bacterial hosts and different time periods (Fig. 2B). This was also reflected in the \( bla_{IMP-4} \) flanking regions and integron sequences. Cluster analysis of the flanking regions up to 5000 bp upstream and downstream of \( bla_{IMP-4} \) using Flanker identified only four flanking regions (excluding singletons) in IncC plasmids, accounting for 67/71 (94%) genomes (Fig. 2B). The \( bla_{IMP-4} \)-containing integrons in the IncC plasmids were also highly similar, with a single SNV profile (TGGTCGACGCCT) accounting for 63/69 (91%) plasmids with intact integrons (Fig. 2B).

Taken together, these findings suggest that \( bla_{IMP-4} \)-containing IncC plasmids dominated the early outbreak period and established endemicity through their ability to rapidly spread across different bacterial hosts while maintaining stability. *S. marcescens* was a persistent host and reservoir for \( bla_{IMP-4} \) IncC during this time period (2002-2010) (Extended Data Fig. 1).

In addition to \( bla_{IMP-4} \) IncC plasmids, chromosomal integration of \( bla_{IMP-4} \) into *P. aeruginosa* ST111 (a global MDR lineage) was also a dominant feature of this early time period (Fig. 1B and Extended Data Fig. 1). These isolates were rapidly noted in nine patients in 2004 and continued to be isolated until 2018. The pseudomonal isolates were highly related (mean pairwise SNV distance of 1.6 SNVs vs 38.7 SNVs between Alfred and publicly available ST111 genomes \([P<0.001]\)) (Extended Data Fig. 2A) but the \( bla_{IMP-4} \)-containing integron and the flanking regions differed depending on the time period of isolation, and were also different to the integron and flanking region sequence of the IncC plasmids (Fig. 2C). Despite being temporally associated, the chromosomal integration of \( bla_{IMP-4} \) in *P. aeruginosa* ST111 with a different integron structure and flanking regions suggests that \( bla_{IMP-4} \) entry into *P. aeruginosa* likely arose independently of the \( bla_{IMP-4} \) IncC plasmids.

**Low endemicity of \( bla_{IMP-4} \) and entry into novel plasmids**

Apart from ongoing isolation of \( bla_{IMP-4} \) IncC plasmids (predominantly in *S. marcescens*) and chromosomal \( bla_{IMP-4} \) *P. aeruginosa*, the evolution of the following time period (2011-2015) was characterised by \( bla_{IMP-4} \) entering novel plasmids in *E. hormaechei* ST114 (a global MDR nosocomial *Enterobacter* clone) (Fig. 1B and Extended Data Fig. 1). At the start of the outbreak, there were a small number of *E. hormaechei* ST114 with \( bla_{IMP-4} \)-carrying IncC plasmids, but during this period, *E. hormaechei* ST114 acquired \( bla_{IMP-4} \)-carrying IncFIb and IncFIA/IncFIb/IncP plasmids (Fig. 1B).
Phylogenomic analysis showed that the *E. hormaechei* ST114 isolates were more diverse (median pairwise SNV distance 35, IQR 27–46) but the Alfred Hospital isolates still clustered more closely than other publicly available genomes (Fig. 3A and Extended Data Table 2). To determine if the IncC plasmids were the source of *bla*<sub>IMP</sub>-4 in the IncFIB and IncFIA/IncFIB/IncP plasmids, we analysed the *bla*<sub>IMP</sub>-4 flanking regions across the three plasmids (Fig. 3B). We noted homology of the 3850 bp upstream and 305 bp downstream regions of the integron in IncC, IncFIA/IncFIB/IncP and IncFIB plasmids (clusters A, B and C) with Tn3 transposons and DNA recombinases (*hin*) located immediately upstream. (Fig. 3B). These three plasmids also shared the same integron SNV profile (TGGTCGACGCCT) (Extended Data Fig. 3). The integron structure and flanking regions were different to the chromosomal *bla*<sub>IMP</sub>-4 genetic setting seen in *P. aeruginosa* (data not shown). Taken together, these findings suggested that as the *bla*<sub>IMP</sub>-4-harbouring IncC plasmids became endemic during the first time period (2002-2010), the outbreak evolved whereby the IncC plasmids served as a *bla*<sub>IMP</sub>-4 reservoir not only for inter-strain plasmid transfer, but also transfer of the *bla*<sub>IMP</sub>-4 integron and flanking regions between IncC, IncFIA/IncFIB/IncP and IncFIB plasmids.

Hyperendemicity and repeat outbreaks driven by clonal expansion of new bacterial strains and inter-species plasmid spread

The most recent time period (2016-2020) was characterised by complex, multi-level transmission dynamics resulting from the emergence of several new and highly successful *bla*<sub>IMP</sub>-4 plasmids and clonal expansion of *E. hormaechei* ST190 and ST93 host strains (Fig. 1B). Ongoing circulation of *bla*<sub>IMP</sub>-4 in genetic settings and bacterial strains from prior periods was also observed (Fig. 1B). Early in this period, *bla*<sub>IMP</sub>-4 was identified in a new plasmid, IncL/M, first in *E. hormaechei* ST114, which was its fourth *bla*<sub>IMP</sub>-4 carrying plasmid, and then in a wide range of other bacterial strains (n=16) (Fig. 1B). All IncL/M plasmids belonged to the same MOB-typer cluster, shared the same SNV profile in the integron (GGGTCGACGCCT) and 14/17 shared the same flanking cluster (Cluster G) (Fig. 3B). The three plasmids with other flanking clusters had minor variations in the Cluster G flanking region leading to them being clustered as singletons. These flanking regions were distinct from all other *bla*<sub>IMP</sub>-4 plasmids, however the same integron SNV profile was also noted in other plasmids (Extended Data Table 5 and Extended Data Fig. 3).
In 2017, blaIMP-4 was detected for the first time in IncHI2A plasmids in a small outbreak of Klebsiella oxytoca ST278 and Klebsiella michiganensis ST50 (Fig. 1B). The first IncHI2A plasmid (type 1 – as defined by MOB-typer cluster AA739 AJ055) then spread to E. hormaechei ST190 and a second IncHI2A plasmid (type 2 – MOB-typer cluster AA739 AJ058) emerged in E. hormaechei ST93, with both bacterial strains undergoing significant clonal expansion and contributing to a repeated outbreak and hyperendemicity from 2017-2020 (Fig. 1B and 4A). Bacterial isolates carrying blaIMP-4 on these two IncHI2A plasmids ultimately accounted for 98/161 (61%) of the sequenced genomes in that period (with 36/98 E. hormaechei ST93 and 43/98 E. hormaechei ST190) (Fig. 4A). The E. hormaechei ST93 and ST190 bacterial hosts were highly clonal with a median pairwise SNV distance of 9 (IQR 2-14) and 3 (IQR 2-4), respectively (Extended Data Fig. 2B and 2C, Extended Data Table 2). In addition to these two strains, the IncHI2A plasmids were found in 13 other strains (Fig. 4A).

Analysis of the blaIMP-4 flanking regions and integrons in the IncHI2A plasmids showed the same flanking sequence (Cluster F) across 36/47 (77%) plasmids and the same integron SNV profile (GGGTGACGCTCT) in 35/47 (74%) plasmids across both IncHI2A plasmid types (Fig. 4A). These flanking sequences and integron SNV profiles were not found in other plasmid types (Extended Data Fig. 3), suggesting that they may have arisen independently of other blaIMP-4 genetic settings.

To understand the rapid appearance of blaIMP-4 IncHI2A plasmids, we compared them to carbapenem-susceptible, non-blaIMP-4 IncHI2A plasmids in single E. hormaechei ST114, ST190 and ST93 genomes from our institution. The plasmids were highly similar between blaIMP-4 and non-blaIMP-4 bacterial strains (Fig. 4B). The IncHI2A type 1 plasmids from blaIMP-4 E. hormaechei ST190 had the addition of a 33 kbp region carrying the class 1 integron with blaIMP-4 (Fig. 4B). The IncHI2A type 2 plasmids from blaIMP-4 E. hormaechei ST93 shared 97.6% homology with the non-blaIMP-4 plasmid, with the blaIMP-4 integrase contained in a mosaic region (Figs. 4B and 4C). These data suggested integration of blaIMP-4 into pre-existing carbapenem-susceptible Enterobacter carrying IncHI2A plasmids, with mobilisation of the regions upstream and downstream. These flanking regions in IncHI2A plasmids were distinct to those in other blaIMP-4 plasmids, with a different IS110-like insertion sequence/Tn3 transposon combination (Fig. 4C).
While the arrival of new IncHI2A and IncL/M plasmids was the major contributor to the repeated outbreak and hyperendemicity during this period, we saw ongoing circulation of bla\textsuperscript{IMP-4} in prior genetic settings. In particular, we saw IncC plasmids circulating (32/161 genomes, 19.9%), including in a novel lineage of S. marcescens in 2019 (lineage 2 - Fig. 2A). We also noted bla\textsuperscript{IMP-4}-harboring P. aeruginosa ST111 (1 genome) and E. hormaechei ST114 with bla\textsuperscript{IMP-4} IncFIA/IncFIB/IncP plasmids (9 genomes). This reflected a cumulative trend where the prior bla\textsuperscript{IMP-4} genetic settings persisted in the context of new bacterial strains and plasmids, rather than waxing and waning over time.

**An environmental reservoir of bla\textsuperscript{IMP-4} plasmids**

In response to the outbreak in the most recent time period, we conducted environmental screening of ICU sinks from 2019–2020 and cultured 34 bla\textsuperscript{IMP-4} isolates (33 S. marcescens and one E. hormaechei), with 11 isolates selected for sequencing. Despite E. hormaechei predominating in bla\textsuperscript{IMP-4} clinical isolates, 6 genomes were S. marcescens lineage 1, 4 S. marcescens lineage 2, and one E. hormaechei ST190. These genomes closely matched clinical isolates, with bla\textsuperscript{IMP-4} being located on IncC plasmids and IncHI2A type 1 plasmids in S. marcescens and E. hormaechei ST190, respectively (Figs. 2B and 4A). This indicated that sinks were a possible reservoir for bla\textsuperscript{IMP-4} and may have explained the persistence of S. marcescens with closely matching IncC plasmids throughout the study (Fig. 2B).

**Within-patient and between-patient bla\textsuperscript{IMP-4} analyses show importance of diversification of genetic settings through plasmid transfer**

Patients with serial bla\textsuperscript{IMP-4}-harbouring genomes available followed different trajectories of bla\textsuperscript{IMP-4} carriage (Fig. 5A). Multiple colonisation events were noted in 4/41 patients, with bla\textsuperscript{IMP-4} being located in distinct genetic settings (i.e. differing plasmids and/or chromosomal integration). Flanking region/integron SNV profiles also differed, making within-patient integron transfer unlikely. Possible within-patient inter-strain transfer of key bla\textsuperscript{IMP-4} plasmids (IncC, IncL/M, IncHI2A types 1 and 2) occurred in 10/41 patients with the same bla\textsuperscript{IMP-4} plasmid types being noted in multiple strains (Fig. 5A). In 7 patients, the evidence was particularly compelling as plasmids had identical flanking sequences and integron SNV profiles across different bacterial hosts. Persisting colonisation was noted in 27 patients, with the same strain and same bla\textsuperscript{IMP-4} genetic setting repeatedly isolated.
We then used patient movement data (available in 127 patients from 2013 onwards) to establish putative transmission events (Fig. 5B), defined as spatiotemporal overlap between patients and genomic evidence of potential transmission. For genomic evidence, we considered both strain transmission (same bacterial host strain carrying \textit{bla}_{\text{IMP-4}} \text{ in the same genetic setting}) and plasmid transmission (detection of the same plasmid by MOB-typer cluster in different bacterial strains). We linked 71/127 (56%) patients using these definitions and identified the ICU as a major transmission site with 36/76 (47%) potential transmission events across 7/16 (44%) transmission networks, including the two largest networks (23 and 9 patients, respectively). While strain transmission contributed significantly, use of long-read sequencing to detect potential plasmid transmission allowed us to detect an additional 5/10 (50%) wards, 7/16 (44%) transmission networks and link 22/71 (31%) patients beyond what was identified for strain transmission alone (Fig. 5B). We measured betweenness centrality to identify key patients involved in transmission\textsuperscript{21}. The patient with the highest betweenness centrality (164.0 vs mean 4.9) had a >4-month ICU admission and was implicated in 7 transmission events, placing them at the centre of the large 23-patient network spanning those two strains and plasmid types (Fig. 5B). A further transmission event to a single patient then occurred during an admission on the Cardiology ward six months later.

**Discussion**

The spread of carbapenemases is the major driver of carbapenem resistance globally\textsuperscript{22} and has been the focus of numerous cross-sectional studies\textsuperscript{12,13,23}. To date, there have been limited efforts to study carbapenem resistance over extended time periods\textsuperscript{24,25}. In this study we had a unique opportunity to analyse two decades of \textit{bla}_{\text{IMP-4}} carbapenemases in our institution and gained important insights into how \textit{bla}_{\text{IMP-4}} caused outbreaks and perpetuated endemicity. \textit{bla}_{\text{IMP-4}} spread occurred through multiple mechanisms including strain transmission, plasmid transmission and transfer of the \textit{bla}_{\text{IMP-4}} class 1 integron. Each of these had a different qualitative and quantitative contribution to \textit{bla}_{\text{IMP-4}} persisting at our institution, highlighting that endemicity is a nuanced process requiring these mechanisms to act in concert. These findings carry important implications for prevention of future carbapenemase endemicity.
From a pathogen perspective, we propose that there are two key conditions required for blaIMP-4 endemicity. Firstly, there is a need for diversification of genetic settings for the resistance determinant, which in our study occurred through extensive inter-strain transmission of key blaIMP-4 plasmids (IncC, IncHIA2 type 1 and type 2, IncL/M), as well as mobilisation of the class I integron and entry into new plasmids (IncFIB, IncFIA/IncFIB/IncP). This ability to diversify led to the initial establishment of endemicity with IncC plasmids, and also to the period of hyperendemicity due to the emergence of a novel context in IncHI2A and IncL/M plasmids. To study this emergence, we demonstrated that blaIMP-4 IncHI2A plasmids were highly similar to non-blaIMP-4 IncHI2A plasmids in E. hormaechei ST93, ST114 and ST190 that may have served as acceptors for the blaIMP-4 integron. In addition, there was possible importation from outside sources: blaIMP-4 has been found in both IncHI2A and IncL/M plasmids in Australian isolates\textsuperscript{2,4,9,26} and blaIMP-4-harbouring IncHI2A plasmids are emerging as a global issue, having been noted in a recent multi-hospital outbreak in the United Kingdom\textsuperscript{27}.

The second condition is propagation of blaIMP-4 through the establishment of high-risk strain-genetic setting pairs we term 'propagators'. While we noted blaIMP-4 in 68 strains during the study, five strains accounted for 190/270 (70\%) genomes and 140/196 (71\%) patients colonised with blaIMP-4. The first of these was S. marcescens lineage 1 (IncC plasmids), which defined the early period of the study and continued to persist throughout. This propagator pair was able to act as a reservoir of blaIMP-4, likely through occupying an environmental niche as we noted during sampling of ICU sinks. The colonisation of hospital plumbing by CPOs has been well documented\textsuperscript{9,28-31}, including blaIMP-4-harbouring S. marcescens in an Australian setting that was unable to be eradicated\textsuperscript{32}. This environmental colonisation enabled not only clonal spread of blaIMP-4 harbouring S. marcescens but also drove diversification through inter-strain transfer of blaIMP-4 IncC plasmids and inter-plasmid transfer of the blaIMP-4 class I integron. Other propagators emerged at various junctures including P. aeruginosa ST111 (chromosome) and E. hormaechei ST114 (IncFIA/IncFIB/IncP plasmids), then E. hormaechei ST190 and ST93 (IncHI2A type 1 and type 2 plasmids, respectively). Clonal spread of propagators was therefore central to establishing and maintaining blaIMP-4 endemicity, as well as leading to a repeated outbreak and hyperendemicity in the final period of the study. These findings broadly fit the ‘multiple lineages, multiple plasmids’ designation proposed by David et al. when analysing carbapenemase spread in K. pneumoniae\textsuperscript{12} but we demonstrated that the dynamics of
Carbapenemase endemicity in our setting were significantly more complex with clonal transmission of propagator strains, inter-strain plasmid transmission and inter-plasmid integron transmission all playing important roles.

In addition to pathogen factors, we were able to analyse patient factors. While genomic surveillance previously focused on lineage-level analysis, long-read sequencing technologies have improved analysis of plasmids and other mobile genetic elements. In our study, these insights proved informative both for understanding within-patient and between-patient blaIMP-4 spread. Within patients, we detected different trajectories of colonisation. Patients who undergo multiple colonisation events may be at the core of multiple transmission networks, as demonstrated by the patient colonised with two IncHI2A plasmids and highest betweenness centrality. Patients with inter-strain plasmid transfer may facilitate the diversification of genetic settings for blaIMP-4, thus increasing the risk of newly successful propagators emerging, in turn fuelling outbreaks. Of note, we did not find clear evidence of blaIMP-4 integron transfer events within-patients, suggesting that they may play a lesser role. Long-read technologies also allowed us to analyse putative plasmid transmission between-patients, which implicated an additional 50% of wards, 44% transmission networks and 31% patients over strain transmission alone. We used detection of blaIMP-4 plasmids of the same MOB-typer primary/secondary clusters as a simple definition but quantitative thresholds incorporating changes in plasmid backbones and large-scale recombination events across a diverse array of bacterial hosts, plasmids and resistance determinants are needed.

Our study had several limitations. Firstly, it was based on an isolate collection that spanned two decades with some patient data from the early part of the study being incomplete. Similarly, approaches to isolate sampling changed during that time, in particular since the inception of a statewide CPO detection program that mandated screening in high-risk areas and likely led to increased detection of CPO colonisation in the last 4 years of the study. Finally, our study focused on blaIMP-4 and was from a single centre, which may limit the generalisability of findings to outbreaks at other centres.
In summary, we showed that blaIMP-4 endemicity and repeated outbreaks were due to diversification of genetic settings through inter-strain blaIMP-4 plasmid transfer and inter-plasmid blaIMP-4 integron transfer in combination with clonal expansion that led to an evolving cascade of high-risk strain-genetic setting pairs. Our findings provide a framework for understanding endemicity of MBL-producing organisms and may have broader applicability to other CPOs. Our study highlights that stopping the spread of CPOs will require adequate surveillance to detect not only the presence of resistance determinants and their bacterial host strains but also their genetic context and plasmid-integron transmission dynamics, thus enabling early detection of novel and potentially hidden threats.
Methods

Isolate selection

The study was reviewed and approved by the Alfred Hospital Ethics Committee. We systematically reviewed an institutional collection spanning all CPO isolates from 2002 to 2020. The collection contained isolates collected as part of routine clinical care, as well as environmental screening of sinks from 2018-2020. Routine antimicrobial susceptibility testing was performed using Vitek2 (BioMérieux). We identified \( \text{bla}^{\text{IMP-4}} \) carriage through polymerase chain reaction (PCR) screening. GoTaq Flexi DNA polymerase (Promega, Wisconsin, USA) was used as per manufacturer’s instructions and 10 µmol of the primers Imp4_screen_F (5′-CCAGGACACACTCCAGATAACC-3′) and Imp4_screen_R (5′-CAAGAGTGATGCGCTCCAGC-3′) in 25 µL reaction volumes. PCR was performed using the following cycle conditions: 98 °C for 2 min, followed by 30 cycles of 98 °C for 30 sec, 55 °C for 30 sec, 72 °C for 30 sec. Amplicons were resolved by agarose gel electrophoresis on a 1% w/v agarose gel.

We selected 277 \( \text{bla}^{\text{IMP-4}} \) isolates for whole genome sequencing (WGS) based on bacterial strain (species/MLST combination) and year of isolation. For species with <30 isolates, we sequenced all available isolates. For species with >30 isolates, we performed WGS on selected isolates based on collection date to ensure that we had sequencing data available for all study periods. We sequenced at least one isolate of all strains across all study periods with both short-read (Illumina) and long-read (Oxford Nanopore) technologies (n=172). We also selected one carbapenem-susceptible \( E. \) hormaechei ST93 and one \( E. \) hormaechei ST190 isolate for short- and long-read WGS.

Culture, DNA extraction and sequencing

All bacterial isolates were grown on cation-adjusted Mueller-Hinton II agar (Becton-Dickinson) for 16 hours at 37°C, and sub-cultured into cation-adjusted Mueller-Hinton broth (Becton-Dickinson) for a further 16 hours at 37°C. Bacterial genomic DNA was extracted from liquid culture using the GenFind V3 Reagent Kit (Beckman Coulter) as per manufacturer’s instructions. Libraries for short read sequencing were prepared using the Nextera Flex DNA Library Prep Kit (Illumina), and 150 bp paired-end sequencing was performed on the NovaSeq 6000 system (Illumina). Libraries for long-read sequencing were prepared using the Ligation Sequencing Kit with Native Barcoding Expansion.
(Oxford Nanopore Technologies) and sequenced on the MinION instrument with an R9.4.1 flow cell
for 48 hours. Basecalling was performed with Guppy v.4.0.14 using
the ‘high accuracy’ basecalling model.

De novo assembly and annotation
We constructed de novo assemblies of all isolates with only short-read data using the Shovill wrapper
for SPAdes, which also utilizes Trimmomatic for read trimming and Pilon for read error correction. For long-read assembly, long reads were filtered using Filtlong v.0.2.0 with the following
parameters: ‘--min_length 1000 --keep_percent 90 --target_bases 500000000’. Hybrid assemblies
incorporating short- and long-read data were created using Unicycler v.0.4.08 with standard
parameters with Unicycler output used to assess circularisation. If blaIMP-4 contigs were non-
circularised, we re-assembled genomes using a long-read-first assembly using a bespoke pipeline
(https://github.com/HughCottingham/clinopore-nf) that incorporates Flye with subsequent polishing
with Medaka, Polypolish and Polca. Assembly quality was checked using Quast and species
identification was performed using GTDB-Tk and checked against isolate identification performed at
time of isolate collection. We excluded genomes (n=7) with a species mismatch, as well as genomes
whose assemblies had >1000 contigs, N50<10000 or assembly length >7.5 Mb.

On the remaining assemblies (n=270), we annotated the genomes using Prokka v1.14.6. We then
performed resistance gene and plasmid replicon detection with Abricate v.1.0.0, using the NCBI
Antibiotic Resistance and PlasmidFinder databases, respectively. We determined in silico multi-
locus sequence type (ST) using ‘mlst’ v.2.19.0. All inconclusive ST calls with ‘mlst’ were checked with
SRST2.

Core genome-based phylogenetic analyses
We performed core genome-based phylogenetic analyses on key STs, defined as those with ≥5
isolates available from our institution. This included E. hormaechei ST93, ST114, and ST190 and P.
aeruginosa ST111. Due to the absence of an MLST schema for S. marcescens we identified all
RefSeq S. marcescens genomes and used Assembly Dereplicator
(https://github.com/rrwick/Assembly-Dereplicator) with a Mash distance threshold of 0.001 to remove
duplicate assemblies. We then used these assemblies, in conjunction with *S. marcescens* genomes from our institution to construct a phylogeny using Mashtree\(^7\). In brief, this tool uses non-alignment based assessment of sequence similarity through use of the min-hash algorithm, as implemented in Mash\(^52\), to generate distance metrics between input sequences. These are then used to cluster sequences using the neighbor joining algorithm. This allowed us to identify that Alfred Hospital genomes belonged to two lineages, for which we conducted the same phylogenetic analyses as we did within STs for other species.

This consisted of identifying RefSeq genomes of the same ST and including them for context in phylogenetic analyses. We chose one completed, closed assembly from our institution for each ST to use as a reference. Mobile genetic elements were excluded from these reference assemblies using PHASTER and IslandViewer 4\(^53,54\). A core chromosomal SNV alignment was generated using Snippy v.4.6.0\(^55\) and recombination was removed using Gubbins\(^56\). We then used this core genome alignment in IQtree v.2.0.3 to generate maximum likelihood phylogenies for each ST\(^57\), with the best-fit model chosen using ModelFinder\(^58\). For each ST, median SNV distances between isolates from our institution were then calculated. Phylogenetic trees were visualized and annotated with metadata using 'ggtree'\(^59\) with additional editing in Adobe Illustrator v2020.24.3.

**Plasmid phylogenetic analyses**

Using Abricate, we identified *bla*\(^\text{IMP-4}\)-harbouring contigs that were putative plasmids in our hybrid assemblies. We then used the MOB-typer tool to determine plasmid replicons present, as well as to assign clusters\(^16\). In addition, we used COPLA to assign plasmid taxonomic units to key plasmid types as determined by MOB-typer\(^60\). We identified possible mosaic plasmids resulting from fusion events by examining plasmid replicon content within MOB-typer cluster and identifying plasmids which had presence of additional plasmid replicons then manually inspecting the assemblies.

We then conducted analyses within key plasmid groups within our dataset, as determined by MOB-typer cluster. These included IncC, IncHI2A type 1, IncHI2A type 2, IncFIA/IncFIB/IncP, IncFIB, IncL/M, and untypeable plasmids from *Acinetobacter* spp. In order to identify SNVs in the plasmid backbone, we used Snippy v.4.6.0\(^55\) to create a core SNV alignment by mapping short reads to a
reference plasmid from our institution from each plasmid group. We then used Mashtree\textsuperscript{17} to generate distance metrics between plasmids belonging to the same group, excluding mosaic plasmids. The R package ‘ggtree’ v3.0.4 was used to visualize the resulting trees\textsuperscript{59} and to annotate with metadata. Adobe Illustrator v2020.24.3 was used to merge different parts of the figures together. We also used fastANI v1.3 to generate pairwise average nucleotide identities between plasmids belonging to the same plasmid group\textsuperscript{61}. We used progressiveMauve to align all plasmids within a plasmid group and assess for structural re-arrangements\textsuperscript{62}, then visualized this in Easyfig v2.2.2\textsuperscript{63}.

**Analysis of bla\textsuperscript{IMP-4} integron and flanking sequences**

We used Flanker\textsuperscript{18} to identify and cluster flanking sequences around bla\textsuperscript{IMP-4} from hybrid contigs. We performed clustering 5000 bp upstream and downstream of the bla\textsuperscript{IMP-4} gene across windows in 500 bp increments. Geneious v10.2.6 (https://www.geneious.com) was used to visualize and assess for structural re-arrangements, with subsequent manual annotation in Adobe Illustrator v2020.24.3. We also assessed for SNVs in the bla\textsuperscript{IMP-4} integron by aligning bla\textsuperscript{IMP-4} genetic settings from completed, circularised assemblies to a previously reported bla\textsuperscript{IMP-4} Integron (GenBank accession number JX101693)\textsuperscript{4} using MUSCLE v3.8.1551\textsuperscript{64}. Assemblies with large scale insertions or deletions in the integron were excluded (e.g. P. aeruginosa ST111 genomes). We extracted SNVs from the resulting alignment using SNP-sites v2.5.1\textsuperscript{65} and grouped plasmids according to the SNV profile.

**Short-read mapping to plasmid sequences**

For genomes which only had short-read data available, we created a database of plasmids from all MOB-typer clusters (described above) and used the Nextflow implementation of the REDDog pipeline (V1.beta10.3; available at https://github.com/scwatts/reddog-nf) to map short-reads to this database. We used the following parameters: ‘mapping_cover_min = 1, mapping_mapped_min = 0.5, mapping_depth_min = 10’ then analysed the data. A read set was considered to have a match to a plasmid in the database if there was >90% coverage of the plasmid with <10 SNVs.

**Patient data and transmission events**

Clinical data were extracted from the electronic medical record. Clinical data were missing for 7 isolates from 2009-2012. Patient movement data were available from 2013 onwards, including
127/196 (65%) patients in the study. As patients did not undergo systematic surveillance for bla\text{IMP-4}, we considered that the patient may have been colonised in the 30 days prior to the first isolation of a bla\text{IMP-4}-harbouring organism and identified overlaps on the same ward at the same time as potential transmission events between patients. We then applied genomic criteria to further confirm potential transmission events. In the first instance, patients would have to have bla\text{IMP-4}-harbouring bacteria of the same strain for a potential transmission event to be considered. These events were then further classified on basis of SNV distance, with a cutoff of 10 SNVs. In the second instance, patients would have to have bla\text{IMP-4} in the same genetic setting (defined as the same MOB-typer primary/secondary cluster), as determined either through completed assemblies or by having a match to a reference plasmid using the short-read mapping approach described above. We then used the R package 'ggraph' v2.0.5 to visualize putative transmission networks with patients as nodes and potential transmission events as edges. Betweenness centrality was calculated using the ‘betweenness’ function in the ‘igraph’ R package v1.2.11.66

Data availability
Illumina/Nanopore read data were deposited in the NCBI SRA under project accession PRJNA924056. Completed genome assemblies were deposited in GenBank; accessions are listed in Supp. Table 1.

Statistical analysis
Categorical variables were compared using χ2 or Fisher’s exact tests and continuous variables were compared using Student’s t-test or Mann-Whitney-Wilcoxon, as appropriate. Statistical analyses were performed in R (v4.1.1).
References

Bacterial Plasmids Confering Carbapenem Resistance. 

Stoesser, N. et al. Genomic Epidemiology of Complex, Multispecies, Plasmid-Borne bla(KPC)


Abrate: mass screening of contigs for antimicrobial resistance or virulence genes (2020).


Fig. 1 – Summary of species, multi-locus sequence types and genetic settings of \textit{bla}\textsubscript{IMP}-4 in sequenced isolates

Fig. 1A: There was a diversity in both the bacterial hosts and the genetic settings of \textit{bla}\textsubscript{IMP}-4 with \textit{bla}\textsubscript{IMP}-4 being noted in 7 bacterial genera and 68 bacterial host strains.

Fig. 1B: We noted \textit{bla}\textsubscript{IMP}-4 temporal trends that allowed definition of three periods. In the first period, \textit{bla}\textsubscript{IMP}-4 rapidly emerged in a promiscuous IncC plasmid and in \textit{P. aeruginosa} ST111. The second period was defined by low-level endemicity with the entry of \textit{bla}\textsubscript{IMP}-4 on a class 1 integron into novel IncFIA/IncFIB/IncP plasmids in \textit{E. hormaechei} ST114. In the final part of the study, a period of hyperendemicity was driven by the emergence of two types of IncHI2A plasmids in \textit{E. hormaechei} ST93 and ST190, as well as \textit{bla}\textsubscript{IMP}-4 IncL/M plasmids in multiple bacterial hosts.

Abbreviations: No. – number; Lin. – lineage; ST – sequence type.
**Fig. 2**  - *bla*\textsuperscript{IMP-4} endemicity established in 2002-2010 due to promiscuous *bla*\textsuperscript{IMP-4} IncC plasmid and *bla*\textsuperscript{IMP-4} entry into *Pseudomonas aeruginosa* ST111

**A**

**B**

**C**

**Fig. 2A:** *S. marcescens* was a persistent host and propagator of *bla*\textsuperscript{IMP-4} on IncC plasmids throughout the entire study period. The inset panel locates *S. marcescens* genomes from the Alfred Hospital in a global *S. marcescens* phylogeny, showing that Alfred Hospital genomes form two distinct lineages, which are then shown in the outer panel. These two *S. marcescens* lineages were detected both in clinical and environmental samples.

**Fig. 2B:** Analysis of IncC plasmids showed promiscuity that enabled entry into 37 bacterial strains. Despite this promiscuity, the IncC plasmid remained stable with four flanking region clusters and one integron SNV profile accounting for 93% and 90% plasmids, respectively.

**Fig. 2C:** Three flanking clusters were detected in *P. aeruginosa* ST111. Differences are shown to the class 1 integron present in contemporaneous *bla*\textsuperscript{IMP-4} IncC plasmids (top), with deletion of *catB3* in flanking cluster 1 (in genomes from 2004-2012), deletion of *qacG, aac4* and *catB3* in flanking cluster 2 (2014-2015) and deletion of *catB3* with distinct flanking regions in flanking cluster 3 (2018).

Abbreviations: bp – base pairs; Lin. – lineage; SNV – single nucleotide variant; ST – sequence type.
Fig. 3 – *E. hormaechei* ST114 was a host of *bla*IMP-4 in diverse genetic settings including due to integron transfer of *bla*IMP-4.

**A**

<table>
<thead>
<tr>
<th>Year</th>
<th>Genetic setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>n=36</td>
</tr>
<tr>
<td>2019</td>
<td>n=25</td>
</tr>
<tr>
<td>2019</td>
<td>n=6</td>
</tr>
<tr>
<td>2019</td>
<td>n=7</td>
</tr>
<tr>
<td>2013</td>
<td>n=8</td>
</tr>
<tr>
<td>2016</td>
<td>n=6</td>
</tr>
<tr>
<td>2019</td>
<td>n=25</td>
</tr>
<tr>
<td>2019</td>
<td>n=40</td>
</tr>
<tr>
<td>2019</td>
<td>n=25</td>
</tr>
<tr>
<td>2019</td>
<td>n=40</td>
</tr>
<tr>
<td>2019</td>
<td>n=6</td>
</tr>
<tr>
<td>2019</td>
<td>n=14</td>
</tr>
<tr>
<td>2012</td>
<td>7x10^5</td>
</tr>
</tbody>
</table>

**B**

- **Genetic setting**
  - *Chromosome*
  - *IncC*
  - *IncFIA/IncFIB/IncP*
  - *IncHI2A type 1*
  - *IncHI2A type 2*
  - *IncL/M*
  - *Class 1 integron*
  - *intI1*
  - *qacG*
  - *aacA4*
  - *catB3*
  - *sul1*
  - *Tn3*
  - *IS6*
  - *yafQ*
  - *hin*
  - *ligA*
  - *hyp*
  - *repA*
  - *sapA*
  - *srpC*
  - *IS110*
  - *rhsC*

**Fig. 3A** – The inset panel places *E. hormaechei* ST114 genomes from the Alfred Hospital in a global *E. hormaechei* ST114 phylogeny, showing that they were clonally related but showed more diversity than other key *bla*IMP-4 bacterial host strains. *E. hormaechei* were able to act as versatile acceptors of *bla*IMP-4 in different genetic settings, including four different plasmid types throughout the study.
Fig. 3B – We conducted a clustering analysis of flanking regions 5000 bp upstream and downstream of *bla*IMP-4. The six key flanking region clusters with >5 genomes are shown on the left and the genetic context is shown on the right. IncFIB and IncFIA/IncFIB/IncP plasmids clustered with IncC plasmids throughout the analysis and shared homology both upstream and downstream of *bla*IMP-4.

Abbreviations: bp. – base pairs; SNV – single nucleotide variant.
Fig. 4 – IncHI2A plasmids drove a period of \textit{bla}\textsubscript{IMP}-4 hyperendemicity from 2016-2020

A

Bacterial strain
- Citrobacter spp.
- Enterobacter cloacae complex - Other STs
- Enterobacter hormaechei ST190
- Enterobacter hormaechei ST114
- Escherichia coli - Other STs
- Klebsiella michiganensis
- Klebsiella oxytoca ST278
- Klebsiella pneumoniae complex - Other STs
- E. hormaechei ST190
- E. hormaechei ST93

Source
- Clinical
- Environmental

Genetic setting
- IncHI2A type 1
- IncHI2A type 2

Flanking cluster
- Absent
- Present

Integron SNV group
- GGGATGACGTCT
- GGGTCGACACCT
- GGGTCGACGCCT
- GGGTCGACGTCT
- GGGTCGATGTCT
- GGGTCGAACCT

C

Enterobacter hormaechei ST190
Enterobacter hormaechei ST114
Enterobacter hormaechei ST93

Fig. 4A – Relatedness analysis of IncHI2A plasmids using Mashtree showed two distinct plasmid types that circulated during the same period and rapidly entered 15 bacterial strains. These plasmids shared flanking regions and integron SNV profiles, which were distinct to those noted in other plasmid types. \textit{E. hormaechei} ST190 and ST93 were propagators for IncHI2A type 1 and type 2 plasmids, respectively.
**Figs. 4B and 4C** – We compared *bla*IMP-4-carrying IncHI2A plasmids to non-*bla*IMP-4 IncHI2A plasmids from *E. hormaechei* ST190, ST93 and ST114. IncHI2A type 1 plasmids from *E. hormaechei* ST190 had homology over 100% of the non-*bla*IMP-4 plasmid, with the addition of a 33 kbp region carrying the class 1 integron in the *bla*IMP-4 plasmid. IncHI2A type 2 plasmids from *E. hormaechei* ST93 shared homology across 97.6% of the non-*bla*IMP-4 plasmid, with the *bla*IMP-4 integron contained in a mosaic region.

Abbreviations: bp – base pairs; SNV – single nucleotide variant; ST – sequence type.
Fig. 5 – Within- and between-patient spread of *bla*IMP-4

**A**

Different trajectories of *bla*IMP-4 colonisation were identified through use of completed, circularised genomes. While the majority of patients had persistent colonization with the same bacterial strains (indicated by point colours) harbouring *bla*IMP-4 in the same genetic setting (indicated by point shapes), potential inter-strain plasmid transfer events occurred in 10 patients with *bla*IMP-4 being noted on the same plasmids in different bacterial strains. In a further four patients, there was evidence of multiple colonization events, as determined by *bla*IMP-4 being found in distinct genetic settings.

**B**

Potential transmission events between patients were identified with individual patients acting as vertices. Edges were drawn on the basis of a combination of spatiotemporal overlap on a hospital wards (indicated by edge colour) and a genomic criterion. In the left panel, the genomic criterion was having the same *bla*IMP-4 bacterial strain with SNV distance from phylogenetic analysis indicated by...
shading. In the right panel, the genomic criterion incorporated both the strain transmission analysis as well as detecting presence of \textit{blaIMP-4} in the same plasmid (as defined by MOB-typer cluster) in different bacterial strains. The additional benefit of plasmid analysis is indicated by the blue shading of the potential transmission events in this panel. The red shading indicates the patient with the highest betweenness-centrality and their contacts.

Abbreviations: ICU – intensive care unit; Lin. – lineage; SNV – single nucleotide variant; ST – sequence type.
### Extended Data Table 1 – bla\textsuperscript{IMP-4} plasmid types with corresponding MOB-typer clusters and plasmid taxonomic units

<table>
<thead>
<tr>
<th>Plasmid type</th>
<th>MOB-typer primary cluster</th>
<th>MOB-typer secondary cluster</th>
<th>Plasmid taxonomic unit (PTU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IncC</td>
<td>AA860</td>
<td>AJ266</td>
<td>PTU-C</td>
</tr>
<tr>
<td>IncFIA/IncFIB/IncP</td>
<td>AD444</td>
<td>AN003</td>
<td>PTU could not be assigned</td>
</tr>
<tr>
<td>IncFIB</td>
<td>AA020</td>
<td>AH567</td>
<td>PTU could not be assigned</td>
</tr>
<tr>
<td>IncHI2A type 1</td>
<td>AA739</td>
<td>AJ055</td>
<td>PTU-HI2</td>
</tr>
<tr>
<td>IncHI2A type 2</td>
<td>AA739</td>
<td>AJ058</td>
<td>PTU-HI2</td>
</tr>
<tr>
<td>IncL/M</td>
<td>AA002</td>
<td>AH532</td>
<td>PTU-L/M</td>
</tr>
<tr>
<td>Untypeable</td>
<td>AC935</td>
<td>AM306</td>
<td>PTU could not be assigned</td>
</tr>
</tbody>
</table>
**Extended Data Table 2 – Phylogenetic distance analysis of key \textit{bla}\textsubscript{IMP-4} bacterial host strains from the Alfred Hospital**

<table>
<thead>
<tr>
<th>Bacterial strain</th>
<th>Mean pairwise SNV distance</th>
<th>Median pairwise SNV distance</th>
<th>Interquartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Enterobacter hormaechei} ST114</td>
<td>39.3</td>
<td>35</td>
<td>27-46</td>
</tr>
<tr>
<td>\textit{Enterobacter hormaechei} ST190</td>
<td>13.6</td>
<td>3</td>
<td>2-4</td>
</tr>
<tr>
<td>\textit{Enterobacter hormaechei} ST93</td>
<td>9.6</td>
<td>9</td>
<td>2-14</td>
</tr>
<tr>
<td>\textit{Klebsiella oxytoca} ST278</td>
<td>2</td>
<td>2</td>
<td>1-3</td>
</tr>
<tr>
<td>\textit{Klebsiella pneumoniae} ST4379</td>
<td>31.2</td>
<td>31.5</td>
<td>14.25-48.75</td>
</tr>
<tr>
<td>\textit{Pseudomonas aeruginosa} ST111</td>
<td>1.6</td>
<td>1</td>
<td>0-2</td>
</tr>
<tr>
<td>\textit{Serratia marcescens} lineage 1</td>
<td>7.8</td>
<td>8</td>
<td>4-11</td>
</tr>
<tr>
<td>\textit{Serratia marcescens} lineage 2</td>
<td>1</td>
<td>1</td>
<td>1-1</td>
</tr>
</tbody>
</table>

Abbreviations: SNV – single nucleotide variant, ST – sequence type.
Extended Data Table 3 – Distance metrics for key *bla*IMP-4 plasmid types

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mash distances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median distance</td>
<td>2.4444×10^{-3}</td>
<td>99.33×10^{-55}</td>
<td>99.49×10^{-4}</td>
<td>33.07×10^{-4}</td>
<td>5.18×10^{-4}</td>
<td>77.91×10^{-3}</td>
</tr>
<tr>
<td>Mean distance</td>
<td>4.0303×10^{-3}</td>
<td>2562.56×10^{-22}</td>
<td>1.49×10^{-3}</td>
<td>44.37×10^{-4}</td>
<td>7.35×10^{-4}</td>
<td>66.17×10^{-22}</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>6.8989×10^{-4}–5.66×10^{-3}</td>
<td>5.57×10^{-5}–9.58×10^{-22}</td>
<td>2.0303×10^{-4}–1.4646×10^{-3}</td>
<td>11.95×10^{-4}–6×10^{-3}</td>
<td>1.09×10^{-3}</td>
<td>66.00×10^{-3}–00.143</td>
</tr>
<tr>
<td><strong>SNV analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total core alignment SNV sites</td>
<td>11</td>
<td>2</td>
<td>25</td>
<td>5</td>
<td>5</td>
<td>1122</td>
</tr>
<tr>
<td>Median pairwise SNV distance</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1013.5</td>
</tr>
<tr>
<td>Mean pairwise SNV distance</td>
<td>0.39</td>
<td>0.25</td>
<td>1.88</td>
<td>0.33</td>
<td>0.69</td>
<td>718.83</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>0–1</td>
<td>0–0</td>
<td>0–4</td>
<td>0–1</td>
<td>0–1</td>
<td>373–1034.5</td>
</tr>
<tr>
<td><strong>Pairwise average nucleotide identity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median nucleotide identity</td>
<td>99.83%</td>
<td>99.9595%</td>
<td>99.95%</td>
<td>99.9797%</td>
<td>99.9595%</td>
<td>99.3434%</td>
</tr>
<tr>
<td>Mean nucleotide identity</td>
<td>99.7777%</td>
<td>982398.23%</td>
<td>99.9393%</td>
<td>99.97%</td>
<td>99.9494%</td>
<td>99.4646%</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>99.7373%–99.8989%</td>
<td>95.48%–99.98%</td>
<td>99.8888%–99.9979%</td>
<td>99.96%–99.9989%</td>
<td>100%</td>
<td>99.80%</td>
</tr>
</tbody>
</table>

Abbreviations: SNV – single nucleotide variant
## Extended Data Table 4 – Characteristics of \textit{bla}^{\text{IMP-4}} mosaic plasmids

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Bacterial host strain</th>
<th>Flanking region cluster</th>
<th>Integron SNV profile</th>
<th>\textit{bla}^{\text{IMP-4}} genetic setting</th>
<th>Plasmid replicon types present</th>
<th>MOB-typer primary cluster</th>
<th>MOB-typer secondary cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPO136</td>
<td>\textit{Klebsiella pneumoniae} ST340</td>
<td>14</td>
<td>N/A</td>
<td>IncC</td>
<td>IncC, IncFIB</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO150</td>
<td>\textit{Klebsiella pneumoniae} ST340</td>
<td>A</td>
<td>GCATCGACGCCT</td>
<td>IncC</td>
<td>IncC, IncFIB</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO161</td>
<td>\textit{Klebsiella pneumoniae} ST4379</td>
<td>11</td>
<td>N/A</td>
<td>IncC</td>
<td>IncC, ColRNA (rep cluster 1987), IncP</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO328</td>
<td>\textit{Klebsiella variicola} ST1582</td>
<td>E</td>
<td>TGGTCGACGCCT</td>
<td>IncC</td>
<td>IncC, IncN</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO422</td>
<td>\textit{Serratia marcescens} lin. 1</td>
<td>A</td>
<td>TGGTCGACGCCT</td>
<td>IncC</td>
<td>IncC, ColRNA (rep cluster 1987)</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO475</td>
<td>\textit{Serratia marcescens} lin. 1</td>
<td>B</td>
<td>TGGTCGACGCCT</td>
<td>IncC</td>
<td>IncC, rep cluster 2335</td>
<td>AA860</td>
<td>AJ266</td>
</tr>
<tr>
<td>CPO504</td>
<td>\textit{Klebsiella pneumoniae} ST1552 (1LV)</td>
<td>A</td>
<td>TGGTCGACGCCT</td>
<td>IncC</td>
<td>IncC, IncFIA, IncFIB, IncFIC</td>
<td>AA861</td>
<td>AJ282</td>
</tr>
<tr>
<td>CPO044</td>
<td>\textit{Enterobacter hormaechei} ST190</td>
<td>F</td>
<td>GGGTCACTGCTCT</td>
<td>IncHI2</td>
<td>IncHI2A, IncR, rep cluster 1088</td>
<td>AA739</td>
<td>AJ055</td>
</tr>
<tr>
<td>CPO062</td>
<td>\textit{Enterobacter hormaechei} ST190</td>
<td>7</td>
<td>GGGTCACTGCTCT</td>
<td>IncHI2</td>
<td>IncHI2A, IncR, rep cluster 1088</td>
<td>AA739</td>
<td>AJ055</td>
</tr>
<tr>
<td>CPO235</td>
<td>\textit{Escherichia coli} ST155</td>
<td>F</td>
<td>GGATGACGTCT</td>
<td>IncHI2</td>
<td>IncHI2A, IncX1, rep cluster 1088</td>
<td>AA739</td>
<td>AJ058</td>
</tr>
</tbody>
</table>

Abbreviations: lin. – lineage; LV – locus variant; ST – Sequence type; N/A – not applicable due to insertions or deletions in integron structure
Extended Data Table 5 – Integron single nucleotide variant profiles in *bla*~IMP-4~ plasmids

<table>
<thead>
<tr>
<th>Integron SNV Profile</th>
<th>Total plasmids</th>
<th>Genetic setting associated with integron SNV profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGTCGACGCCT</td>
<td>75</td>
<td>Chromosome, IncC, IncFIA/IncFIB/IncP, IncFIB</td>
</tr>
<tr>
<td>GGGTCGACGTCT</td>
<td>38</td>
<td>Chromosome, IncHI2A type 1, IncHI2A type 2</td>
</tr>
<tr>
<td>GGGTCGACGCCT</td>
<td>25</td>
<td>Chromosome, IncC, IncHI2A type 1, IncL/M</td>
</tr>
<tr>
<td>GGGTCGACACCT</td>
<td>4</td>
<td>IncHI2A type 1</td>
</tr>
<tr>
<td>GGGATGACGTCT</td>
<td>2</td>
<td>IncHI2A type 1, IncHI2A type 2</td>
</tr>
<tr>
<td>GCATCGACGCCT</td>
<td>1</td>
<td>IncC</td>
</tr>
<tr>
<td>TGGTCGGCGCCT</td>
<td>1</td>
<td>IncC</td>
</tr>
<tr>
<td>TGGTTGACGCCT</td>
<td>1</td>
<td>IncC</td>
</tr>
<tr>
<td>TGGTCGACGCTC</td>
<td>1</td>
<td>IncFIA/IncFIB/IncP</td>
</tr>
<tr>
<td>GGGTGCATGTCT</td>
<td>1</td>
<td>IncHI2A type 1</td>
</tr>
<tr>
<td>GGGATTACGTCT</td>
<td>1</td>
<td>IncHI2A type 2</td>
</tr>
<tr>
<td>GGGTCTACGTCT</td>
<td>1</td>
<td>IncHI2A type 2</td>
</tr>
</tbody>
</table>

Abbreviations: SNV – single nucleotide variant.
Extended Data Fig. 1 – Temporal trends of \textit{bla\textsc{IMP-4}} genetic settings over two decades

\textit{bla\textsc{IMP-4}} was initially noted in an outbreak in IncC plasmids and in clonal \textit{P. aeruginosa} ST111 with \textit{bla\textsc{IMP-4}} chromosomal integration. This was followed by low-level endemicity and diversification of \textit{bla\textsc{IMP-4}} genetic settings into novel IncFIA/IncFIB/IncP plasmids in \textit{E. hormaechei} ST114. We then noted a period of hyperendemicity driven by the emergence of two types of IncHI2A plasmids, as well as \textit{bla\textsc{IMP-4}} IncL/M plasmids in multiple bacterial hosts. During this time, we saw ongoing circulation of \textit{bla\textsc{IMP-4}} in prior genetic settings.

Abbreviations: No. – number.
We conducted a phylogenetic analysis of the key $bla_{\text{IMP}-4}$ bacterial host strains, *Pseudomonas aeruginosa* ST111 and *Enterobacter hormaechei* ST190 and ST93. The inset panel locates genomes from the Alfred Hospital in a phylogeny of global publicly available genomes of the same sequence type. In each inset, Alfred Hospital genomes form distinct lineages, which are then shown in the outer panel. Genomes for each of these three strains were clonally related, with median pairwise single nucleotide variant distances <20.

**Abbreviations:** ST – Sequence type.
We noted both flanking clusters and integron single nucleotide variant (SNV) profiles shared across multiple plasmid types (indicated by dotted plasmid lines). In particular, this was the case for IncC, IncFIA/IncFIB/IncP and IncFIB plasmids.

Abbreviations: No. – number; SNV – single nucleotide variant.