

# The rise and global spread of IMP carbapenemases (1996-2023): a genomic epidemiology study

## Authors

Ben Vezina<sup>1+</sup>, Bhargava Reddy Morampalli<sup>1+</sup>, Hoai-An Nguyen<sup>1</sup>, Angela Gomez-Simmonds<sup>2</sup>, Anton Y. Peleg<sup>1,3,4</sup>, Nenad Macesic<sup>\*1,3,5</sup>

<sup>1</sup> Department of Infectious Diseases, The Alfred Hospital and School of Translational Medicine, Monash University, Melbourne, Australia

<sup>2</sup> Division of Infectious Diseases, Department of Internal Medicine, UC Davis Health, Sacramento, California, USA

<sup>3</sup> Centre to Impact AMR, Monash University, Clayton, Australia

<sup>4</sup> Infection Program, Monash Biomedicine Discovery Institute, Department of Microbiology, Monash University, Clayton, Australia

<sup>5</sup> Infection Prevention & Healthcare Epidemiology, Alfred Health, Melbourne, Australia.

+ These authors contributed equally

\* Corresponding Author Dr. Nenad Macesic, Department of Infectious Diseases, The Alfred Hospital and School of Translational Medicine, Monash University, Level 1 Alfred Lane House, Alfred Hospital, 55 Commercial Rd, Melbourne, VIC 3004, Australia. Email: [nenad.macesic1@monash.edu](mailto:nenad.macesic1@monash.edu)

# Abstract

**Background:** IMP carbapenemases confer extensive drug resistance and are increasingly noted worldwide. Despite this, little is known regarding the global epidemiology of IMP carbapenemases.

**Methods:** We comprehensively identified *bla*<sub>IMP</sub> genes in all publicly available bacterial genomes, then systematically analysed the distribution of variants across species, lineages, plasmids and mobile elements, examining patterns over time, across geographic regions and by source. Structural analysis of IMP variants was performed.

**Findings:** 4,556 *bla*<sub>IMP</sub>-containing genomes were identified from 1996-2023, including 52 *bla*<sub>IMP</sub> variants across 93 bacterial species. Key variants (*bla*<sub>IMP-1</sub>, *bla*<sub>IMP-4</sub>, *bla*<sub>IMP-7</sub>, *bla*<sub>IMP-8</sub> and *bla*<sub>IMP-13</sub>) achieved global endemicity, while *bla*<sub>IMP-26</sub> and *bla*<sub>IMP-27</sub> were regionally endemic in Southeast Asia and North America, respectively. *bla*<sub>IMP</sub> dissemination was driven by horizontal gene transfer, facilitating inter-species spread. Proliferation of multidrug-resistant *Enterobacter hormaechei*, *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* lineages led to local outbreaks. Dereplication removed 3,175/4,556 (69.9%) genomes, indicating that most *bla*<sub>IMP</sub>-containing genomes were highly related. *bla*<sub>IMP</sub> variants were associated with mobile genetic element combinations including class 1 integrons and insertion sequences (99.7%), aiding mobilisation into ≥52 plasmid clusters, predominantly IncHI2A, IncN, IncL/M and IncC. Genomes of environmental and animal origin accounted for 10.0% and 1.1% of the dataset, respectively. Evidence of cross-source transmission was limited, with most spillover occurring between genomes of human and environmental origin. Structural analysis revealed a conserved carbapenemase structure (mean IDDT 0.977), with convergent missense mutations at seven catalytically relevant sites.

**Interpretation:** Global analysis enabled us to historically reconstruct the emergence and variant-specific epidemiologies of *bla*<sub>IMP</sub> carbapenemase genes. Intersecting mobile elements enabled *bla*<sub>IMP</sub> genes to spread across multiple plasmids and bacterial genera, facilitating global and multi-source spread within a One Health framework. Additionally, convergent evolutionary patterns indicate that IMP variants may continue evolving, potentially evading novel beta-lactam antimicrobial agents.

**Funding:** NHMRC EL1 (APP1176324) to N.M.; NHMRC PF (APP117940) to A.Y.P.; NIH/NIAID R01AI175414 to A.G-S.

## Research in context panel

### ***Evidence before this study***

Despite being a major cause of carbapenem resistance in Gram negative infections, little is known about the global epidemiology of IMP carbapenemases. IMP carbapenemases are metallo-beta-lactamases that were first identified in 1991 and have evolved into 96 different IMP variants. On May 21 2025, we searched all published reports available in PubMed using the terms “IMP' and 'carbapenemase' genomics NOT (Review[Publication Type]) NOT (Case Reports[Publication Type]) NOT PCR” with no language restrictions and no publication date restrictions. We identified 223 articles, 62 and 121 of which reported single species or a single study centre/country, respectively. Only 6 articles employed genomics to examine multi-species and multi-geographical isolates, though this was in the context of carbapenem resistance more broadly rather than IMP carbapenemases specifically. The most relevant study included 38 globally distributed genomes across four species and tracked seven *bla*<sub>IMP</sub> variants across mobile genetic elements.

### ***Added value of this study***

To our knowledge, this global characterisation provides the most comprehensive account of *bla*<sub>IMP</sub> carbapenemase gene epidemiology. To analyse the global distribution and diversity of *bla*<sub>IMP</sub> genes, we compiled all available public genome data resulting in a dataset of 4,646 genomes. This has allowed us to identify local, regional and international spread of *bla*<sub>IMP</sub> variants and determine the contributions of clonal expansion, plasmid proliferation and co-localised mobile genetic elements. We demonstrated that key *bla*<sub>IMP</sub> variants display global (IMP-1, IMP-4, IMP-7, IMP-8 and IMP-13) and regional (IMP-26 within Southeast Asia and IMP-27 within North America) endemicity and that these patterns have been previously unacknowledged, reframing the previous understanding that IMP carbapenemases were largely confined to the Asia-Pacific region. Our observation of convergent evolutionary patterns raise concern that IMP variants may continue to evolve, potentially evading new  $\beta$ -lactam antimicrobials. This analysis has revealed the under-recognised contribution IMP carbapenemases make to global carbapenem resistance.

## 77 ***Implications of all the available evidence***

78 These findings provide the first comprehensive atlas of *bla*<sub>IMP</sub> carbapenemase gene dissemination  
 79 and underscore the silent global spread of IMP carbapenemases. We note the critical need for  
 80 enhanced surveillance systems, particularly in low- and middle-income countries, that can detect  
 81 complex plasmid-mediated and mobile genetic element-associated spread, as we noted with *bla*<sub>IMP</sub>  
 82 carbapenemase genes. Moreover, our analyses show that systematic sampling across human,  
 83 animal, and environmental reservoirs is crucial to address the One Health dimensions of emerging  
 84 antimicrobial resistance threats. The study provides a framework for future interventions aimed at  
 85 tracking and stopping the spread of IMP carbapenemases and calls for co-ordinated, real-time public  
 86 health responses to this growing challenge.

87



# Introduction

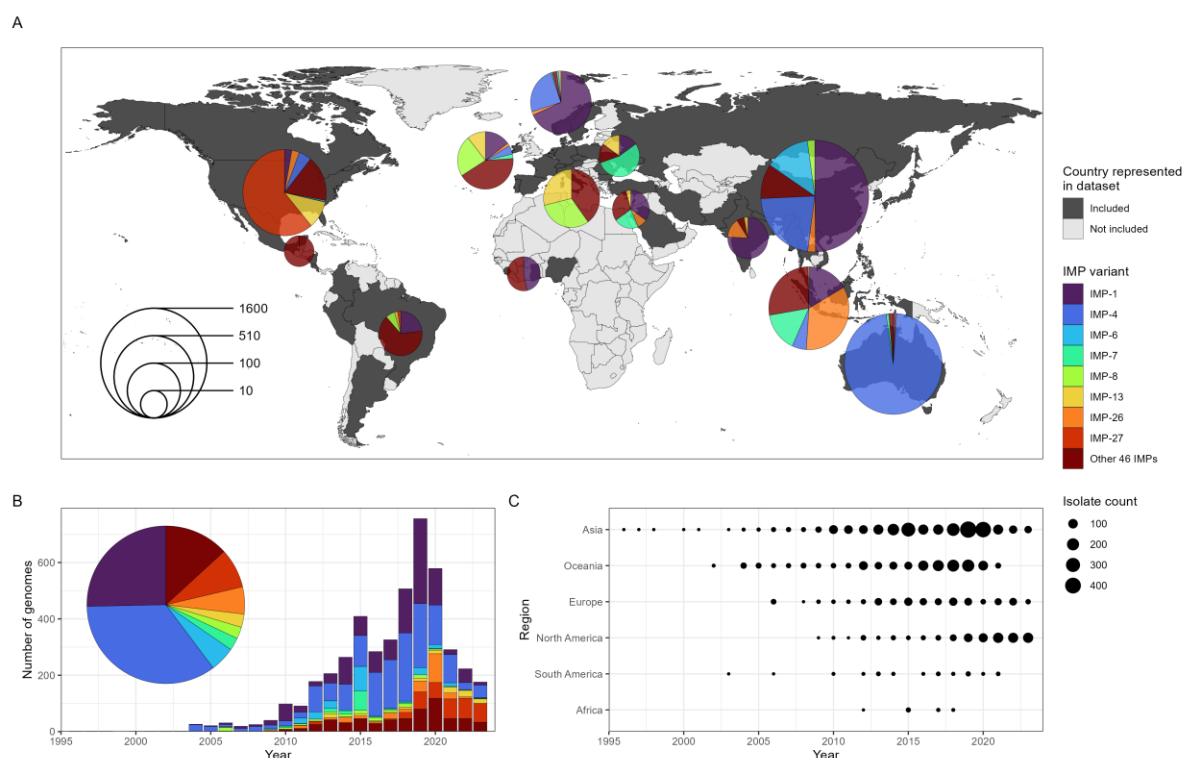
Carbapenemase-producing organisms (CPOs) are a significant threat to global health and have been deemed critical priority pathogens by the World Health Organization (1). Five key carbapenemase classes (KPC, NDM, OXA, VIM and IMP) cause the majority of global infections (2,3). Despite being included amongst these, little is known about the global epidemiology of IMP carbapenemases. IMP carbapenemases are metallo-beta-lactamases (MBLs) that were first identified in 1991 in *Pseudomonas aeruginosa* in Japan (4). To date, 96 different IMP carbapenemase variants have been identified and IMP carbapenemases are now endemic to Asia and Australia (5–10). In addition, outbreaks of IMP-carrying organisms are increasingly reported across several regions including Europe and the Americas (11–15). This is a highly concerning development given the paucity of treatment options for infections caused by these organisms, including resistance to novel agents with activity against other MBLs such as cefepime-taniborbactam (16).

Carbapenemases disseminate through various mechanisms, including transposon-mediated transfer between plasmids (*bla*<sub>NDM</sub>), stable association with successful clonal lineages (*bla*<sub>KPC</sub>), rapid expansion of a single epidemic plasmid across multiple bacterial lineages (*bla*<sub>OXA-48</sub>), and transient associations involving diverse plasmids and numerous lineages (17,18). Our prior work indicated that in Australia *bla*<sub>IMP-4</sub> spreads both clonally and through horizontal transfer via mobile genetic elements (19). However, current data on *bla*<sub>IMP</sub> dissemination remain limited: most prior studies have focused on a single IMP carbapenemase type and/or a specific geographical region (5,19–23).

We therefore aimed to comprehensively determine the genomic epidemiology of *bla*<sub>IMP</sub> carbapenemase genes. Specifically, we dissected the dynamics underlying *bla*<sub>IMP</sub> dissemination and evaluated the contributions of genomic factors, outbreak events, structural determinants and One Health-related influences. We analysed all publicly-available *bla*<sub>IMP</sub>-carrying genomes (n=4,556) spanning almost three decades (1996-2023), uncovering global expansion and regional endemicity of diverse *bla*<sub>IMP</sub> variants. Collectively, this work creates an atlas of *bla*<sub>IMP</sub> carbapenemase genes that highlights their transition from initial endemic foci in the Asia-Pacific region to a worldwide public health threat and emphasises the pressing need for integrated strategies to combat their further spread.

# Results

We identified 4,556 genomes (4,020 assembled from short- and 536 from long-read sequencing data) isolated globally from 1996-2023 carrying 52 distinct *bla*<sub>IMP</sub> variants across 26 bacterial genera (Fig. 1 and Table S1). This revealed a remarkable diversity of both *bla*<sub>IMP</sub> genes and their bacterial hosts, totalling 4,559 total *bla*<sub>IMP</sub> genes, with three genomes carrying two *bla*<sub>IMP</sub> variants each. *bla*<sub>IMP-4</sub> and *bla*<sub>IMP-1</sub> were the most frequent variants, found in 1,592/4,559 (34.9%) and 1,155/4,559 (25.3%) genomes, respectively (Table S2). The most prevalent species included 1,053 *Enterobacter hormaechei* (23.1%), 977 *Pseudomonas aeruginosa* (21.4%) and 681 *Klebsiella pneumoniae* (14.8%), together accounting for 59.4% of the dataset. Of the 4,556 *bla*<sub>IMP</sub>-carrying genomes, 728 (16%) carried mobile colistin resistance (*mcr*) genes and 345 (7.6%) carried other carbapenemase genes (Table S1).



**Fig. 1: Geographical and temporal spread of *bla*<sub>IMP</sub> carbapenemase genes.**

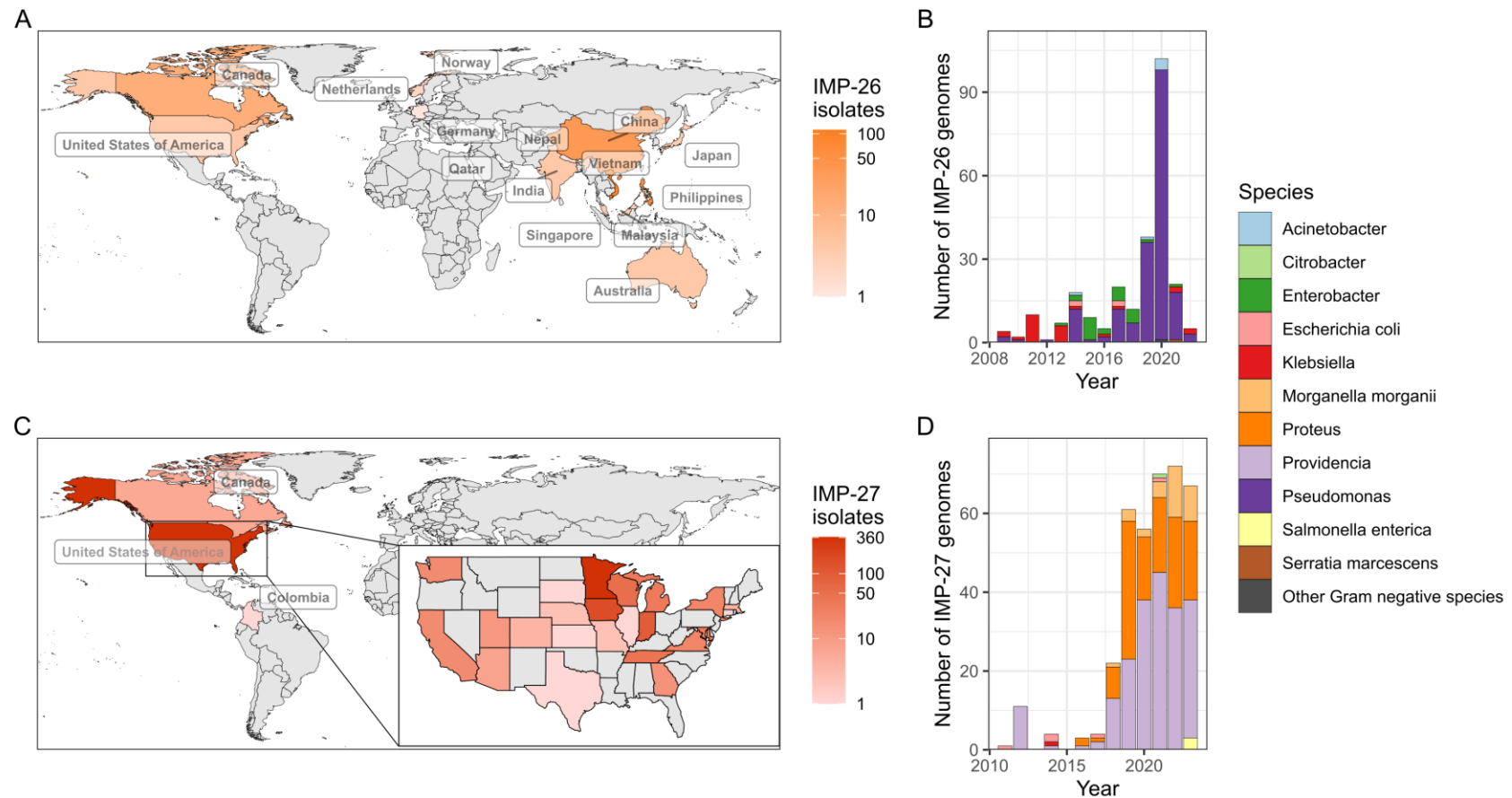
**A:** World map showing breakdown of *bla*<sub>IMP</sub>-carrying genomes and their global species distribution. Size of pie charts indicates number of isolates. *bla*<sub>IMP</sub> variants with <100 total observations are grouped into 'Other 46 IMPs'. **B:** Global prevalence of most common *bla*<sub>IMP</sub> variants over time. **C:** Dot plot showing prevalence of *bla*<sub>IMP</sub>-positive genomes within the United Nations geoscheme regions over time.

## Tracking the global spread of *bla*<sub>IMP</sub> carbapenemase genes

Our dataset enabled us to reconstruct the global expansion of *bla*<sub>IMP</sub> genes from origins in the Asia-Pacific region to an increasing number of variants detected across multiple regions as defined by the United Nations geoscheme (**Fig. 1A, C**).

From 1996-2011, *bla*<sub>IMP-1</sub>, *bla*<sub>IMP-4</sub> and *bla*<sub>IMP-6</sub> emerged in Asia and Australia, together accounting for 79.44% genomes studied. *bla*<sub>IMP-1</sub> was identified across Eastern and South Eastern Asia (939/1,155 *bla*<sub>IMP-1</sub> genomes) during the entire study period, most predominantly Japan, Singapore and China. After 2014, regional *bla*<sub>IMP-1</sub> outbreaks were increasingly noted outside of Asia, including in Europe, Western Africa and North America. *bla*<sub>IMP-4</sub> was initially noted in China in 1998 and continued to be isolated there through the study period (379/1,592 *bla*<sub>IMP-4</sub> genomes). However, from 2002 it was well established in Australia (1,105/1,592 genomes), predominantly on the east coast. Similarly to *bla*<sub>IMP-1</sub>, sporadic *bla*<sub>IMP-4</sub> outbreaks were noted outside these regions between 2014-2023 in Europe and Northern America (62/1,592 *bla*<sub>IMP-4</sub> genomes). *bla*<sub>IMP-6</sub> was identified in Japan from 2000 and remained focused there until end of the study period (224/235 *bla*<sub>IMP-6</sub> genomes), with sporadic isolation in South Korea, UK and USA between 2017-2023 (n=11 genomes).

From 2009, *bla*<sub>IMP-26</sub> and *bla*<sub>IMP-27</sub> emerged as regional *bla*<sub>IMP</sub> genes in South Eastern Asia and North America, respectively (**Fig. 2**). *bla*<sub>IMP-26</sub> was first noted in 2009 and became established in *P. aeruginosa* in Vietnam and Philippines. *bla*<sub>IMP-26</sub> was highly associated with *P. aeruginosa* ST235 (151/254 *bla*<sub>IMP-26</sub> genomes). This variant-lineage combination led to subsequent spread internationally across six geographic subregions from South-Eastern Asia to Southern Asia, Australia, Northern Europe, Western Europe and Northern America from 2009-2022, possibly due to travel-associated importation (**Fig. 2A, Table S3**). *bla*<sub>IMP-27</sub> emerged from 2011 as the dominant carbapenemase in North America, specifically the US (genomes from 21 states) (**Fig. 2B**). In addition to geographical location, this *bla*<sub>IMP</sub> variant displayed a distinct epidemiology characterised by associations with agricultural animals from 2011-2014 before emerging in human-origin genomes from 2016 onwards (**Table S3**). 77.6% (287/370 North American *bla*<sub>IMP-27</sub> genomes resulted from expansion of local lineages of *Providencia rettgeri* spp. 1, 2 and 3, *Providencia stuartii* spp. 1 and 2, *Providencia huaxiensis*, *Proteus mirabilis* and *Morganella morganii*.



**Fig. 2: *bla*<sub>IMP-26</sub> and *bla*<sub>IMP-27</sub> epidemiology**

**A:** Global distribution of *bla*<sub>IMP-26</sub>, showing regional endemicity in Asia and putative travel-association with other geographic regions. **B:** Species breakdown of *bla*<sub>IMP-26</sub> genomes over time. **C:** Global distribution of *bla*<sub>IMP-26</sub>, with panel showing regional endemicity concentrated within the US. **D:** Species breakdown of *bla*<sub>IMP-27</sub> genomes over time.

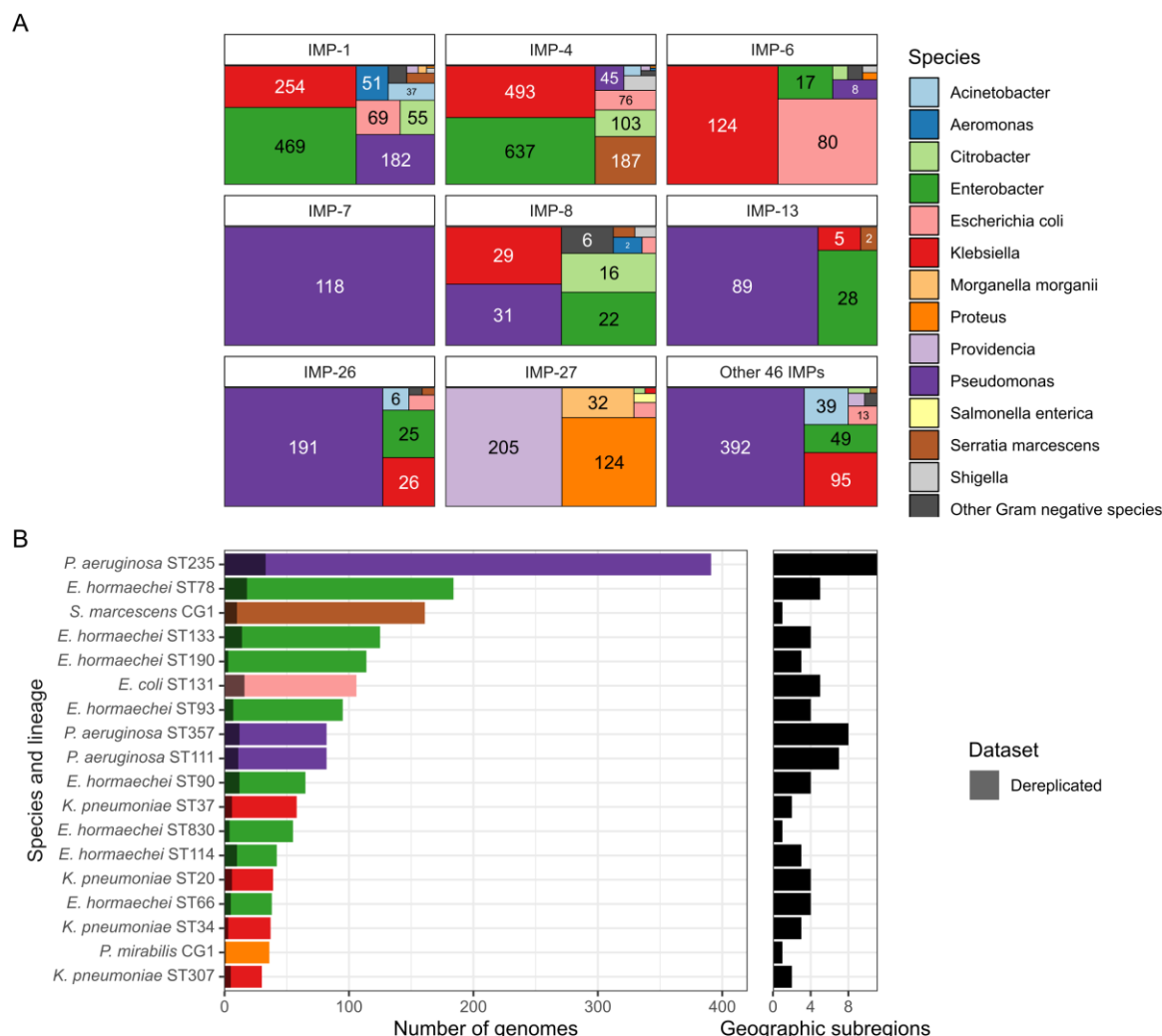
From 2018 to 2023, several *bla*<sub>IMP</sub> variants achieved global (*bla*<sub>IMP-1</sub>, *bla*<sub>IMP-4</sub>, *bla*<sub>IMP-7</sub>, *bla*<sub>IMP-8</sub>, *bla*<sub>IMP-13</sub>) and regional (*bla*<sub>IMP-6</sub>, *bla*<sub>IMP-26</sub>, *bla*<sub>IMP-27</sub>) endemicity, whereby there was evidence of ongoing spread outside of sporadic outbreaks (Table S4). During this period, the global dissemination of *bla*<sub>IMP</sub> carbapenemase genes was demonstrated with their detection in 42 countries spanning all regions, with 23 countries reporting  $\geq 2$  *bla*<sub>IMP</sub> variants and 8 countries reporting  $\geq 5$  *bla*<sub>IMP</sub> variants.

## ***bla*<sub>IMP</sub> carbapenemase genes found in diverse bacterial hosts with over-representation of multidrug-resistant lineages**

Having determined that *bla*<sub>IMP</sub> genes achieved global spread, we wanted to understand how genomic factors shaped this spread and hence adopted a multi-level approach focusing on bacterial hosts, plasmids and finally other mobile genetic elements. On a bacterial host level, there were 93 species carrying 4,556 total *bla*<sub>IMP</sub> genes but we noted that *bla*<sub>IMP</sub> variants were associated with specific species (Fig. 3A, Table S5). *bla*<sub>IMP-1</sub>, *bla*<sub>IMP-4</sub> and *bla*<sub>IMP-6</sub> were predominantly associated with *Enterobacterales* (specifically *E. hormaechei*, *K. pneumoniae*, *E. coli*), while *bla*<sub>IMP-7</sub>, *bla*<sub>IMP-13</sub> and *bla*<sub>IMP-26</sub> were predominantly noted in *P. aeruginosa*. *bla*<sub>IMP-27</sub> had a unique epidemiology dominated by *Providencia*, *Proteus* and *Morganella* spp.

*P. aeruginosa* sequence type ST235 was the most frequent single lineage isolated (391/4,556 genomes, 8.6%) and carried the greatest number of *bla*<sub>IMP</sub> variants (n=17) (Table S5). *P. aeruginosa* ST235 is a global multidrug-resistant (MDR) lineage, recognised for its ability to harbour a high diversity of acquired resistance genes (24). While there was a close association with *bla*<sub>IMP-26</sub> (151/391 *P. aeruginosa* ST235 genomes), ST235 also carried other key *bla*<sub>IMP</sub> variants including *bla*<sub>IMP-1</sub> (55/391 genomes), *bla*<sub>IMP-51</sub> (51/391 genomes), and *bla*<sub>IMP-7</sub> (46/391 genomes). We noted *bla*<sub>IMP</sub> presence in several other MDR lineages: *E. hormaechei* ST78 had a close association with *bla*<sub>IMP-1</sub> (162/4,556 [3.6%] genomes, 160 from Japan); *E. coli* ST131 with *bla*<sub>IMP-6</sub> (64/4,556 [1.4%] genomes, all from Japan); *K. pneumoniae* ST307 harboured *bla*<sub>IMP-38</sub> in a limited number of genomes (n=22, 21/22 from China) (Fig. 3B, Table S6).

We then quantified the impact of clonal bias and potential outbreaks by only including a single representative from each cluster of closely-related genomes (i.e. 'dereplication') (see Methods). We



**Fig. 3: Species and lineages associated with *bla*<sub>IMP</sub> carriage.**

**A:** Species breakdown across each *bla*<sub>IMP</sub> variant. **B:** Highly prevalent lineages within the dataset, showing the impact of dereplicating IMP-clusters. Dereplicated bars are shown as grey overlays over coloured genome counts, not stacked bars. Only lineages with  $\geq 30$  genomes are shown. Number of geographic subregions each lineage was detected in is shown as a companion plot. Full data found in **Table S5** and **Table S6**.

defined these 'IMP-clusters' as clusters of genomes which shared the same *bla*<sub>IMP</sub> variant, species, lineage (ST or clonal group [CG], if no MLST schema available) and were within a species-specific threshold of 5 single nucleotide variants (SNV) per Mb (see **Methods**). This resulted in 1,381 IMP-clusters, with 3,175/4,556 (69.9%) closely-related genomes removed by dereplication (**Fig. 3B**, **Table S6**). Indeed, the majority of *bla*<sub>IMP</sub> genomes were clonally linked to at least one other genome: 700/1,381 IMP-clusters contained  $>1$  genome (**Fig. 3B**, **Fig. S1**), potentially reflecting sequencing conducted in outbreak settings. This included multiple *P. aeruginosa* ST235 IMP-clusters (*bla*<sub>IMP-26</sub> - 127 genomes,

*bla*<sub>IMP-51</sub> - 51 genomes, *bla*<sub>IMP-1</sub> - 50 genomes, *bla*<sub>IMP-31</sub> - 39 genomes) (**Fig. S2A, Table S6**). Some IMP-clusters also accounted for high proportions of observations for that gene: *K. pneumoniae* ST37 and *E. coli* ST131 *bla*<sub>IMP-6</sub> IMP-clusters accounted for 14.5% and 13.2% of all *bla*<sub>IMP-6</sub> observations, while *E. hormaechei* ST190 and ST78 IMP-clusters accounted for 6.91% and 6.67% of *bla*<sub>IMP-4</sub> and *bla*<sub>IMP-1</sub> observations. We then assessed IMP-clusters for possible spread between countries and geographic subregions and noted that only 108/1,381 (7.8%) were found across >1 country. From the perspective of regional spread, 5.7% (78/1,381) IMP-clusters were found across ≥2 geographic regions, including 24 independent *P. aeruginosa* IMP-clusters (**Table S6**).

## Diverse plasmids facilitate global dissemination of *bla*<sub>IMP</sub> carbapenemase genes

To determine the genetic context of *bla*<sub>IMP</sub> genes, we first analysed all available long-read genomes in the dataset (n=539) (see **Methods**). *bla*<sub>IMP</sub> genes were located on plasmids in 436/539 genomes (80.9%). These were divided into 52 plasmid ‘clusters’ as defined by MOB-suite plasmid replicon clusters (25), which carried a total of 20 *bla*<sub>IMP</sub> variants (**Table S7**). IncHI2A, IncC and IncN plasmids collectively accounted for 56.0% (244/436) plasmids and each carried four, two and five different *bla*<sub>IMP</sub> variants. These plasmids were found across Asia and Australia with sporadic cases in Europe (**Table 1 and Table S7**). Other key *bla*<sub>IMP</sub> plasmids were IncU, IncFIB, IncFIB/IncFII-type 2, IncL/M, IncP and an untyped *A. baumannii* plasmid. These carried 2-4 *bla*<sub>IMP</sub> variants and were found in ≥2 geographic regions.

There were 181 plasmid cluster-*bla*<sub>IMP</sub> variant combinations and we evaluated their geographic distributions. We noted diversity of IncN, IncHI2A and IncP plasmids, with multiple distinct plasmid clusters (n=8, n=5 and n=4, respectively). These were predominantly associated with one *bla*<sub>IMP</sub> variant and one region (**Table S7**). Only 54/181 *bla*<sub>IMP</sub> variant-plasmid cluster combinations were found across different countries, while 13/181 were found in ≥3 geographic regions, driven by successful expansion of key IMP-clusters carrying broad host-range plasmids (IncHI2A, IncL/M, IncC, IncN). This indicated that while subregional spread of closely-related plasmids may have occurred, spread between regions was not detected in most cases. The regional *bla*<sub>IMP-26</sub> gene was found in IncHI2A, IncU and five untypeable plasmids (MOB-suite clusters AA739\_AJ059, AC212\_AL309, AC213\_AL312.



AC935\_AM305 and AD068\_AM495), respectively, however long read data was limited (13 genomes). Plasmids also displayed clear, species-specific host ranges. Most notably, IncHI2A plasmids were overwhelmingly associated with *E. hormaechei* (n=101), while IncC plasmids were associated with *S. marcescens* (n=30) and *K. pneumoniae* (n=18).

To gain insights from genomes with short-read data only, we clustered short-read contigs to long-read plasmid clusters (see **Methods**), with the caveat of reduced predictive confidence. In this analysis, *bla*<sub>IMP</sub> variants were found on plasmids in 2,909/4,559 (63.8%) genomes, the chromosome in 97 genomes, with 1,553 remaining unclassified (**Table S7**). Representation of most plasmids was similar between following inclusion of short-read assemblies (**Fig. S3**). Notable exceptions included IncC plasmids decreasing from 16.3% to 8.3%, while IncN and IncL/M plasmids increased (10.8% to 19.7% and 8.3% to 14%, respectively). IncN increases resulted from short-read *K. pneumoniae* species complex genomes largely absent in the long-read/hybrid dataset, while the IncL/M increase was driven by additional *Enterobacter* spp., *K. pneumoniae* species complex and *E. coli* short-read genomes. Of the 181 *bla*<sub>IMP</sub> variant-plasmid cluster combinations, 55 were found across multiple countries. A key example were *bla*<sub>IMP-4</sub> IncL/M plasmids that were detected in 249 genomes across Australia, US, UK and Philippines in 19 bacterial species (predominantly *Enterobacterales*). The remaining 126/181 *bla*<sub>IMP</sub> variant-plasmid cluster combinations were specific to a single country.

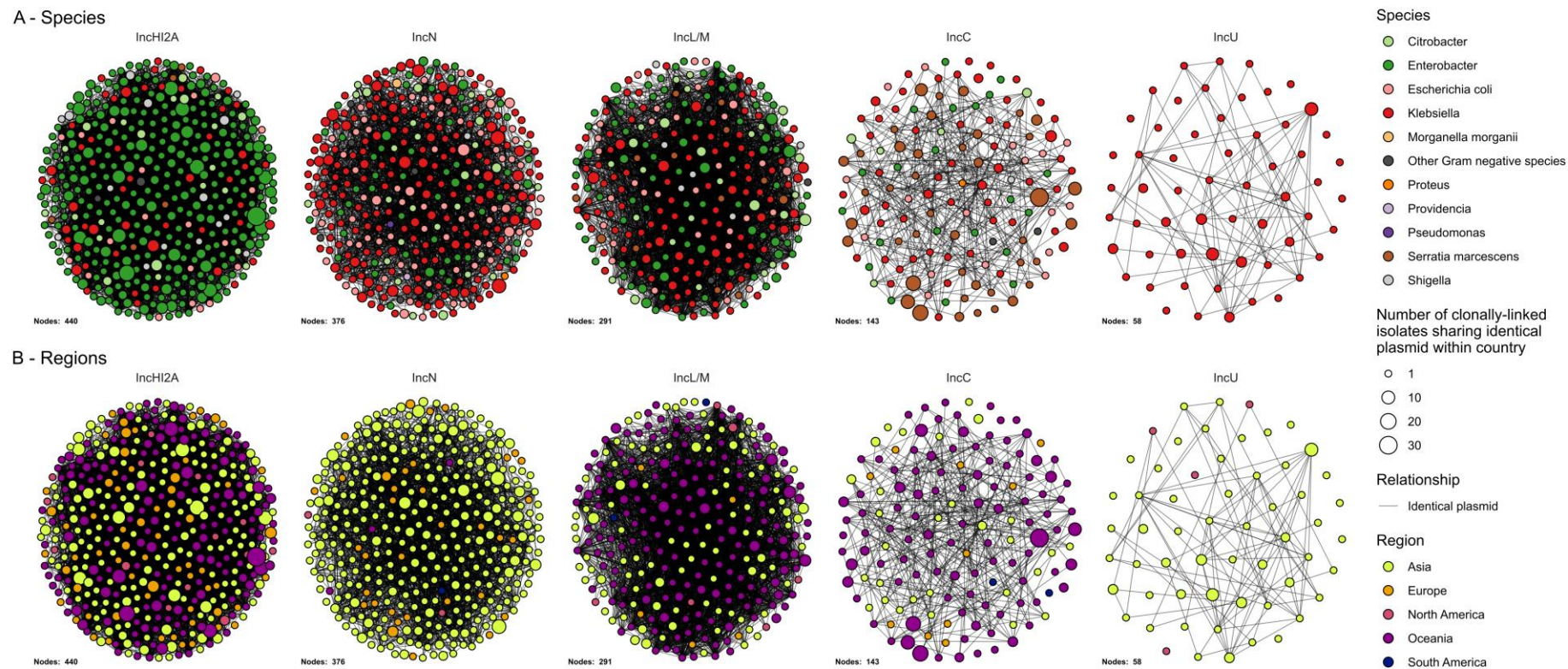
We then assessed relationships of *bla*<sub>IMP</sub> plasmids with bacterial host lineages. In our prior work, we noted successful lineage-plasmid pairings that we termed 'propagators' (19). There were 35 lineages possibly acting as propagators with ≥10 genomes including *S. marcescens* CG1 carrying *bla*<sub>IMP-4</sub> IncC plasmids (145 genomes), and *bla*<sub>IMP-1</sub> IncHI2A plasmids associated with *E. hormaechei* ST78 and *E. hormaechei* ST133 (n=116 and n=88, respectively) (**Table S8**). Additionally, we identified 64 'connector' lineages capable of harbouring ≥3 *bla*<sub>IMP</sub> plasmid clusters (**Table S8**), which could serve as an opportunity for transfer of *bla*<sub>IMP</sub>-containing integrons (19). *E. hormaechei* ST78 and *E. coli* ST131 were the most prominent, carrying 7 plasmid clusters each (**Table S8**).



271 **Table 1:** Summary of key plasmid clusters (n≥100) and associated features. Full data can be found in **Table S1** and **Table S7**.

MOB-suite cluster	Inc replicon	Year range	<i>bla</i> <sub>IMP</sub> variant	Countries	Species
AA739_AJ055	IncHI2A (n=396)	2004-2022	IMP-4, IMP-26, IMP-1, IMP-22, IMP-13	Australia, Philippines, Japan, Ireland, UK, Spain, US	<i>K. michiganensis</i> , <i>K. oxytoca</i> , <i>E. hormaechei</i> , <i>K. pneumoniae</i> , <i>E. cloacae</i> , <i>C. youngae</i> , <i>E. asburiae</i> , <i>E. kobei</i> , <i>R. planticola</i> , <i>E. roggenkampii</i> , <i>E. coli</i> , <i>C. amalonaticus</i> , <i>K. aerogenes</i> , <i>C. freundii</i> , <i>E. bugandensis</i> , <i>K. variicola</i> , <i>S. boydii</i> , <i>S. sonnei</i> , <i>C. koseri</i> , <i>C. farmeri</i> , <i>E. chengduensis</i> , <i>C. portucalensis</i>
AA552_AI753	IncN (n=302)	2004-2023	IMP-6, IMP-1, IMP-66, IMP-26, IMP-22, IMP-4, IMP-74, IMP-11	Japan, Philippines, UK, Singapore, Spain, Portugal, Germany, Peru, China	<i>K. pneumoniae</i> , <i>E. coli</i> , <i>C. freundii</i> , <i>E. asburiae</i> , <i>E. hormaechei</i> , <i>E. kobei</i> , <i>E. ludwigii</i> , <i>K. michiganensis</i> , <i>R. ornithinolytica</i> , <i>K. variicola</i> , <i>S. sonnei</i> , <i>E. chengduensis</i> , <i>P. stuartii</i> Spp 1, <i>L. adecarboxylata</i> , <i>P. mirabilis</i> , <i>K. quasipneumoniae</i> , <i>E. cloacae</i> , <i>K. grimontii</i> , <i>C. braakii</i>
AA739_AJ057	IncHI2A (n=276)	2007-2021	IMP-1, IMP-13	Japan, US, UK	<i>E. hormaechei</i> , <i>C. freundii</i> , <i>E. asburiae</i> , <i>E. kobei</i> , <i>K. michiganensis</i> , <i>E. cloacae</i> , <i>E. coli</i> , <i>E. chengduensis</i> , <i>K. pneumoniae</i> , <i>S. marcescens</i> , <i>E. roggenkampii</i> , <i>K. oxytoca</i>
AA860_AJ266	IncC (n=260)	2002-2022	IMP-4, IMP-1, IMP-60, IMP-8, IMP-15, IMP-23	Australia, Japan, Netherlands, UK, Spain, China, Brazil, Singapore	<i>S. marcescens</i> , <i>K. pneumoniae</i> , <i>E. chengduensis</i> , <i>E. hormaechei</i> , <i>K. oxytoca</i> , <i>C. freundii</i> , <i>E. asburiae</i> , <i>E. coli</i> , <i>V. alginolyticus</i> , <i>K. michiganensis</i> , <i>K. variicola</i> , <i>S. sonnei</i> , <i>E. cloacae</i> , <i>P. mirabilis</i> , <i>C. koseri</i> , <i>E. bugandensis</i> , <i>K. quasipneumoniae</i>
AA002_AH532	IncL/M (n=253)	2006-2023	IMP-4, IMP-34, IMP-1, IMP-22, IMP-59	Australia, Japan, Philippines, Portugal, US, UK	<i>K. pasteurii</i> , <i>K. quasipneumoniae</i> , <i>C. freundii</i> , <i>K. pneumoniae</i> , <i>E. hormaechei</i> , <i>E. coli</i> , <i>K. michiganensis</i> , <i>S. sonnei</i> , <i>S. marcescens</i> , <i>C. murlinae</i> , <i>K. aerogenes</i> , <i>E. asburiae</i> , <i>K. variicola</i> , <i>S. boydii</i> , <i>C. koseri</i> , <i>E. kobei</i> , <i>L. adecarboxylata</i> , <i>A. subterranea</i> , <i>C. farmeri</i>
AA552_AI757	IncN (n=153)	2008-2023	IMP-6, IMP-4, IMP-38, IMP-26	Japan, China, US	<i>K. pneumoniae</i> , <i>K. michiganensis</i> , <i>K. pasteurii</i> , <i>A. subterranea</i> , <i>C. freundii</i> , <i>E. asburiae</i> , <i>E. hormaechei</i> , <i>K. grimontii</i> , <i>E. coli</i> , <i>E. soli</i> , <i>R. ornithinolytica</i> , <i>K. quasipneumoniae</i> , <i>K. variicola</i> , <i>R. planticola</i> , <i>C. amalonaticus</i> , <i>C. portucalensis</i> , <i>H. chinensis</i> , <i>C. braakii</i> , <i>E. roggenkampii</i>
AA739_AJ059	IncHI2A (n=134)	2008-2023	IMP-6, IMP-1, IMP-4, IMP-22, IMP-8, IMP-13, IMP-26, IMP-19	UK, Portugal, Taiwan, US, China, Australia, Poland	<i>C. youngae</i> , <i>E. hormaechei</i> , <i>E. asburiae</i> , <i>A. hermannii</i> , <i>E. cloacae</i> , <i>E. coli</i> , <i>K. pneumoniae</i> , <i>E. bugandensis</i> , <i>K. oxytoca</i> , <i>C. freundii</i> , <i>E. kobei</i> , <i>K. aerogenes</i> , <i>S. marcescens</i> , <i>P. vulneris</i>
AA739_AJ058	IncHI2A (n=113)	2010-2022	IMP-4, IMP-6, IMP-1, IMP-13	Australia, Japan, Spain, US, Singapore, China	<i>E. hormaechei</i> , <i>K. oxytoca</i> , <i>K. pneumoniae</i> , <i>E. coli</i> , <i>E. asburiae</i> , <i>K. michiganensis</i> , <i>E. kobei</i> , <i>C. amalonaticus</i> , <i>K. variicola</i>

272



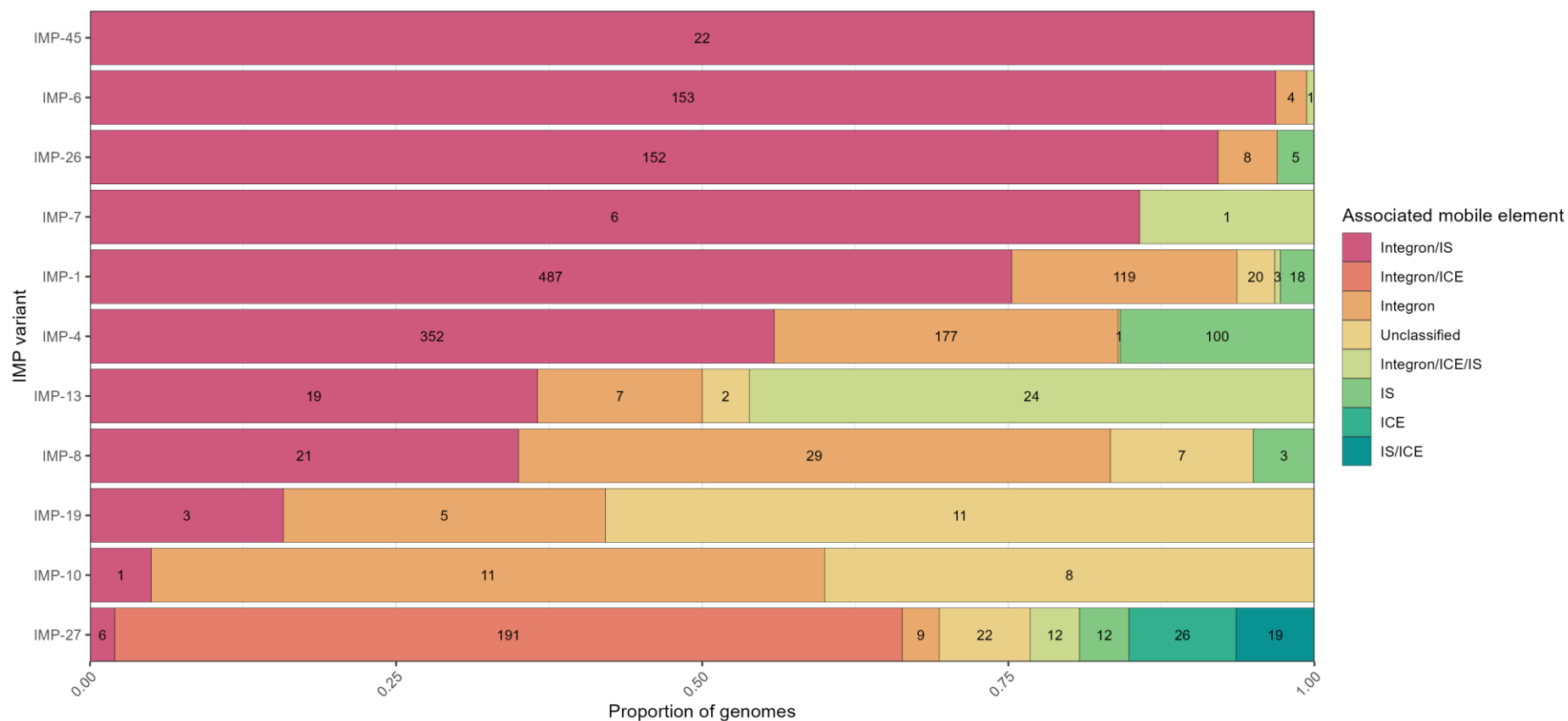
**Fig. 4: Key *bla*<sub>IMP</sub> plasmids by bacterial species and geographic region**

**A:** Network of related plasmids, separated by plasmid Inc types and coloured by host species. Edges show identical plasmids between IMP-clusters (nodes). Size of node represents number of genomes within an IMP-cluster and country also carrying the same plasmid. **B:** Identical networks but coloured by geographical region.

## Mobile elements associated with *bla*<sub>IMP</sub> carbapenemase genes

*bla*<sub>IMP</sub> genes are frequently associated with integrons (26), integrative and conjugative elements (ICEs) (27), insertion sequences (IS) (28) and transposons (29). We therefore sought to determine if these mobile elements could provide insight into the spread of *bla*<sub>IMP</sub> genes. We analysed the genetic context 10kb up- and downstream of each *bla*<sub>IMP</sub> gene. To maximise identification of mobile elements, only contigs ≥6kb in length were considered for this analysis which resulted in 2,314/4,559 eligible *bla*<sub>IMP</sub>-containing contigs (**Fig. 5**). Initially, we examined the total presence of mobile elements across the dataset. Of these, 2,053/2,314 (88.7%) contigs showed direct association with integron elements (850 intact and 1,203 with *attC* clusters only) and almost all were class 1 integrons (795/850). When taking other mobile elements into account, integrons were found alone in 478/2,314 (20.6%) contigs but were co-located with other mobile elements in the majority, including IS (1,332/2,314 [57.6%]) and ICE (192/2,314 [8.3%]) (**Fig. 5**).

Co-location of multiple mobile elements with *bla*<sub>IMP</sub> genes may have contributed to their dissemination across populations due to increased mobility. Integron/IS-associated *bla*<sub>IMP</sub> genes were most frequent (across 56 species, 39 IMP variants and 41 plasmid clusters), dating back to the earliest *bla*<sub>IMP</sub> isolates in the late 1990s (**Fig. S4**). We saw no major differences between *bla*<sub>IMP</sub> variants and their putative associated mobile element/s. *bla*<sub>IMP-27</sub> was a notable exception: it had the most variable mobile element associations, characterised by presence of ICE alone or in combination with integrons and/or other IS (**Fig. 5**).

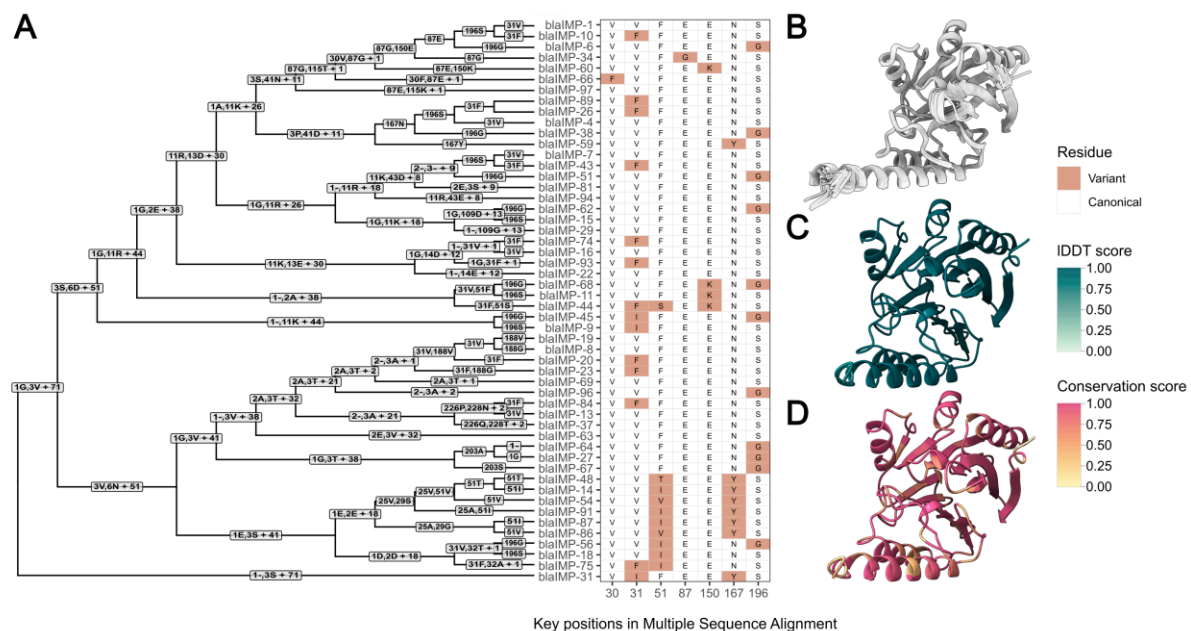


**Fig. 5: The majority of *bla*<sub>IMP</sub> variants are associated with integrons and IS elements.**

Column graph showing the proportion of each *bla*<sub>IMP</sub> variant and its association with mobile elements including integrative conjugative elements (ICE), integrons, insertion sequences (IS), and unclassified. Numbers within bars show *n* of each group. Full data found in **Table S1**.

## IMP variants display convergent evolutionary trends

Beyond analysing genomic contributors to IMP dissemination, we analysed sequence and structural variation of IMP variants in our study (**Fig. 6A and B**). Critically, this variation is closely linked with  $\beta$ -lactam hydrolytic specificity and alters minimum inhibitory concentrations of carbapenems (30–33). This led us to hypothesise that it may influence *bla*<sub>IMP</sub> dissemination. We noted IMP sequences were highly structurally conserved, with a mean local distance difference test (IDDT) score of 0.977 (**Fig. 6C**). Despite this, we saw reduced amino acid conservation scores at key positions throughout the structure (**Fig. 6D**), indicative of varying substrate specificity/activity. We examined mutations in key residue positions (30, 31, 51, 87, 150, 167, 196) previously shown to impact carbapenem hydrolytic specificity and noted a consistent pattern of convergent evolutionary mutations throughout the protein phylogeny at these key residues. Many of these convergent mutations appear to have been acquired independently (**Fig. 6A**), such as 31F found in multiple IMP sequences.



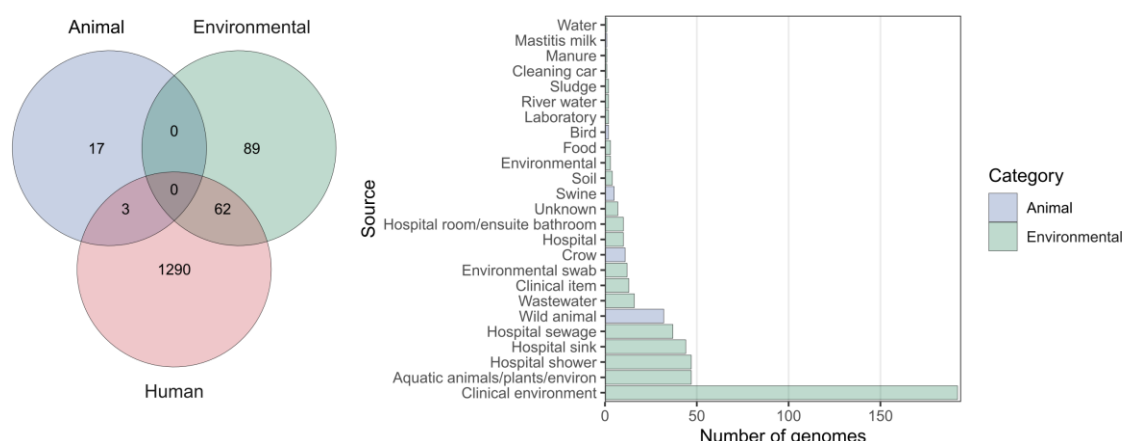
**Fig. 6: IMP variants have key residue changes which alter catalytic activity**

**A:** Multiple sequence alignment guide tree with associated amino acid changes. Key residues which have demonstrated impact on antimicrobial resistance profiles (30,33–35) are shown as a heatmap. **B:** Predicted AlphaFold2 structures of each 52 IMP variants, structurally aligned. **C:** IDDT scores for each residue after structural alignment with FoldMason, indicating level of structural conservation. **D:** Amino acid conservation score, as calculated using residue similarity and the bio3d sub matrix. The Multiple Sequence Alignment is found at Figshare: 10.6084/m9.figshare.28440992.



## Non-human reservoirs of *bla*<sub>IMP</sub> carbapenemase genes

We then adopted a One Health approach by analysing sources of *bla*<sub>IMP</sub> genomes to determine the potential contribution of environmental or animal reservoirs to *bla*<sub>IMP</sub> dissemination. Genomes of human origin accounted for the majority (4,051/4,556, 88.9%), while genomes of environmental and animal origin accounted for 10.0% (454/4,556) and 1.1% (51/4,556), respectively (Fig. 7A). Nine samples were unable to be classified. Most environmental genomes (353/454, 77.8%) came from healthcare settings, with at least 128/353 (36.3%) coming from hospital aquatic environments (Fig. 7B, Table S9). Animal data were limited but *bla*<sub>IMP-4</sub> and *bla*<sub>IMP-38</sub> were detected in Australian seagulls (n=31 and n=1, respectively) (36), and *bla*<sub>IMP-27</sub> in US and Canadian genomes of pig origin (n=4 and n=1, respectively). 65/1,381 (4.7%) IMP-clusters were found across multiple source categories (Fig. 7A), most notably *bla*<sub>IMP-26</sub> *P. aeruginosa* ST235 and *bla*<sub>IMP-4</sub> *S. marcescens*. Only three IMP-clusters had genomes of both human and animal origin (all birds), including *bla*<sub>IMP-4</sub> *E. coli* ST58, *bla*<sub>IMP-27</sub> *P. rettgeri* spp. 2 CG5 and *bla*<sub>IMP-45</sub> *P. aeruginosa* ST313.



**Fig. 7: Human, animal and environmental origins of *bla*<sub>IMP</sub> genomes.**

**A:** Venn diagram showing the intersections of IMP-clusters (as previously defined) between genomes of human, environmental and animal origins, indicating either clonal spread between categories or discrete isolation. **B:** Number of genomes isolated from specific non-human sources.

## Discussion

We have systematically determined the global distribution, diversity and epidemiology of 52 *bla*<sub>IMP</sub> variants across 4,556 publicly-available genomes. This enabled us to reconstruct a complete picture of the emergence of IMP carbapenemases as a global and multi-faceted problem across all geographic

regions. Previous studies have either focused on a single *bla*<sub>IMP</sub> variant or geographical site (5,19,22,37–39), leaving gaps in our understanding of the genomic epidemiology of these genes. Our analysis has revealed the under-recognised diversity and global distribution of *bla*<sub>IMP</sub> carbapenemase genes. We have shown that *bla*<sub>IMP-1</sub> and *bla*<sub>IMP-4</sub> were the most common variants due to their early emergence (40,41) and wide dissemination, achieving global endemicity along with *bla*<sub>IMP-7</sub>, *bla*<sub>IMP-8</sub>, *bla*<sub>IMP-13</sub> (**Fig. 1**). In contrast, several *bla*<sub>IMP</sub> variants displayed distinct regional patterns including *bla*<sub>IMP-26</sub> (**Fig. 2A**) in Southeast Asia and *bla*<sub>IMP-27</sub> in North America (**Fig. 2C**), achieving regional endemicity. These patterns have been previously unrecognised and challenges earlier perceptions that IMP carbapenemases are largely confined to the Asia-Pacific region (42).

Horizontal gene transfer, more than clonal spread, enabled global dissemination of *bla*<sub>IMP</sub> genes. While previous work has demonstrated associations of *bla*<sub>IMP</sub> and *bla*<sub>GES-5</sub> genes with integrons (19,43,44) and co-location with IS (45,46), we demonstrate these intersections systematically at scale. We found that *bla*<sub>IMP</sub> genes were almost invariably embedded in class 1 integrons, frequently flanked by IS. This genetic context promoted *bla*<sub>IMP</sub> mobility, enabling entry into a vast array of at least 52 plasmid clusters. These included broad host-range IncHI2A, IncC, and IncN plasmids, all found to carry multiple *bla*<sub>IMP</sub> variants. The coupling of integrons and IS with these broad-host range plasmids likely drove interspecies spread of *bla*<sub>IMP</sub> genes into diverse bacterial hosts, specifically *Enterobacteriales* (e.g. *Enterobacter* spp., *Klebsiella* spp. but also *Providencia* spp. and *Proteus* spp.). The regional *bla*<sub>IMP-27</sub> gene was not associated with any known plasmids using our approach (which relied on completed plasmid references), but has been shown to be associated with the conjugative pPM187 (IncX8 replicon type) and pPR1 (no known replicon type) plasmids, allowing experimentally-verified inter-species spread (47).

Despite horizontal gene transfer being a key driver of *bla*<sub>IMP</sub> spread, we also noted that proliferation of certain bacterial host lineages was an important contributor to *bla*<sub>IMP</sub> dissemination. *bla*<sub>IMP-1</sub> *P. aeruginosa* ST235 was a successful global lineage but other lineages were more geographically limited. Instead, we noted numerous local outbreaks and limited inter-country spread of successful plasmid-lineage combinations that we previously termed ‘propagators’ (19). We identified at least 35 propagators that sustained and amplified local *bla*<sub>IMP</sub> outbreaks (e.g. *bla*<sub>IMP-4</sub> *S. marcescens*-IncC, *bla*<sub>IMP-1</sub> *E.*

*hormaechei*-IncHI2A). This is in contrast to other carbapenemases such as KPC and OXA-48, which are associated with globally disseminated clones such as *K. pneumoniae* ST258 and ST11, respectively (48). In addition, we propose the idea of ‘connectors’: lineages that can accept multiple plasmids and serve as bridges for *bla*<sub>IMP</sub> gene transfer through mobile genetic elements, without themselves causing local outbreaks. We identified 64 such lineages such as *E. hormaechei* ST78 and *E. coli* ST131. The interplay of propagators and connectors helps explain how *bla*<sub>IMP</sub> genes can disseminate and also repeatedly establish in new hosts. The *bla*<sub>IMP</sub> threat is therefore multidimensional: it spreads by many local expansions rather than a single dominant lineage crossing borders, thus posing difficulties for both detection and control.

Beyond the various levels of *bla*<sub>IMP</sub> gene transfer, we also analysed the structures of IMP enzymes to examine whether adaptive changes in the enzymes themselves may be contributing to spread. Despite the diversity of 52 known IMP variants, we detected convergent evolutionary patterns, with repeated missense mutations at specific sites. These changes likely offer functional advantages by altering carbapenem hydrolysis activity and thus minimum inhibitory concentrations (30,33–35). For example, IMP-26, which we identified as key regional endemic clone in Asia, is more effective at hydrolysing meropenem and doripenem than older variants and thus displays a broader and more effective carbapenemase phenotype (49). While we cannot determine which selective pressures led to the rise of each *bla*<sub>IMP</sub> variant, we speculate that this could be driven by antimicrobial usage (30) and species-specific susceptibility (50). The fact that similar mutations have arisen independently in different *bla*<sub>IMP</sub> variants implies ongoing adaptive evolution. Similar convergent mutation patterns have been shown in the *K. pneumoniae* extended-spectrum  $\beta$ -lactamase *bla*<sub>SHV</sub> (51). These evolutionary patterns raise concern that IMP enzymes may continue to evolve, not only with increasing carbapenemase activity but also potentially evading even the newest  $\beta$ -lactamase inhibitors.

Finally, we adopted a One Health lens to help understand and hopefully control *bla*<sub>IMP</sub> spread. Although most genomes came from human clinical isolates, *bla*<sub>IMP</sub> genes were also found in hospital environments (e.g., surfaces, wastewater), as well as animal samples (e.g., in cats and birds). Our fine-grained IMP-cluster analysis allowed detection of movement between these One Health categories, predominantly between human clinical isolates and clinical environments, and to a lesser extent



between environmental/animal sources. This illustrates how healthcare-associated resistance genes can spill over into animals or the environment, creating secondary reservoirs. These reservoirs may, in turn, seed new infections back into humans, forming a cycle that blurs traditional epidemiological boundaries. The presence of *bla*<sub>IMP</sub> genes in *Serratia* spp., *Pseudomonas* spp. and other environmental bacteria also raises the possibility of environmental persistence. *bla*<sub>IMP</sub> genes therefore have access to multiple ecological niches (human, animal, and environmental) that constantly interact, further complicating control efforts. Furthermore, non-clinical isolation sources are likely under-sampled, thus underestimating the true extent of non-human *bla*<sub>IMP</sub> reservoirs and underscoring the need for a One Health approach for surveillance and control.

Our study had ambitious reach, aiming at analysing the global genomic epidemiology of *bla*<sub>IMP</sub> genes, but this was also the source of several limitations. Our analysis relied on publicly-available genomes that represented biases in whole genome sequencing activity and genomic surveillance. This led to overrepresentation of countries with genomic infrastructure, with corresponding under-representation of low- and middle-income countries that are likely disproportionately impacted by infections caused by IMP carbapenemases (52). On a technical level, we were reliant on the quality of genomic data submitted. Most data were derived from short-read sequencing, introducing uncertainty for determining the genetic contexts of *bla*<sub>IMP</sub> genes. We tried to compensate for this by conducting a mapping-based analysis but acknowledge the inherent limitations of this approach. Similarly, metadata were frequently poorly curated, thus preventing discovery of additional non-clinical spread and limiting geographic/temporal inferences. Moreover, the dataset includes 52 of 96 known IMP variants, leaving rare variants uncharacterised.

In summary, we have demonstrated that the emergence of IMP carbapenemases has largely ‘flown under the radar’, despite the establishment of *bla*<sub>IMP</sub> variants with global endemicity (IMP-1, IMP-4, IMP-7, IMP-8, IMP-13), as well as regional endemicity (IMP-26 and IMP-27). We have shown that this global spread was the result of a complex interplay of genomic drivers of dissemination (at the mobile genetic element, plasmid and bacterial host lineage levels), with horizontal gene transfer playing a more substantial role than global spread of specific lineages. We noted convergent evolution in IMP carbapenemase enzymes, suggesting adaptation under evolutionary pressures and posing an ongoing

challenge with the possibility of adapting to and conferring resistance to new antimicrobial treatments. Finally, we noted the presence of One Health reservoirs of IMP carbapenemases, although detailed analysis was limited by the small number of non-clinical genomes. These findings provide a detailed atlas of IMP carbapenemases and their global spread that also carry implication for other carbapenemases and emerging forms of antimicrobial resistance. In addition to casting light on the extent IMP carbapenemases have silently become a global problem, we have provided a roadmap for future interventions to disrupt their future spread. Our work highlights the need for more robust and sophisticated surveillance approaches that address gaps in low- and middle-income countries, incorporate methodology to detect plasmid and mobile genetic element transmission and conduct more systematic sampling of One Health reservoirs. We have carried out a systematic global analysis of IMP carbapenemases using three decades of data but the critical threat they pose now requires a proactive, real-time and co-ordinated public health response.

## Methods

### Genome acquisition, species identification and genotyping

NCBI Pathogen Isolate Browser (53) was searched using the term "AMR\_genotypes:blaIMP\*" to retrieve assemblies containing all IMP variants. This initial search identified a total of 4,063 genomes from 517 BioProjects. Of these, 3,177 assemblies were downloaded from Pathogen Isolate Browser and 886 reads were obtained from Sequence Read Archive (SRA) and assembled with Unicycler v.0.4.8 (54) with standard parameters.

To identify additional publicly available assemblies, protein accession numbers for all IMP variants were obtained and queried against NCBI GenBank databases using BLAST v2.15 (55). Assemblies retrieved from this search were compared against the Pathogen Isolate Browser dataset, and non-duplicates were retained, resulting in an additional 583 assemblies. Collectively, this approach yielded a total of 4,556 assemblies, which were used for all subsequent analyses. Metadata was extracted from the BioSample accessions. In cases where no 'collection date' was available and a BioProject did not contribute  $\geq 50$  isolates, the date of sequence upload was used as a proxy. Assembly quality was

checked using Quast v5.2.0 (56) and species identification was performed using Speciator (57). Genomes were annotated using Prokka v1.14.6 (58). We then performed resistance gene detection with AMRFinderPlus v3.12.8 (59). We determined *in silico* multi-locus sequence type (ST) using ‘mlst’ v.2.19.0 (60). All inconclusive ST calls were checked with SRST2 v0.2.0 (61).

We note that several key *bla*<sub>IMP</sub>-containing species such as *Providencia rettgeri*, *Providencia stuartii*, *Morganella morganii* and *Proteus mirabilis* do not have well-defined MLST schema. As such, we employed a two-tiered approach to classify lineages for this study. Initially, untyped species were assigned MASH clusters if they were within  $\leq 0.05$  of each other to accurately group independent species (62), then PopPUNK v4.2.0 (63) was used on each of these MASH clusters to assign lineages for within-species comparisons. The ‘create-db’ function was used with the following options: ‘--sketch-size 1000000 --min-k 15 --max-k 29 --qc-filter prune’. Then the ‘fit-model’ function was used with the following options: ‘bgmm --ranks 1,2,3,5 --graph-weights’. Various --K values were used to obtain the best-scoring model fitting. Additionally, the ‘poppunk\_visualise’ function was used, with the ‘--distances’ and ‘--previous-clustering’ to output a neighbour-joining core tree.

We identified that species identified as *Providencia rettgeri* and *Providencia stuartii* were considerably divergent and the taxonomy of these species requires detailed revision, with five and two independent species groupings outside the 0.05 standard threshold. These displayed MASH distances of up to 0.19 ‘within species’, indicating that although these ‘species’ are related, they are distinct species or potentially genera. However, as this is beyond the scope of our current work, we refer to these as *Providencia rettgeri* spp. 1, *Providencia rettgeri* spp. 2 etc.

For assemblies without recorded sequencing technology, a heuristic approach was adopted: assemblies with fewer than 20 contigs were classified as long-read or hybrid, while those with 20 or more contigs were marked as short-read. We assumed that long-read sequencing or assemblies with both short- and long-read data typically yields fewer contigs compared to short-read sequencing data alone. The United Nations geoscheme was used to classify countries into regions and subregions. This resulted in a dataset comprising 536 long and 4,022 short read assemblies.

## Assembly dereplication and IMP-cluster definition

To determine the impact of potential clonal outbreaks, a dereplicated dataset was employed. Initially, all genomes were split into their respective STs/CGs, then pairwise SNV distances were calculated using the 'fasta' command from SKA v1.0.0 (64), followed by the 'distance' command. Clonally-linked genomes were detected if two genomes had  $\leq 5$  pairwise SNVs per Mb. igraph v2.0.3 (65) was then used to perform graph-based clustering using pairwise SNV differences as edges, then membership data extracted for any isolates forming a distinct sub-clusters, allowing identification of putative spread clusters. 'IMP-clusters' were defined as genomes which shared the same *bla<sub>IMP</sub>* variant, species, ST/CG, and were within the SNV threshold as previously defined.

## Plasmid analysis

Initially, the mob\_recon function from MOB-suite v3.1.8 (25) was used to assign contigs to plasmids and identify replicon types and other information to long read assemblies only. These 433 plasmid-positive contigs were used to construct a custom *bla<sub>IMP</sub>*-positive plasmid database via the "mob\_typer -multi", then "mob\_cluster --mode build" commands. Short-read assemblies were then screened against this custom database with mob\_recon.

## Genetic context of mobile genetic elements

Flanker v0.1.5 (66) was used to determine genetic context of *bla<sub>IMP</sub>* genes with the following commands: "--flank both --window 10000 --gene blaIMP --include\_gene --cluster". For analysis of mobile elements, only sliced contigs with  $\geq 6$  kb were used. Integrations were identified using Integron\_finder v2.0.5 (67) with the "--local-max --func-annot --union-integrases" options. IS were screened using ISEScan v1.7.2.3 (68). ICEs were identified using the ICEberg 2.0 (69) database as a query against each genome with minimap2 v2.26 (70).

## Protein and structural analysis

SignalP v6.0 (71) was used to process the signal sequences from IMP protein sequences, then Clustal Omega v1.2.4 (72) was used to align these sequences. This was used as input to calculate amino acid conservation scores using the conserve() command from bio3d v2.4-4 (73), with the following parameters: 'method="similarity", sub.matrix="bio3d"'. Colabfold v1.5.5 (74) was used to generate

AlphaFold2 (75) structure predictions of mature IMP sequences, with the rank 1 structures kept. All structures were used as input for FoldMason v1.763a428 (76) to generate per-residue structural conservation IDDT scores. Scores were used to colour structures by residue using ChimeraX v1.8 (77) and custom scripts at [https://github.com/bananabenana/ChimeraX\\_scripts](https://github.com/bananabenana/ChimeraX_scripts) (78).

## Statistical analysis

Statistics were performed in R v4.4.1 (79) and RStudio v2024.09.0 (80). The following R packages were used: tidyverse v2.0.0 (81), colorspace v2.1-1 (82), viridis v0.6.5 (83), ggh4x v0.2.8 (84), ggstream v0.1.0 (85), maps v3.4.2 (86), scatterpie 0.2.4 (87), sf v 1.0-17 (88), rnaturalearth v1.0.1 (89), ggnewscale v0.5.0 (90), treemapify v2.5.6 (91), patchwork v1.3.0 (92), igraph v2.0.3 (65), qgraph v1.9.8 (93), ggraph v2.2.1 (94), ggforce v0.5.0 (95), ggalluvial v0.12.5 (96), ggtree v3.12.0 (97), treeio v1.28.0 (98) and applot v0.2.3 (99). All R code can be found at Figshare: 10.6084/m9.figshare.28440992.

## Competing interests

The author(s) declare no competing interests.

## Data availability

All data used in this study is available as supplemental material (Figs. S1-4, Table S1-9). Additionally, all analysis code is available at Figshare: 10.6084/m9.figshare.28440992, along with additional supplemental material. Custom scripts can also be found at [https://github.com/bananabenana/residue\\_structure\\_colour\\_scripts](https://github.com/bananabenana/residue_structure_colour_scripts).

## Author contributions

Conceptualization: NM, AYP

Data Curation: BRM, BV

Formal Analysis: BRM, BV, HAN

545 Funding Acquisition: NM, AYP  
 546 Methodology: BV, BRM, NM  
 547 Project Administration: NM, AYP  
 548 Resources: NM, AYP  
 549 Supervision: NM, AYP  
 550 Writing – Original Draft Preparation: BRM, BV, NM  
 551 Writing – Review & Editing: BRM, BV, HAN, AG-S, AYP, NM  
 552 Both first-authors (BV, BRM) can present as first-author on paper for grants and resumes.

## 553 Acknowledgements

554 This research/work was supported by Monash eResearch capabilities, including M3 and Research Data  
 555 Storage.

## 556 Funding

557 This work was supported by an NHMRC Emerging Leader 1 Fellowship (APP1176324) awarded to  
 558 N.M., NHMRC Practitioner Fellowship (APP1117940) awarded to A.Y.P., and NIH/NIAID  
 559 R01AI175414 awarded to A.G-S.

## 560 References

- 561 1. WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to  
 562 guide research, development and strategies to prevent and control antimicrobial resistance  
 563 [Internet]. [cited 2024 Oct 9]. Available from:  
 564 <https://www.who.int/publications/i/item/9789240093461>  
 565
- 566 2. Woodford N, Wareham DW, Guerra B, Teale C. Carbapenemase-producing Enterobacteriaceae  
 567 and non-Enterobacteriaceae from animals and the environment: an emerging public health risk of  
 568 our own making? *Journal of Antimicrobial Chemotherapy*. 2014 Feb 1;69(2):287–91.
- 569 3. Bennet KF, Guy RL, Gerver SM, Hopkins KL, Puleston R, Brown CS, et al. Determining the impact  
 570 of professional body recommendations on the screening of acquired carbapenemase-producing  
 571 Enterobacterales in England. *Infection Prevention in Practice*. 2023 Jun;5(2):100281.
- 572 4. Watanabe M, Iyobe S, Inoue M, Mitsuhashi S. Transferable imipenem resistance in *Pseudomonas*  
 573 *aeruginosa*. *Antimicrob Agents Chemother*. 1991 Jan;35(1):147–51.

- 574 5. Kizny Gordon A, Phan HTT, Lipworth SI, Cheong E, Gottlieb T, George S, et al. Genomic  
575 dynamics of species and mobile genetic elements in a prolonged blaIMP-4-associated  
576 carbapenemase outbreak in an Australian hospital. *J Antimicrob Chemother.* 2020 Apr  
577 1;75(4):873–82.
- 578 6. Roberts LW, Catchpoole E, Jennison AV, Bergh H, Hume A, Heney C, et al. Genomic analysis of  
579 carbapenemase-producing Enterobacteriaceae in Queensland reveals widespread transmission of  
580 bla IMP-4 on an IncHI2 plasmid. *Microb Genom.* 2019 Dec 20;6(1):e000321.
- 581 7. Peleg AY, Franklin C, Bell JM, Spelman DW. Dissemination of the metallo-beta-lactamase gene  
582 blaIMP-4 among gram-negative pathogens in a clinical setting in Australia. *Clin Infect Dis.* 2005  
583 Dec 1;41(11):1549–56.
- 584 8. Tsukada M, Miyazaki T, Aoki K, Yoshizawa S, Kondo Y, Sawa T, et al. The outbreak of  
585 multispecies carbapenemase-producing Enterobacterales associated with pediatric ward sinks:  
586 IncM1 plasmids act as vehicles for cross-species transmission. *American Journal of Infection  
587 Control.* 2024 Jul 1;52(7):801–6.
- 588 9. Hirabayashi A, Yahara K, Kajihara T, Sugai M, Shibayama K. Geographical distribution of  
589 Enterobacterales with a carbapenemase IMP-6 phenotype and its association with antimicrobial  
590 use: An analysis using comprehensive national surveillance data on antimicrobial resistance. *PLoS  
591 One.* 2020;15(12):e0243630.
- 592 10. Liu C, Dong N, Zhang Y, Sun Q, Huang Y, Cai C, et al. Phenotypic and genomic  
593 characteristics of clinical IMP-producing *Klebsiella* spp. Isolates in China. *Commun Med.* 2024 Feb  
594 21;4(1):1–9.
- 595 11. Freire MP, de Oliveira Garcia D, Cury AP, Spadão F, Di Gioia TSR, Francisco GR, et al.  
596 Outbreak of IMP-producing carbapenem-resistant *Enterobacter gergoviae* among kidney transplant  
597 recipients. *J Antimicrob Chemother.* 2016 Sep;71(9):2577–85.
- 598 12. Limbago BM, Rasheed JK, Anderson KF, Zhu W, Kitchel B, Watz N, et al. IMP-Producing  
599 Carbapenem-Resistant *Klebsiella pneumoniae* in the United States ▽. *J Clin Microbiol.* 2011  
600 Dec;49(12):4239–45.
- 601 13. Gual-de-Torrella A, López-Causape C, Alejo-Cancho I, Rojo-Molinero E, Sanchez-García G,  
602 Cortés-Lara S, et al. Molecular characterization of a suspected IMP-type carbapenemase-  
603 producing *Pseudomonas aeruginosa* outbreak reveals two simultaneous outbreaks in a tertiary-  
604 care hospital. *Infect Control Hosp Epidemiol.* 2023 Nov;44(11):1801–8.
- 605 14. Wan Y, Myall AC, Boonyasiri A, Bolt F, Ledda A, Mookerjee S, et al. Integrated Analysis of  
606 Patient Networks and Plasmid Genomes to Investigate a Regional, Multispecies Outbreak of  
607 Carbapenemase-Producing Enterobacterales Carrying Both blaIMP and mcr-9 Genes. *Journal of  
608 Infectious Diseases.* 2024 Jul 15;230(1):e159–70.
- 609 15. Izdebski R, Biedrzycka M, Urbanowicz P, Żabicka D, Gniadkowski M. Genome-Based  
610 Epidemiologic Analysis of VIM/IMP Carbapenemase-Producing *Enterobacter* spp., Poland. *Emerg  
611 Infect Dis.* 2023 Aug;29(8):1618–26.
- 612 16. Terrier CL, Gruenig V, Fournier C, Nordmann P, Poirel L. NDM-9 resistance to taniborbactam.  
613 *The Lancet Infectious Diseases.* 2023 Apr 1;23(4):401–2.
- 614 17. David S, Cohen V, Reuter S, Sheppard AE, Giani T, Parkhill J, et al. Integrated chromosomal  
615 and plasmid sequence analyses reveal diverse modes of carbapenemase gene spread among  
616 *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A.* 2020 Oct 6;117(40):25043–54.
- 617 18. Acman M, Wang R, van Dorp L, Shaw LP, Wang Q, Luhmann N, et al. Role of mobile genetic  
618 elements in the global dissemination of the carbapenem resistance gene blaNDM. *Nat Commun.*  
619 2022 Mar 3;13(1):1131.



- 620 19. Macesic N, Hawkey J, Vezina B, Wisniewski JA, Cottingham H, Blakeway LV, et al. Genomic  
621 dissection of endemic carbapenem resistance reveals metallo-beta-lactamase dissemination  
622 through clonal, plasmid and integron transfer. *Nat Commun.* 2023 Aug 8;14(1):4764.
- 623 20. Roberts LW, Harris PNA, Forde BM, Ben Zakour NL, Catchpoole E, Stanton-Cook M, et al.  
624 Integrating multiple genomic technologies to investigate an outbreak of carbapenemase-producing  
625 *Enterobacter hormaechei*. *Nat Commun.* 2020 Jan 24;11(1):466.
- 626 21. Partridge SR, Ginn AN, Paulsen IT, Iredell JR. pEI1573 Carrying blaIMP-4, from Sydney,  
627 Australia, Is Closely Related to Other IncL/M Plasmids. *Antimicrob Agents Chemother.* 2012  
628 Nov;56(11):6029–32.
- 629 22. Pongchaikul P, Mongkolsuk P. Comprehensive Analysis of Imipenemase (IMP)-Type Metallo-  
630  $\beta$ -Lactamase: A Global Distribution Threatening Asia. *Antibiotics (Basel).* 2022 Feb 11;11(2):236.
- 631 23. Macesic N, Blakeway LV, Stewart JD, Hawkey J, Wyres KL, Judd LM, et al. Silent spread of  
632 mobile colistin resistance gene mcr-9.1 on IncHI2 “superplasmids” in clinical carbapenem-resistant  
633 *Enterobacterales*. *Clin Microbiol Infect.* 2021 Dec;27(12):1856.e7-1856.e13.
- 634 24. Treepong P, Kos VN, Guyeux C, Blanc DS, Bertrand X, Valot B, et al. Global emergence of  
635 the widespread *Pseudomonas aeruginosa* ST235 clone. *Clin Microbiol Infect.* 2018 Mar;24(3):258–  
636 66.
- 637 25. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of  
638 plasmids from draft assemblies. *Microbial Genomics [Internet].* 2018 Aug 1 [cited 2024 Oct 9];4(8).  
639 Available from:  
640 <https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000206>
- 641 26. Tavares RDS, Tacão M, Henriques I. Integrons are key players in the spread of beta-  
642 lactamase-encoding genes. *International Journal of Antimicrobial Agents.* 2024 Dec 20;107421.
- 643 27. Bientz L, Guyet U, Guiraud J, Metifiot M, Moulieras M, Aillerie S, et al. Mobilization of an  
644 ICE $clc$ -like element as a potential mechanism for the spread of IMP-13 carbapenemase in  
645 *Pseudomonas aeruginosa*. *Journal of Global Antimicrobial Resistance [Internet].* 2024 Dec 18  
646 [cited 2025 Jan 6]; Available from:  
647 <https://www.sciencedirect.com/science/article/pii/S221371652400465X>
- 648 28. Wang Y, Lo WU, Lai RWM, Tse CWS, Lee RA, Luk WK, et al. IncN ST7 epidemic plasmid  
649 carrying blaIMP-4 in Enterobacteriaceae isolates with epidemiological links to multiple  
650 geographical areas in China. *Journal of Antimicrobial Chemotherapy.* 2017 Jan 1;72(1):99–103.
- 651 29. Pagani L, Colinon C, Migliavacca R, Labonia M, Docquier JD, Nucleo E, et al. Nosocomial  
652 Outbreak Caused by Multidrug-Resistant *Pseudomonas aeruginosa* Producing IMP-13 Metallo- $\beta$ -  
653 Lactamase. *Journal of Clinical Microbiology.* 2005 Aug;43(8):3824–8.
- 654 30. Cheng Z, Bethel CR, Thomas PW, Shurina BA, Alao JP, Thomas CA, et al. Carbapenem Use  
655 Is Driving the Evolution of Imipenemase 1 Variants. *Antimicrobial Agents and Chemotherapy.* 2021  
656 Mar 18;65(4):10.1128/aac.01714-20.
- 657 31. Yamamoto K, Tanaka H, Kurisu G, Nakano R, Yano H, Sakai H. Structural insights into the  
658 substrate specificity of IMP-6 and IMP-1 metallo- $\beta$ -lactamases. *The Journal of Biochemistry.* 2023  
659 Jan 1;173(1):21–30.
- 660 32. Pal A, Tripathi A. An in silico approach for understanding the molecular evolution of clinically  
661 important metallo-beta-lactamases. *Infection, Genetics and Evolution.* 2013 Dec 1;20:39–47.
- 662 33. LaCuran AE, Pegg KM, Liu EM, Bethel CR, Ai N, Welsh WJ, et al. Elucidating the Role of  
663 Residue 67 in IMP-Type Metallo- $\beta$ -Lactamase Evolution. *Antimicrobial Agents and Chemotherapy.*  
664 2015 Nov 17;59(12):7299.



- 665 34. Brown NG, Horton LB, Huang W, Vongpunsawad S, Palzkill T. Analysis of the Functional  
666 Contributions of Asn233 in Metallo- $\beta$ -Lactamase IMP-1. *Antimicrobial Agents and Chemotherapy*.  
667 2011 Dec;55(12):5696.
- 668 35. Iyobe S, Kusadokoro H, Ozaki J, Matsumura N, Minami S, Haruta S, et al. Amino Acid  
669 Substitutions in a Variant of IMP-1 Metallo- $\beta$ -Lactamase. *Antimicrobial Agents and Chemotherapy*.  
670 2000 Aug;44(8):2023.
- 671 36. Dolejska M, Masarikova M, Dobiasova H, Jamborova I, Karpiskova R, Havlicek M, et al. High  
672 prevalence of Salmonella and IMP-4-producing Enterobacteriaceae in the silver gull on Five  
673 Islands, Australia. *Journal of Antimicrobial Chemotherapy*. 2016 Jan 1;71(1):63–70.
- 674 37. Zheng Z, Liu L, Ye L, Xu Y, Chen S. Genomic insight into the distribution and genetic  
675 environment of blaIMP-4 in clinical carbapenem-resistant *Klebsiella pneumoniae* strains in China.  
676 *Microbiol Res*. 2023 Oct;275:127468.
- 677 38. Dolejska M, Papagiannitsis CC, Medvecky M, Davidova-Gerzova L, Valcek A.  
678 Characterization of the Complete Nucleotide Sequences of IMP-4-Encoding Plasmids, Belonging  
679 to Diverse Inc Families, Recovered from Enterobacteriaceae Isolates of Wildlife Origin. *Antimicrob*  
680 *Agents Chemother*. 2018 May;62(5):e02434-17.
- 681 39. Hendrickx APA, Schade RP, Landman F, Bosch T, Schouls LM, van Dijk K. Comparative  
682 analysis of IMP-4- and OXA-58-containing plasmids of three carbapenemase-producing  
683 *Acinetobacter ursingii* strains in the Netherlands. *J Glob Antimicrob Resist*. 2022 Dec;31:207–11.
- 684 40. Ito H, Arakawa Y, Ohsuka S, Wacharotayankun R, Kato N, Ohta M. Plasmid-mediated  
685 dissemination of the metallo-beta-lactamase gene blaIMP among clinically isolated strains of  
686 *Serratia marcescens*. *Antimicrobial Agents and Chemotherapy*. 1995 Apr;39(4):824–9.
- 687 41. Houang ETS, Chu YW, Lo WS, Chu KY, Cheng AFB. Epidemiology of Rifampin ADP-  
688 Ribosyltransferase (arr-2) and Metallo- $\beta$ -Lactamase (blaIMP-4) Gene Cassettes in Class 1  
689 Integrins in *Acinetobacter* Strains Isolated from Blood Cultures in 1997 to 2000. *Antimicrobial*  
690 *Agents and Chemotherapy*. 2003 Apr;47(4):1382–90.
- 691 42. Kanj SS, Kantecki M, Arhin FF, Gheorghe M. Epidemiology and outcomes associated with  
692 MBL-producing Enterobacterales: A systematic literature review. *International Journal of*  
693 *Antimicrobial Agents*. 2025 Apr 1;65(4):107449.
- 694 43. Zhao WH, Hu ZQ. IMP-type metallo- $\beta$ -lactamases in Gram-negative bacilli: distribution,  
695 phylogeny, and association with integrons. *Critical Reviews in Microbiology* [Internet]. 2011 Aug 1  
696 [cited 2025 Feb 20]; Available from:  
697 <https://www.tandfonline.com/doi/abs/10.3109/1040841X.2011.559944>
- 698 44. Matlock W, Shaw LP, Stoesser N. Global genomic epidemiology of blaGES-5  
699 carbapenemase-associated integrons. *Microbial Genomics*. 2024;10(12):001312.
- 700 45. Rui Y, Lu W, Li S, Cheng C, Sun J, Yang Q. Integrons and insertion sequence common  
701 region 1 (ISCR1) of carbapenem-non-susceptible Gram-negative bacilli in fecal specimens from  
702 5000 patients in southern China. *International Journal of Antimicrobial Agents*. 2018 Nov  
703 1;52(5):571–6.
- 704 46. Fu Y, Morris FC, Pereira SC, Kostoulas X, Jiang Y, Vidor C, et al. Mechanisms of blaIMP-4  
705 dissemination across diverse carbapenem-resistant clinical isolates. *Journal of Global*  
706 *Antimicrobial Resistance*. 2025 Mar 1;41:189–94.
- 707 47. Potter RF, Wallace MA, McMullen AR, Prusa J, Stallings CL, Burnham C a. D, et al. blaIMP-  
708 27 on transferable plasmids in *Proteus mirabilis* and *Providencia rettgeri*. *Clinical Microbiology and*  
709 *Infection*. 2018 Sep 1;24(9):1019.e5-1019.e8.

- 710 48. Lee CR, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global Dissemination of  
711 Carbapenemase-Producing *Klebsiella pneumoniae*: Epidemiology, Genetic Context, Treatment  
712 Options, and Detection Methods. *Front Microbiol* [Internet]. 2016 Jun 13 [cited 2025 Apr 16];7.  
713 Available from:  
714 [https://www.frontiersin.orghttps://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2](https://www.frontiersin.orghttps://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2016.00895/full)  
715 016.00895/full
- 716 49. Tada T, Nhung PH, Miyoshi-Akiyama T, Shimada K, Tsuchiya M, Phuong DM, et al.  
717 Multidrug-Resistant Sequence Type 235 *Pseudomonas aeruginosa* Clinical Isolates Producing  
718 IMP-26 with Increased Carbapenem-Hydrolyzing Activities in Vietnam. *Antimicrob Agents*  
719 *Chemother*. 2016 Oct 21;60(11):6853–8.
- 720 50. European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for  
721 interpretation of MICs and zone diameters, version 15.0, 2025 [Internet]. EUCAST; 2025 Jan.  
722 Available from: [http://www.eucast.org/clinical\\_breakpoints/](http://www.eucast.org/clinical_breakpoints/)
- 723 51. Tsang KK, Lam MMC, Wick RR, Wyres KL, Bachman M, Baker S, et al. Diversity, functional  
724 classification and genotyping of SHV  $\beta$ -lactamases in *Klebsiella pneumoniae*. *Microbial Genomics*.  
725 2024;10(10):001294.
- 726 52. Naghavi M, Vollset SE, Ikuta KS, Swetschinski LR, Gray AP, Wool EE, et al. Global burden of  
727 bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *The*  
728 *Lancet*. 2024 Sep 28;404(10459):1199–226.
- 729 53. The NCBI Pathogen Detection Project [Internet]. 2016 [cited 2024 Oct 9]. Available from:  
730 <https://www.ncbi.nlm.nih.gov/pathogens/>
- 731 54. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies  
732 from short and long sequencing reads. *PLOS Computational Biology*. 2017 Jun 8;13(6):e1005595.
- 733 55. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+:  
734 architecture and applications. *BMC Bioinformatics*. 2009 Dec 15;10(1):421.
- 735 56. A G, V S, N V, G T. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*  
736 (Oxford, England) [Internet]. 2013 Apr 15 [cited 2024 Oct 9];29(8). Available from:  
737 <https://pubmed.ncbi.nlm.nih.gov/23422339/>
- 738 57. Gladstone RA, Lo SW, Goater R, Yeats C, Taylor B, Hadfield J, et al. Visualizing variation  
739 within Global Pneumococcal Sequence Clusters (GPSCs) and country population snapshots to  
740 contextualize pneumococcal isolates. *Microbial Genomics* [Internet]. 2020 May 1 [cited 2024 Oct  
741 9];6(5). Available from:  
742 <https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000357>
- 743 58. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014 Jul  
744 15;30(14):2068–9.
- 745 59. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, et al.  
746 AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links  
747 among antimicrobial resistance, stress response, and virulence. *Sci Rep*. 2021 Jun  
748 16;11(1):12728.
- 749 60. Seemann T. tseemann/mlst [Internet]. 2024 [cited 2024 Oct 9]. Available from:  
750 <https://github.com/tseemann/mlst>
- 751 61. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid  
752 genomic surveillance for public health and hospital microbiology labs. *Genome Medicine*. 2014  
753 Nov 20;6(11):90.

754 62. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast  
755 genome and metagenome distance estimation using MinHash. *Genome Biology*. 2016 Jun  
756 20;17(1):132.

757 63. Lees JA, Harris SR, Tonkin-Hill G, Gladstone RA, Lo SW, Weiser JN, et al. Fast and flexible  
758 bacterial genomic epidemiology with PopPUNK. *Genome Res*. 2019 Jan 2;29(2):304–16.

759 64. Harris SR. SKA: Split Kmer Analysis Toolkit for Bacterial Genomic Epidemiology [Internet].  
760 bioRxiv; 2018 [cited 2025 Jan 9]. p. 453142. Available from:  
761 <https://www.biorxiv.org/content/10.1101/453142v1>

762 65. Csárdi G, Nepusz T, Müller K, Horvát S, Traag V, Zanini F, et al. igraph for R: R interface of  
763 the igraph library for graph theory and network analysis [Internet]. Zenodo; 2024 [cited 2024 Oct 9].  
764 Available from: <https://zenodo.org/doi/10.5281/zenodo.7682609>

765 66. Matlock W, Lipworth S, Constantinides B, Peto TEA, Walker AS, Crook D, et al. Flanker: a  
766 tool for comparative genomics of gene flanking regions. *Microb Genom*. 2021 Sep 24;7(9):000634.

767 67. Néron B, Littner E, Haudiquet M, Perrin A, Cury J, Rocha EPC. IntegronFinder 2.0:  
768 Identification and Analysis of Integrons across Bacteria, with a Focus on Antibiotic Resistance in  
769 *Klebsiella*. *Microorganisms*. 2022 Apr;10(4):700.

770 68. Xie Z, Tang H. ISEScan: automated identification of insertion sequence elements in  
771 prokaryotic genomes. *Bioinformatics*. 2017 Nov 1;33(21):3340–7.

772 69. Liu M, Li X, Xie Y, Bi D, Sun J, Li J, et al. ICEberg 2.0: an updated database of bacterial  
773 integrative and conjugative elements. *Nucleic Acids Research*. 2019 Jan 8;47(D1):D660–5.

774 70. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018 Sep  
775 15;34(18):3094–100.

776 71. Teufel F, Armenteros JJA, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0  
777 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*.  
778 2022 Jan 3;40(7):1023.

779 72. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein  
780 sequences. *Protein Science*. 2018;27(1):135–45.

781 73. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for  
782 the comparative analysis of protein structures. *Bioinformatics*. 2006 Nov 1;22(21):2695–6.

783 74. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making  
784 protein folding accessible to all. *Nat Methods*. 2022 Jun;19(6):679–82.

785 75. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate  
786 protein structure prediction with AlphaFold. *Nature*. 2021 Aug;596(7873):583–9.

787 76. Gilchrist CLM, Mirdita M, Steinegger M. Multiple Protein Structure Alignment at Scale with  
788 FoldMason [Internet]. bioRxiv; 2024 [cited 2025 Jan 9]. p. 2024.08.01.606130. Available from:  
789 <https://www.biorxiv.org/content/10.1101/2024.08.01.606130v1>

790 77. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF  
791 ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*.  
792 2018;27(1):14–25.

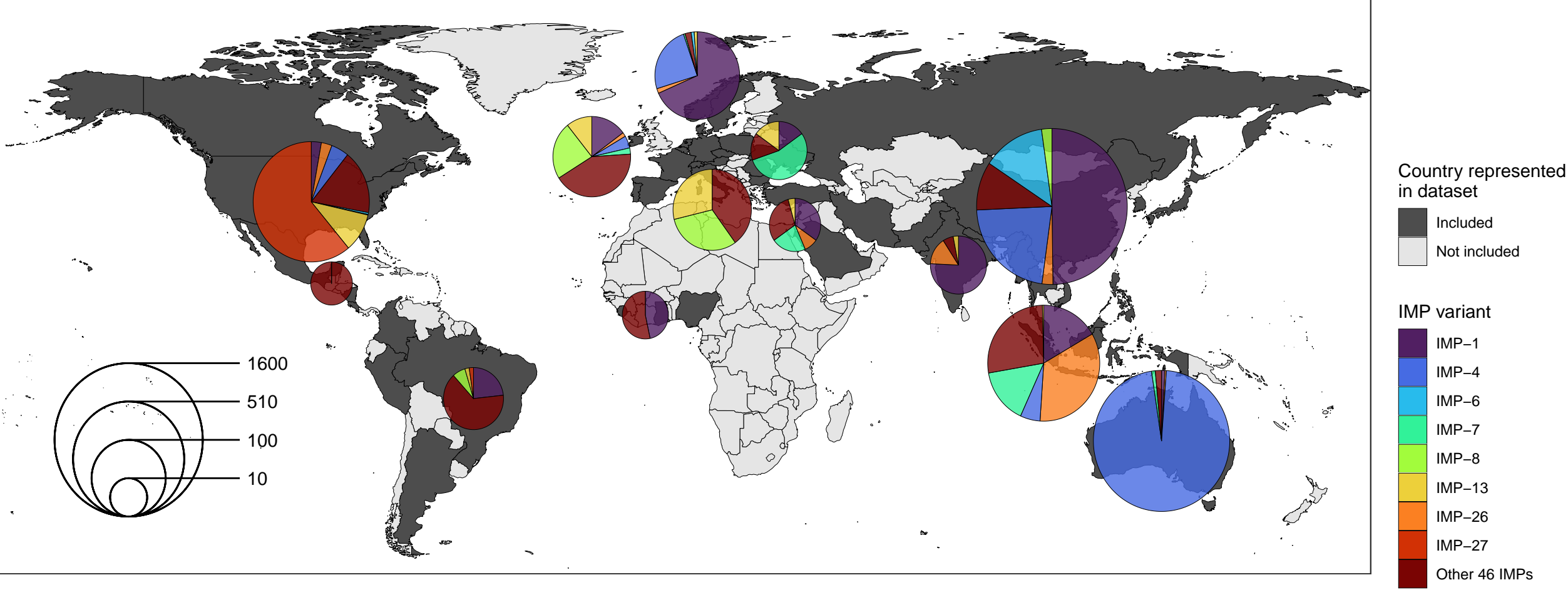
793 78. Vezina B. ChimeraX\_scripts [Internet]. Zenodo; 2025. Available from:  
794 <https://doi.org/10.5281/zenodo.14891324>

- 795 79. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna,  
796 Austria: R Foundation for Statistical Computing; 2024. Available from: <https://www.R-project.org/>
- 797 80. Posit team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: Posit  
798 Software, PBC; 2024. Available from: <http://www.posit.co/>
- 799 81. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the  
800 Tidyverse. Journal of Open Source Software. 2019 Nov 21;4(43):1686.
- 801 82. Zeileis A, Fisher JC, Hornik K, Ihaka R, McWhite CD, Murrell P, et al. colorspace: A Toolbox  
802 for Manipulating and Assessing Colors and Palettes. J Stat Soft [Internet]. 2020 [cited 2025 Jan  
803 9];96(1). Available from: <http://arxiv.org/abs/1903.06490>
- 804 83. Garnier S, Ross N, Rudis B, Sciaini M, Camargo AP, Scherer C. viridis: Colorblind-Friendly  
805 Color Maps for R [Internet]. 2024 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/viridis/index.html)  
806 [project.org/web/packages/viridis/index.html](https://cran.r-project.org/web/packages/viridis/index.html)
- 807 84. Brand T van den. teunbrand/ggh4x [Internet]. 2025 [cited 2025 Jan 9]. Available from:  
808 <https://github.com/teunbrand/ggh4x>
- 809 85. Sjoberg D. ggstream: Create Streamplots in “ggplot2” [Internet]. 2021 [cited 2025 Jan 9].  
810 Available from: <https://cran.r-project.org/web/packages/ggstream/index.html>
- 811 86. Deckmyn OS code by RAB and ARWR version by RBE by TPM and A. maps: Draw  
812 Geographical Maps [Internet]. 2024 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/maps/index.html)  
813 [project.org/web/packages/maps/index.html](https://cran.r-project.org/web/packages/maps/index.html)
- 814 87. Yu G, Xu S. scatterpie: Scatter Pie Plot [Internet]. 2024 [cited 2025 Jan 9]. Available from:  
815 <https://cran.r-project.org/web/packages/scatterpie/index.html>
- 816 88. Pebesma E, Bivand R, Racine E, Sumner M, Cook I, Keitt T, et al. sf: Simple Features for R  
817 [Internet]. 2024 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/sf/index.html)  
818 [project.org/web/packages/sf/index.html](https://cran.r-project.org/web/packages/sf/index.html)
- 819 89. Massicotte P, South A, Hufkens K. rnaturalearth: World Map Data from Natural Earth  
820 [Internet]. 2023 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/rnaturalearth/index.html)  
821 [project.org/web/packages/rnaturalearth/index.html](https://cran.r-project.org/web/packages/rnaturalearth/index.html)
- 822 90. Campitelli E. ggnewscale: Multiple Fill and Colour Scales in “ggplot2” [Internet]. 2024 [cited  
823 2025 Jan 9]. Available from: <https://cran.r-project.org/web/packages/ggnewscale/index.html>
- 824 91. Wilkins D, Rudis B. treemapify: Draw Treemaps in “ggplot2” [Internet]. 2023 [cited 2025 Jan  
825 9]. Available from: <https://cran.r-project.org/web/packages/treemapify/index.html>
- 826 92. Pedersen TL. patchwork: The Composer of Plots [Internet]. 2024 [cited 2025 Jan 9]. Available  
827 from: <https://cran.r-project.org/web/packages/patchwork/index.html>
- 828 93. Epskamp S, Costantini G, Haslbeck J, Isvoranu A, Cramer AOJ, Waldorp LJ, et al. qgraph:  
829 Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation  
830 [Internet]. 2023 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/qgraph/index.html)  
831 [project.org/web/packages/qgraph/index.html](https://cran.r-project.org/web/packages/qgraph/index.html)
- 832 94. Pedersen TL, RStudio. ggraph: An Implementation of Grammar of Graphics for Graphs and  
833 Networks [Internet]. 2024 [cited 2025 Jan 9]. Available from: [https://cran.r-](https://cran.r-project.org/web/packages/ggraph/index.html)  
834 [project.org/web/packages/ggraph/index.html](https://cran.r-project.org/web/packages/ggraph/index.html)
- 835 95. Pedersen TL, RStudio. ggforce: Accelerating “ggplot2” [Internet]. 2024 [cited 2025 Jan 9].  
836 Available from: <https://cran.r-project.org/web/packages/ggforce/index.html>

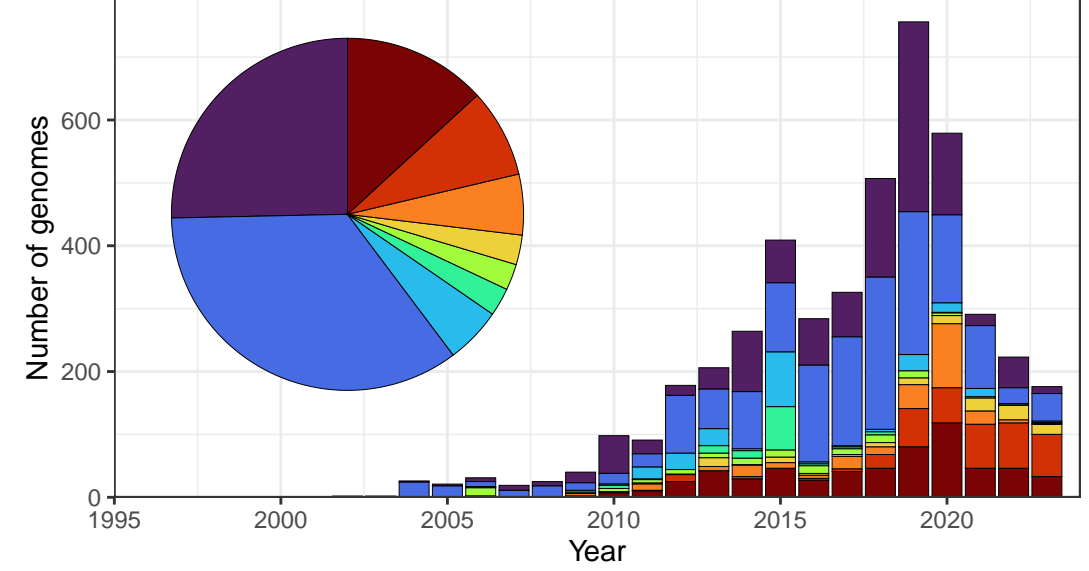
- 837 96. Brunson JC, Read QD. ggalluvial: Alluvial Plots in “ggplot2” [Internet]. 2023 [cited 2025 Jan  
838 9]. Available from: <https://cran.r-project.org/web/packages/ggalluvial/index.html>
- 839 97. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an r package for visualization and  
840 annotation of phylogenetic trees with their covariates and other associated data. Methods in  
841 Ecology and Evolution. 2017;8(1):28–36.
- 842 98. Wang LG, Lam TTY, Xu S, Dai Z, Zhou L, Feng T, et al. Treeio: An R Package for  
843 Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. Molecular Biology  
844 and Evolution. 2020 Feb 1;37(2):599–603.
- 845 99. Yu G, Xu S, Hackl T. aplot: Decorate a “ggplot” with Associated Information [Internet]. 2024  
846 [cited 2025 Jan 9]. Available from: <https://cran.r-project.org/web/packages/aplot/index.html>

847

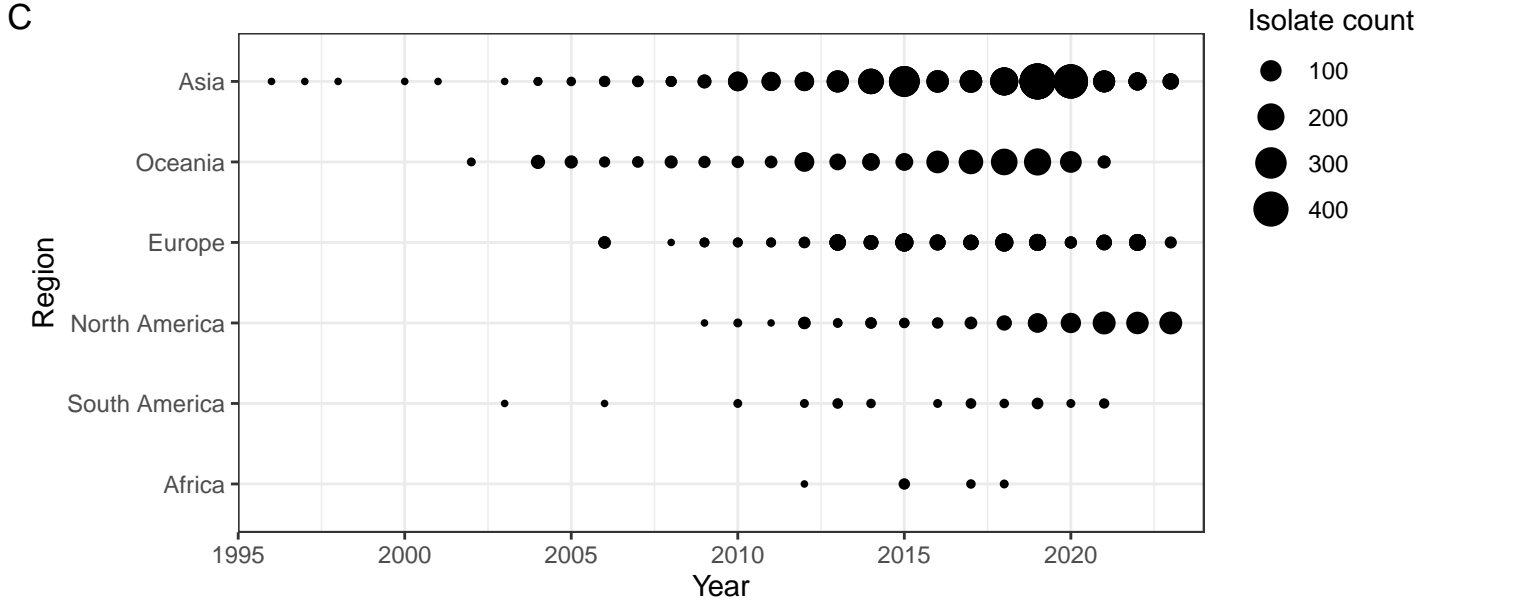
A



B

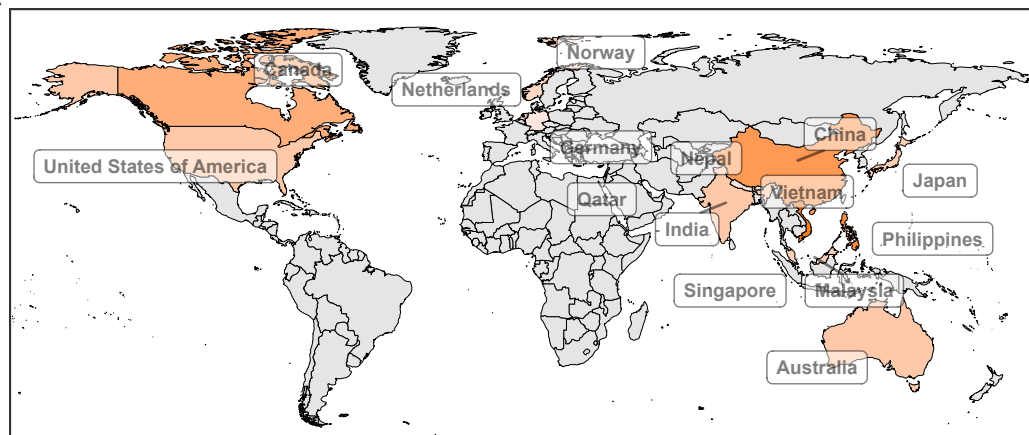


C

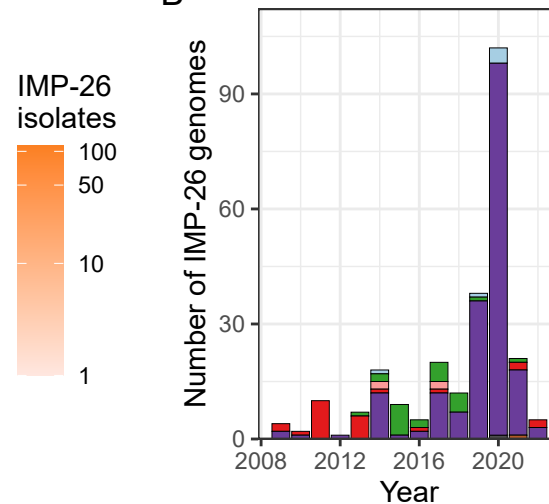




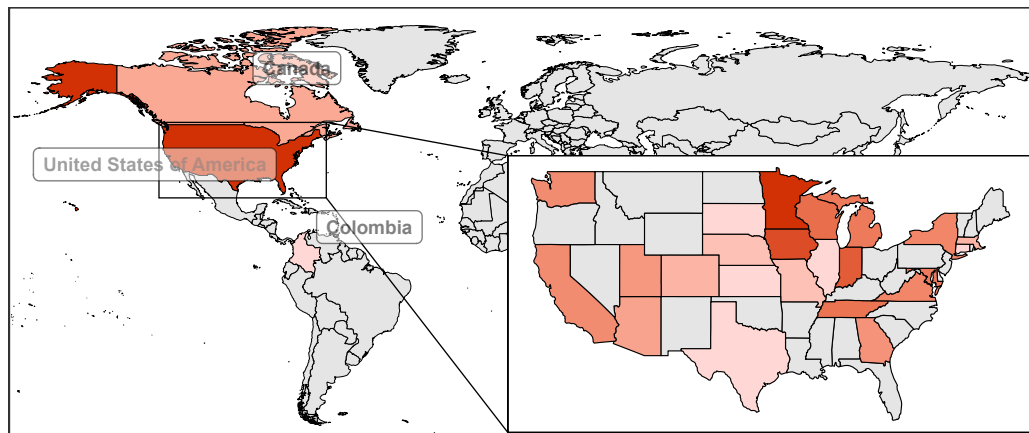
A



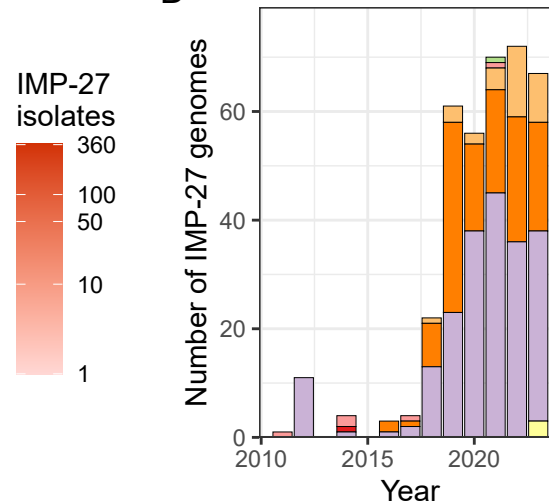
B



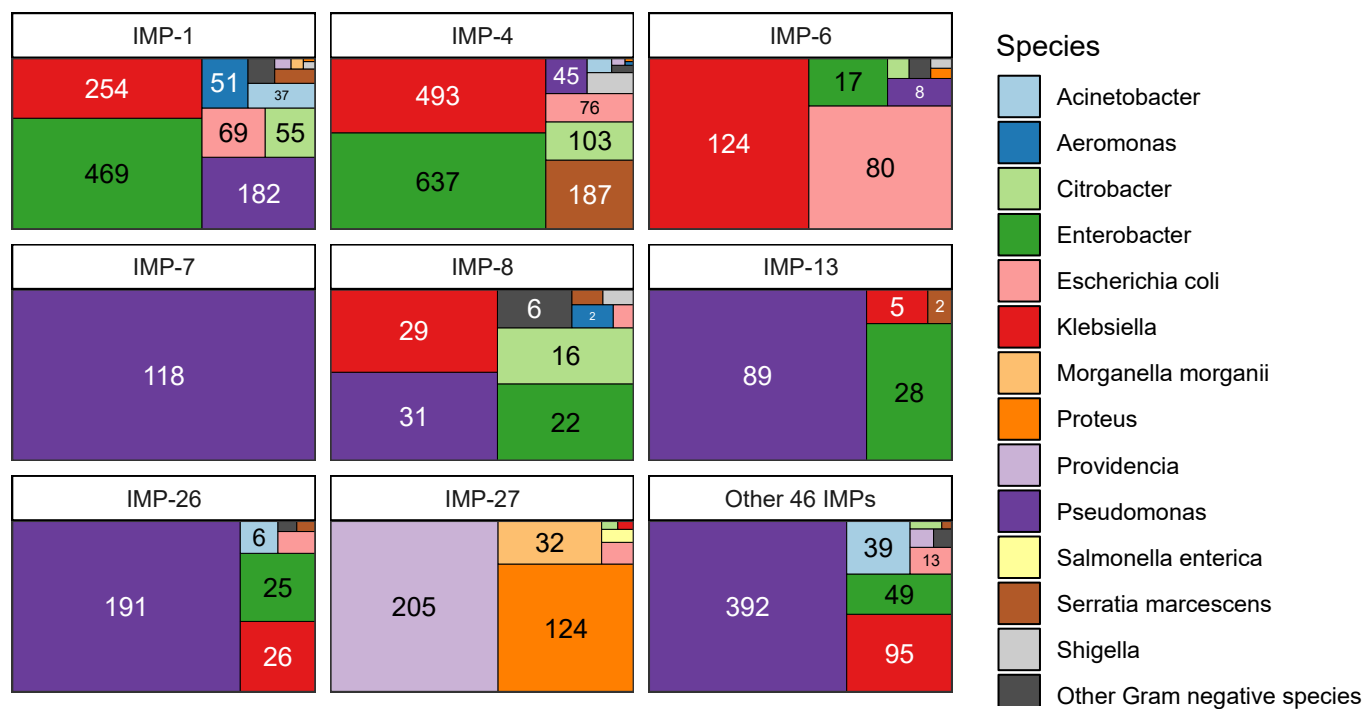
C



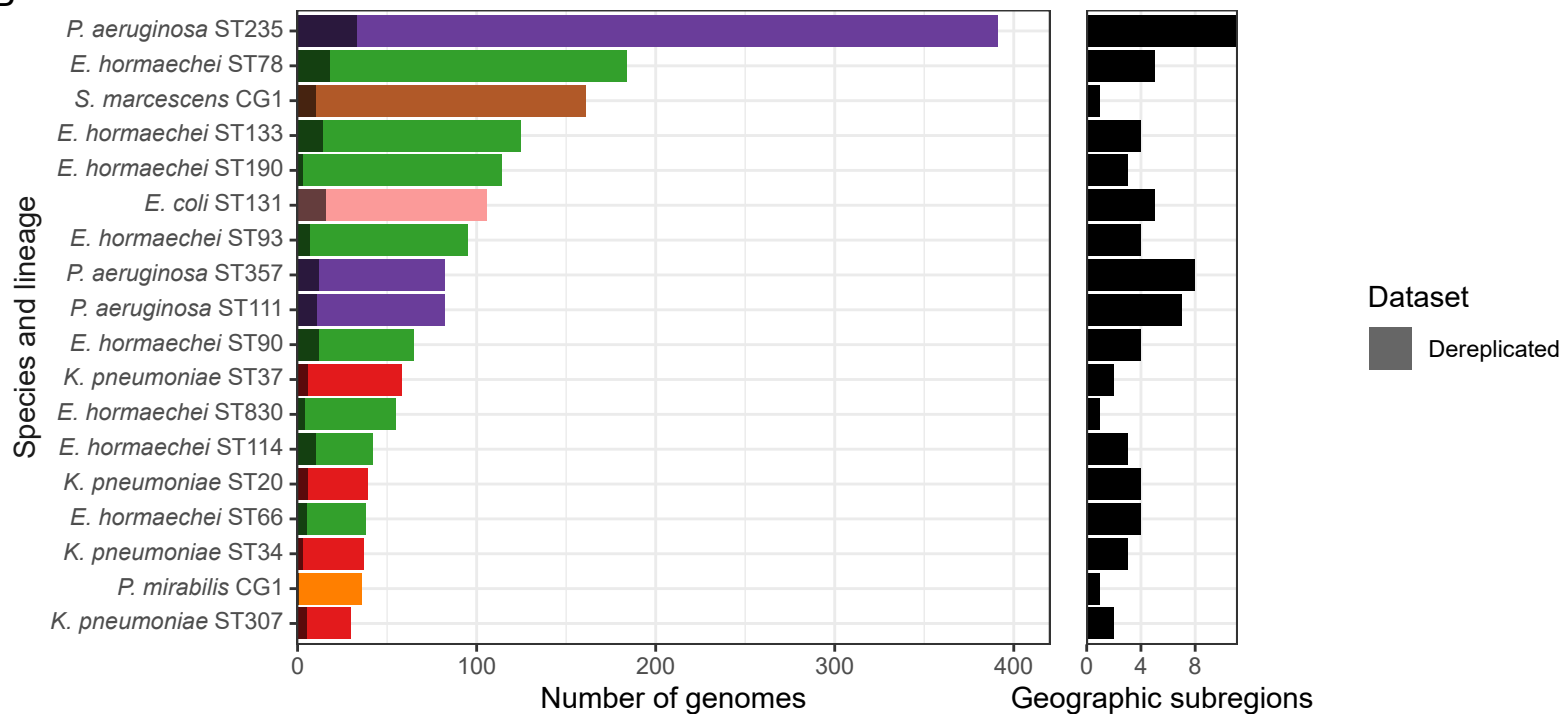
D



A

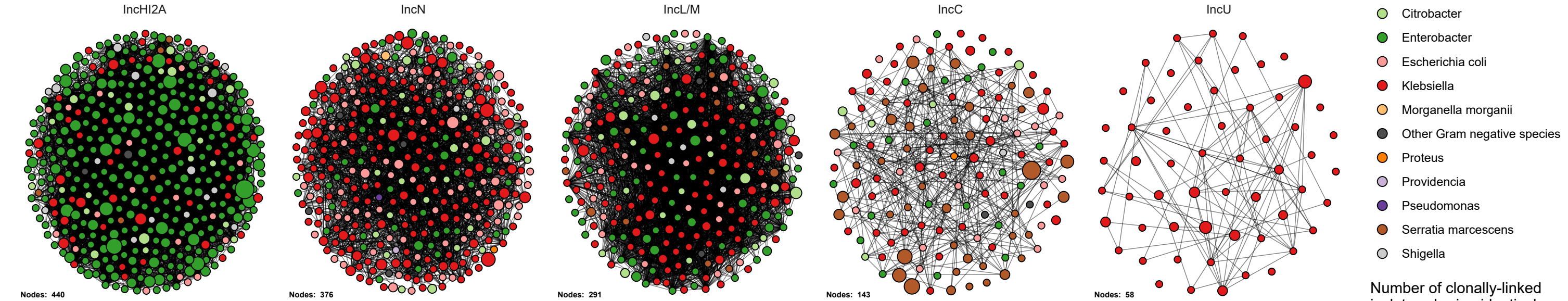


B

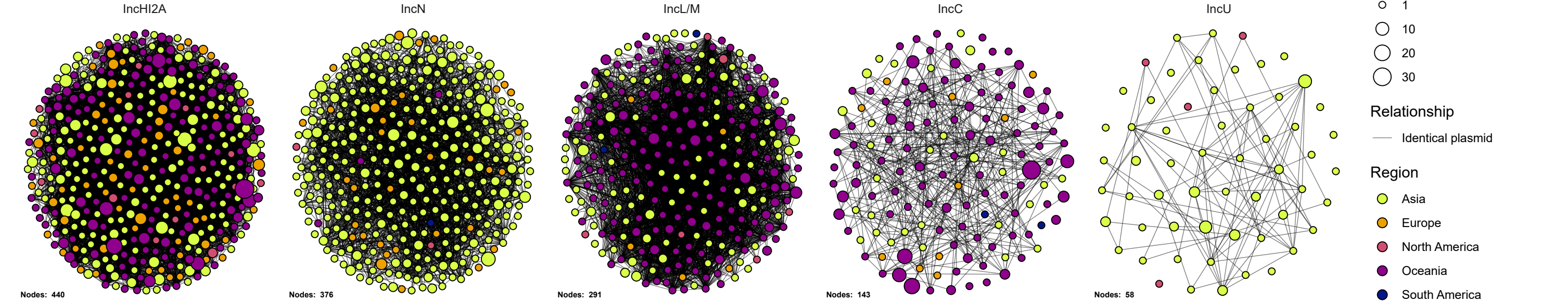


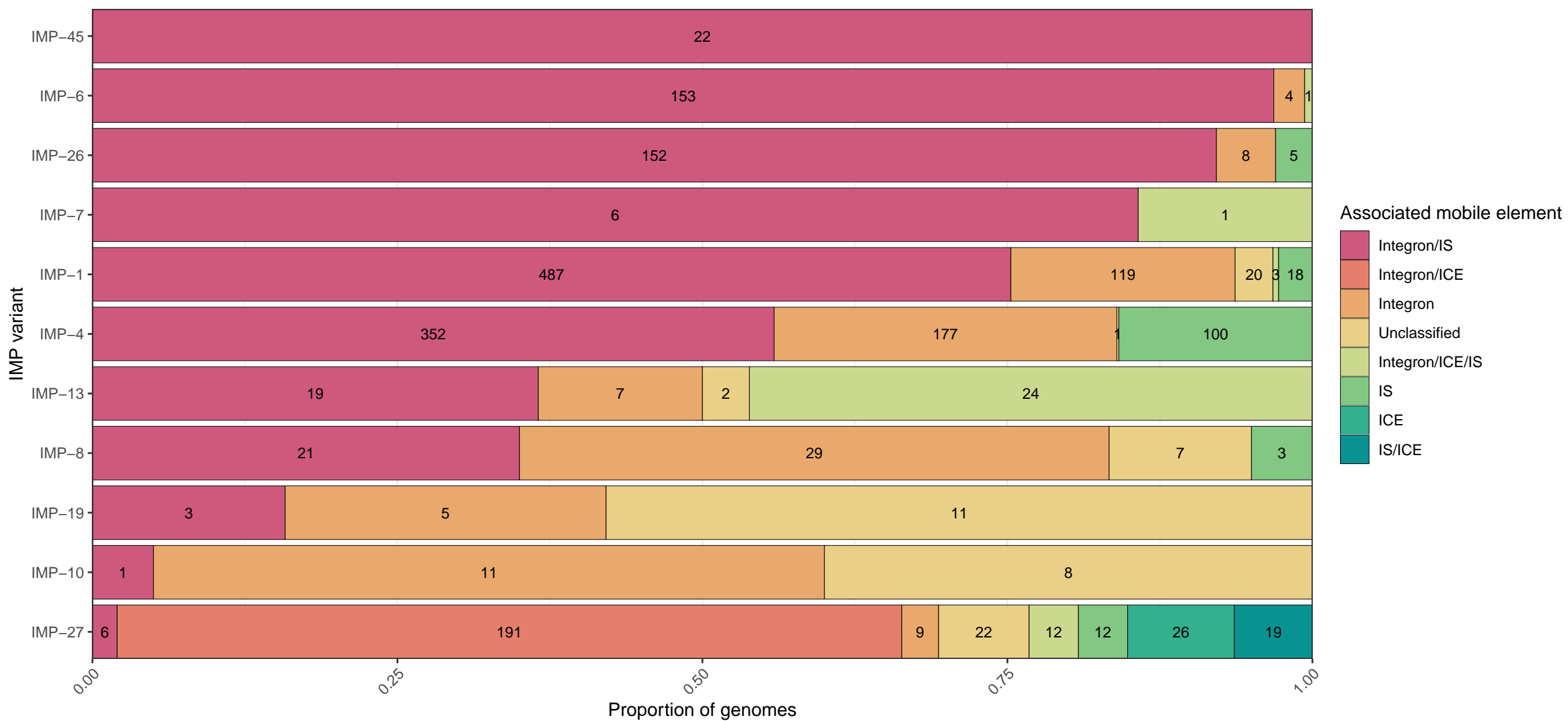


A - Species

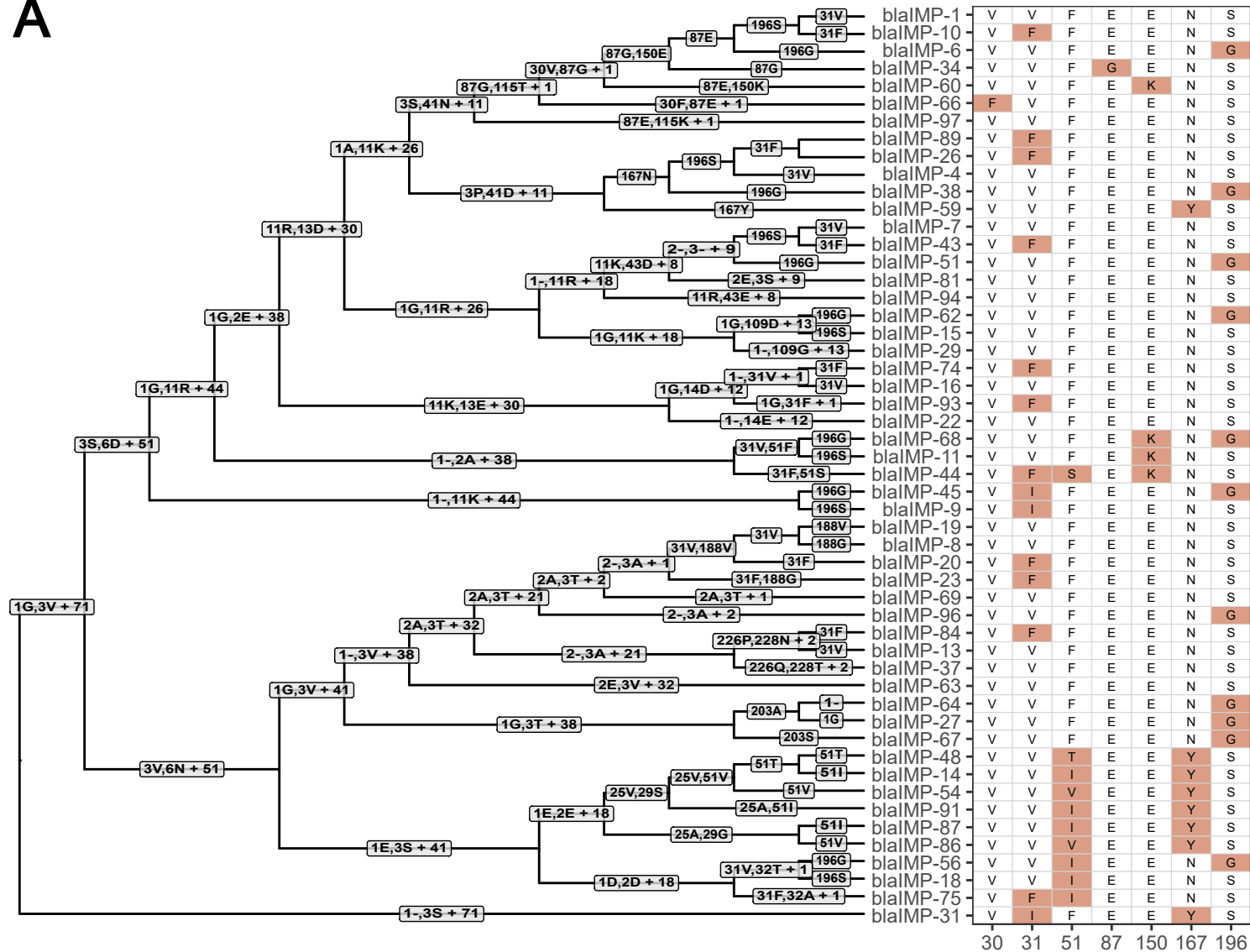


B - Regions



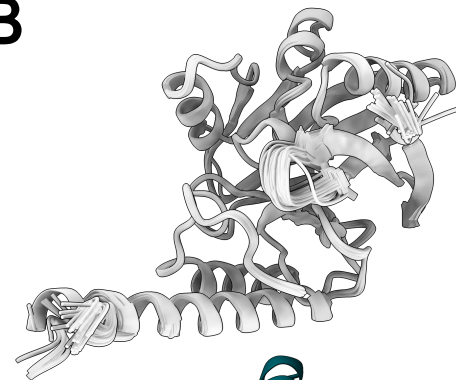


A

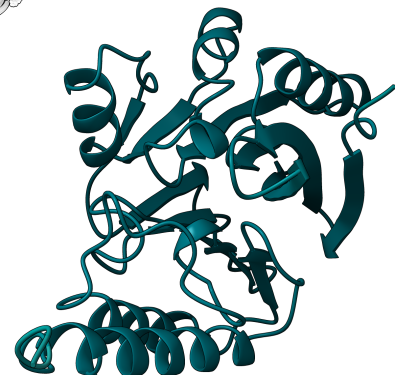


Key positions in Multiple Sequence Alignment

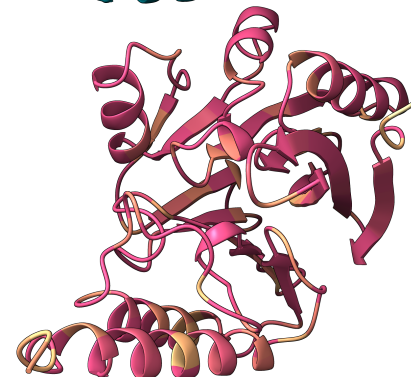
B



C



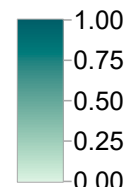
D



Residue



IDDT score



Conservation score

