UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie
"L. Spallanzani"

Genomic epidemiology of emerging and re-emerging pathogens

Greta Petazzoni

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
Ciclo XVIII – A.A. 2023-2025

Minds are like parachutes: they only function when open
- Thomas Dewar

Table of contents

ABSTRACT **1**

ABBREVIATIONS **2**

CHAPTER 1: **4**

Introduction

1. The epidemiology of healthcare-associated infections **4**

1.1 Healthcare-associated infections: epidemiology and etiology 4

1.2 A focus on sepsis 6

1.3 Antimicrobial resistance 7

1.4 Italian epidemiology of multidrug-resistant microorganisms 10

1.5 Surveillance of HAIs: an italian experience 11

2. Genomic epidemiology **13**

2.1 The evolution of genomics 13

2.2 Next-Generation Sequencing in clinical microbiology 14

2.3 Genomic-based surveillance 16

2.3.1 Bacterial typing and transmission dynamics 17

2.3.2 Antimicrobial resistance and virulence genes 18

2.3.3 Horizontal gene transfer: a focus on plasmids 19

CHAPTER 2: **21**

Aim of the research

CHAPTER 3: **23**

Longitudinal phylogenomic surveillance of carbapenem-resistant Klebsiella pneumoniae reveals clonal replacement and segregated plasmid dissemination

Abstract **24**

Introduction **25**

Materials and Methods **26**

Study design, isolate selection, and sequencing 26

Genomic characterization, phylogeny, and statistical analysis 27

Plasmid landscape analysis 28

Results **29**

Carbapenem-resistant *K. pneumoniae* population 29

Genomic epidemiology of CR-Kp 30
Phylogeny and transmission dynamics 31
Plasmidome landscape 33
Discussion **37**
Conclusion **39**
Acknowledgment **39**
Data availability **39**
Supplementary material **39**
Ethical statement **39**
References **40**

CHAPTER 4: **43**
Empowering active surveillance of carbapenem-resistant Enterobacterales: insights from one year of genomic surveillance

Abstract **44**
Introduction **45**
Materials and methods **46**
Criteria of inclusion 46
DNA extraction and sequencing 46
In vitro evaluation of carbapenem resistance retention capability 47
Results **48**
Dataset characterisation 48
Transmission clusters 50
Klebsiella pneumoniae *50*
K. pneumoniae ST307 *51*
Other highly prevalent CR-Kp clones 54
Virulence factors and Inc groups in *K. pneumoniae* 56
Klebsiella spp. *57*
Escherichia coli *57*
E. coli ST131 *57*
Other prevalent CR-Ec clones 58
Other species 58
Discussion **59**
Conclusion **60**
Acknowledgement **61**
Data availability **61**
Ethical statement **61**

References **61**

CHAPTER 5: **64**

Rapid spread of a novel NDM-producing clone of Klebsiella pneumoniae CC147, Northern Italy, February to August 2023

Introduction **65**

Event description **65**

Clone characterisation **67**

Antimicrobial susceptibility profiling 67

Typing and phylogenetic analysis 67

Strain characterisation 69

Discussion **70**

Conclusion **71**

Data availability **71**

Ethical statement **71**

References **71**

CHAPTER 6: **74**

The COVID-19 pandemic sparked off a large-scale outbreak of carbapenem-resistant Acinetobacter baumannii from the endemic strains at an italian hospital

Abstract **75**

Importance **75**

Introduction **76**

Results **77**

Dataset characterisation 77

Phylogenetic analyses 78

Antimicrobial resistance profile 79

Virulence determinants and biofilm formation 80

Plasmidome analysis 80

Discussion **81**

Materials and methods **83**

Criteria of inclusion 83

Genomic DNA extraction and sequencing 83

Plasmidome characterisation 84

Hospital surveillance program 84

Biofilm production assay 84

Phylogenetic analyses 85

Acknowledgment **86**

Data availability **86**
Ethical statement **86**
References **86**

CHAPTER 7: **91**
Time is of the essence: a new 24/7 laboratory organization and a fast-track diagnostic pathway for blood culture of critical patients

Abstract **92**
Introduction **93**
Materials and methods **94**
Definitions 94
Study design 95
Laboratory workflow 96
Data processing and statistical analysis 97
Results **98**
Gram stain turnaround time (Gram-TAT) 99
Microorganism identification turnaround time (ID-TAT) 102
Resistance markers turnaround time (Marker-TAT) 105
Discussion **105**
Conclusion **107**
Acknowledgement **107**
Ethical statement **107**
References **107**

CHAPTER 8: **111**
Gut microbiota and B Cell receptor inhibitors for the treatment of Chronic Lymphocytic Leukemia: Is biodiversity correlated with clinical response or immune-related adverse event occurrence? A Cross-Sectional study

Abstract **112**
Introduction **113**
Materials and methods **115**
Study design and patient cohort 115
Objectives 115
Sample collection and sequencing 116
Bioinformatic and statistical analysis 116
Results **117**
Taxonomic profiling and diversity in clinical response groups 118
Taxonomic profiling and diversity in relation to immune-related

adverse events 121
Discussion **125**
Conclusions **126**
Acknowledgments **127**
Ethical statement **127**
Data availability **127**
References **127**

CHAPTER 9: **130**
General discussion

REFERENCES **134**

ACKNOWLEDGMENT **142**

ABSTRACT

The application of Next-Generation Sequencing technologies and advanced bioinformatics is becoming increasingly widespread in clinical microbiology to enhance epidemiological surveillance and optimize patient management. This PhD thesis explores its integration in practice and is articulated through three levels of investigation: genomic surveillance, outbreak investigation, and translational clinical applications.

The first level (Chapters 3 and 4) explores the long- and short-term epidemiology of MDR pathogens. Chapter 3 presents nine-year genomic surveillance of carbapenem-resistant *Klebsiella pneumoniae* which revealed a dynamic population structure. The population was characterized by the replacement of historical lineages with emerging high-risk clones, whose evolutionary success is sustained by high plasmid plasticity. Chapter 4 extends the analysis to all carbapenem-resistant *Enterobacterales* isolated during one year, highlighting the role of low-prevalence species as cryptic resistance reservoirs.

The second level (Chapters 5 and 6) applies high-resolution genomics to outbreak investigation. Chapter 5 describes the emergence of the NDM-1-producing ST6668 clone, marking a critical local epidemiological shift towards the dissemination of metallo-beta-lactamases. Chapter 6 presents the investigation of an *Acinetobacter baumannii* outbreak during the COVID-19 pandemic. This analysis showed that transmission was not driven by intrinsic bacterial characteristics, but rather by the extrinsic pressure exerted by the health emergency on infection control practices.

Finally, the third level (Chapters 7 and 8) shifts the focus from epidemiology to precision medicine. In Chapter 7 is evaluated the clinical impact of reorganizing the laboratory into a 24/7 workflow, demonstrating how reduced Turnaround Times facilitate timely antimicrobial stewardship. In Chapter 8, targeted metagenomics was applied to study the gut microbiota in hematological patients, identifying taxonomic signatures associated with therapeutic response, suggesting the potential of the microbiota as a non-invasive predictive biomarker.

ABBREVIATIONS

AE	Adverse events
AMR	Antimicrobial resistance
ARG	Antimicrobial resistant gene
AST	Antimicrobial susceptibility testing
BC	Blood culture
BCbs	Blood culture bottles
BCRi	B-cell receptor inhibitors
BSI	Bloodstream infection
BTK	Bruton tyrosine kinase
CC	Clonal complex
CG	Clonal group
CLL	Chronic lymphocytic leukemia
CML	Clinical Microbiology Laboratory
CR-Ab	Carbapenem-resistant *Acinetobacter baumannii*
CR-Ec	Carbapenem-resistant *Escherichia coli*
CRE	Carbapenem-resistant *Enterobacterales*
CR-Kp	Carbapenem-resistant *Klebsiella pneumoniae*
CUS	Chaperone-Usher secretion
ECDC	European Centre for Disease Prevention and Control
ECs	Epidemic clusters
E.R.	Emergency Room
ESBL	Extended-spectrum β-lactamase
EUCAST	European Committee on Antimicrobial Susceptibility Testing
EU/EEA	European Union and European Economic Area
GM	Gut microbiota
GNB	Gram-negative bacilli
GPB	Gram-positive bacilli
GPC	Gram-positive cocci
GPCC	Gram-positive cocci in chains
Gram-TAT	Gram stain turnaround time
HAIs	Healthcare-associated infections
HSM	Fondazione IRCCS Policlinico San Matteo (Pavia, Italy)
ICIs	Immune checkpoint inhibitors
ICUs	Intensive Care Units
ID-TAT	Microorganism identification turnaround time
IPC	Infection prevention and control

irAEs	Immune-related adverse events
Kp	*Klebsiella pneumoniae*
KPC	Klebsiella pneumoniae carbapenemase
LIS	Laboratory Information System
LOS	Length of stay
LPD	Chronic lymphoproliferative disorders
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation Time-Of-Flight
Marker-TAT	Resistance markers turnaround time
MDR	Multidrug-resistant
MEWS	Modified Early Warning Score
MIC	Minimum Inhibitory Concentration
MLST	Multilocus sequence typing
MRSA	Metchillin-resistant *Staphylococcus aureus*
NAE	Non-Adverse Events
NDM	New Delhi metallo-β-lactamase
NEWS	National Early Warning Score
NGS	Next-Generation Sequencing
Non Urg-BC	Non urgent blood culture
NR	Non-Responder Group
OXA	Oxacillinase
Pavia-HAU	Pavia Health Authority Unit
PC	Plasmidome clusters
PCoA	Principal Coordinate Analysis
PFS	Progression-free survival
PI3K	Phosphatidylinositol 3-kinase
PNAG	Poly-β-(1,6)-N-acetylglucosamine
PPE	Personal protective equipment
qSOFA	quick Sequential Organ Failure Assessment
R	Responder Group
RI	Responder Group treated with Ibrutinib
RT-PCR	Real-time reverse transcriptase PCR
SCFAs	Short Chain Fatty Acids
SLV	Single-locus variant
SNP	Single Nucleotide Polymorphism
ST	Sequence Type
TATs	Turnaround times
TCs	Transmission clusters
TGS	Third-Generation Sequencing
TTD	Time to detection
Urg-BC	Urgent Blood Cultures
WGS	Whole-genome sequencing
WHO	World Health Organization

CHAPTER 1:

Introduction

1. The epidemiology of healthcare-associated infections

Modern healthcare settings are characterized by a complex dual nature: designed as structures for healing, they simultaneously function as distinct ecological niches for the transmission of infectious agents. Within this unique environment, pathogens do not exist in isolation but circulate through a dynamic "chain of transmission" that bridges patients, healthcare personnel, and surfaces [1]. Unlike community settings, the hospital ecosystem exerts a selective pressure driven by high antimicrobial consumption level. This pressure influences the composition of the local microbial communities, favoring the selection and proliferation of multidrug-resistant (MDR) organisms capable of persisting on patients and the environment [2]. Consequently, the hospital itself becomes a persistent reservoir, necessitating a deeper understanding of microbial ecology and developing effective infection prevention and control strategies to safeguard public health.

1.1 Healthcare-associated infections: epidemiology and etiology

Healthcare-associated infections (HAIs), historically known as nosocomial infections, represent a group of infectious diseases acquired during the process of receiving medical care. By definition, these infections are absent at the time of admission and typically manifest after a latency period of at least 48 to 72 hours; they may also present clinically after patient discharge, thereby complicating diagnostic and therapeutic management.

The impact of HAIs is substantial both clinically and economically. According to the most recent surveillance data released by the European Centre for Disease Prevention and Control (ECDC), it is estimated that approximately 4.3 million patients annually acquire at least one HAI within hospitals across the European Union and European Economic Area (EU/EEA), corresponding to 7.1% of total hospitalized patients (Figure 1) [3]. The acquisition of an HAI triggers a cascade of complications, such as prolonged hospitalization, elevated rates of morbidity and mortality, and

increased therapeutic costs. Consequently, this clinical burden translates also into an economic deficit; financial estimates indicate that HAIs cost European healthcare systems at least 7 billion euros per year, diverting essential resources away from standard patient care [4].

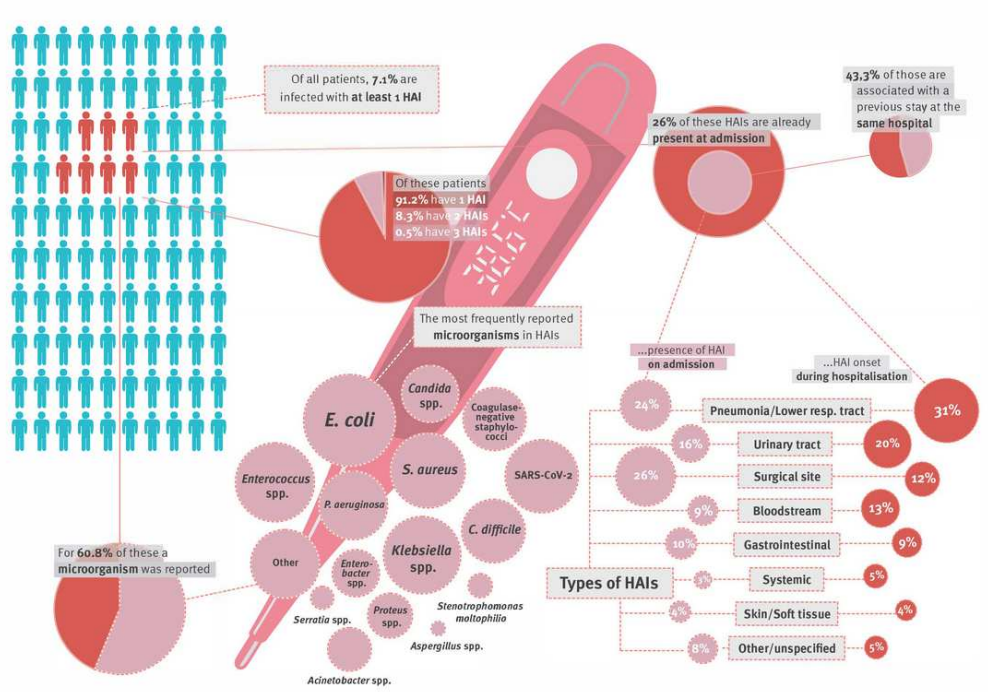


Figure 1 - Infographic of ECDC based on Point Prevalence survey (2022-2023 data) regarding HAIs in EU/EEA hospitals [5].

In terms of clinical presentation, HAIs most frequently compromise the respiratory and urinary tracts, surgical sites, and the bloodstream, alongside gastrointestinal tract. While viruses and fungi act as contributing pathogens, bacteria remain the predominant etiological agents, responsible for approximately 90% of all cases [6]. Within this bacterial landscape, the most significant therapeutic challenge is posed by the ESKAPE group (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species). As the acronym suggests, these bacteria often possess capabilities to "escape" the action of antimicrobial agents, driving the global antimicrobial resistance crisis.

Recognizing this threat, the World Health Organization (WHO) prioritizes these bacteria in its *Bacterial Priority Pathogens List*. Notably, the 2024

update classifies carbapenem-resistant *A. baumannii* (CR-Ab) and carbapenem-resistant *Enterobacterales* (including *Klebsiella* and *Enterobacter spp.*) as "Critical Priority" targets, a designation that underscores the urgent necessity for novel therapeutic development to prevent untreatable infections [7].

1.2 A focus on sepsis

The clinical progression of HAIs can culminate in severe and life-threatening complications, most notably sepsis. According to the Third International Consensus Definitions (Sepsis-3), sepsis is defined as life-threatening organ dysfunction caused by a dysregulated host response to infection [8]. While clinical identification relies on warning scores such as MEWS, NEWS, or qSOFA [9], it is crucial to distinguish this syndrome from bloodstream infection (BSI). Although frequently co-occurring and strongly correlated, they represent distinct entities: sepsis is a systemic physiological manifestation, whereas BSI is a strictly microbiological event characterized by the presence of viable pathogens in the blood. The latter is one of the most critical infectious conditions and is the most prevalent cause of sepsis.

The mortality associated with this condition is time-dependent and severity-dependent: sepsis mortality is estimated between 10% and 20% [10] but it rises precipitously to more than 30% in cases of septic shock [11]. Its management is dependent on early recognition and the timeliness of effective treatment. Literature suggests that even a one-hour delay in initiating appropriate antibiotic therapy adversely affects patient survival in a significant way [10, 12]. Therefore, prompt administration of empiric antimicrobial therapy is essential to mitigate this risk, a strategy that frequently relies on broad-spectrum antibiotics. The choice of the empiric therapy is informed by the infection site, patient history, and local epidemiological data. However, in the current antimicrobial resistance era, this empiric approach faces a challenge: if the causative agent is a MDR organism, the empiric therapy demands a broader therapeutic coverage, relying on the combinations of multiple antibiotics [13].

In addition, the rapid clinical response must be mirrored by laboratory efficiency. Blood cultures (BCs) remain the gold standard for the microbiological diagnosis of BSIs. The timely and accurate identification of the causative pathogens, along with the detection of resistance mechanisms, is essential to transition from empiric to targeted therapy [13].

The capability of modern microbiology to provide timely results relies on the integration of rapid technologies with traditional methods. Indeed, automated, continuous-monitoring blood culture systems (such as the BD BACTEC™ FX), analyze bottles to detect bacterial growth at the earliest

possible moment. Upon a positive signal, the diagnostic workflow is immediately initiated with Gram staining. In critical cases, molecular syndromic panels (e.g., multiplex PCR as FilmArray) may be employed directly from the positive BCs to detect species and resistance genes. However, standard procedure involves culturing into solid media, and once microbial growth is observed, rapid pathogen identification could be achieved via, for instance, Matrix Assisted Laser Desorption/Ionisation Time-Of-Flight (MALDI-TOF) mass spectrometry. Depending on the isolated pathogen, the identification is often complemented by immunochromatographic tests for the detection of specific resistance enzymes (e.g., carbapenemases). Ultimately, regardless of the rapid methods used, a complete antimicrobial susceptibility testing (AST) is always performed to determine the minimum inhibitory concentration (MIC) for guiding the final therapeutic optimization.

1.3 Antimicrobial resistance

Antimicrobial resistance (AMR) is a natural evolutionary phenomenon wherein microorganisms, including bacteria, viruses, and fungi, undergo genomic modifications that make antimicrobials ineffective. This process is driven by the acquisition of resistance traits through spontaneous mutations or, more frequently, through the acquisition of mobile genetic elements from other bacteria or the environment.

Bacteria, in particular, have evolved an arsenal of mechanisms to neutralize antimicrobials. These strategies include the enzymatic degradation of the molecule (as in the case of the aforementioned carbapenemases), the alteration of cell wall permeability to prevent molecule entry (e.g., membrane porins alteration or loss), the modification of molecular targets to inhibit binding, and the overexpression of efflux pumps to actively extrude toxic compounds from the cell [14]. All these AMR mechanisms evolved as competitive strategies for resource acquisition and ecological niche colonization in the environment, well before the introduction of modern antibiotics [15].

The “Golden age” (1940s) marked the discovery of new numerous synthetic antibiotic classes. However, their introduction in the market, their misuse and overuse in human medicine, veterinary, and agriculture have accelerated the diffusion of AMR. In fact, the employment of a new antibiotic class has been followed by the rapid emergence and dissemination of the relative resistance, creating an evolutionary 'arms race'. For instance, surveillance data has documented a relentless upward trend in resistance among Gram-negative bacteria over the last two decades, with resistance traits increasingly moving from hospital settings into the community [16].

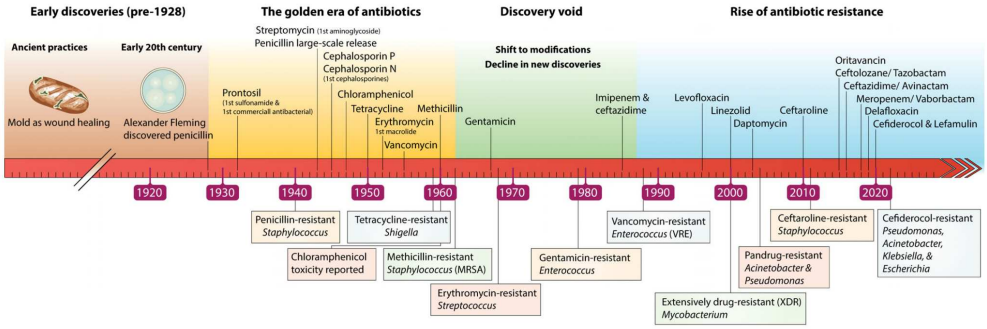


Figure 2 - Antibiotic discovery vs. resistance development. Timeline illustrating the discovery of major antibiotics and their resistances, as well as the appearance of important resistant microorganisms [17].

Concurrently with the rise of AMR levels, the pharmaceutical discovery process has effectively stalled since the 1980s, creating a significant "discovery void." This widening gap between the rapid pace of bacterial evolution and the stagnation in the availability of novel effective treatments has led to AMR being recognized as a global health emergency, pictured as a "silent pandemic", which is anything but silent [18]. This designation reflects the insidious nature of its spread through healthcare settings and communities, often driven by colonized patients who do not present clinical symptoms. However, the term "pandemic" is fully justified by the sheer scale of its impact: in 2021 alone, bacterial AMR was associated with nearly 4.71 million deaths worldwide [19], and it is forecasted to cause cumulatively 39 million deaths by 2050 [20].

In a concerted effort to mitigate this escalating threat, the ECDC established ambitious targets to be achieved by 2030 [21]. These goals focus on reducing total human antibiotic consumption by 20% and ensuring that at least 65% of prescribed antibiotics belong to the 'Access' group of the WHO AWaRe classification. Furthermore, specific reduction targets were set for the incidence of BSIs caused by key pathogens: 5% for carbapenem-resistant *K. pneumoniae* (CR-Kp), 10% for third-generation cephalosporin-resistant *E. coli*, and 15% for methicillin-resistant *S. aureus* (MRSA).

However, current epidemiological trends reveal a divergence from most of these goals. Since the 2019 baseline, the estimated incidence of BSIs caused by CR-Kp has surged by more than 60%, standing in contrast to the reduction goal. Similarly, infections caused by resistant *E. coli* have risen by

over 5%, drifting further away from the target. An exception is MRSA, which has exhibited a decreasing trend, aligning more closely with the established reduction trajectory [22]. Regarding stewardship, challenges remain critical: total antibiotic consumption has increased rather than decreased, and the proportion of 'Access' antibiotics used has remained stagnant at approximately 60%, falling short of the 65% benchmark [23].

This persistent difficulty in reversing resistance trends through traditional control measures underscores the urgency of innovation. Consequently the research focus has shifted towards non-conventional strategies beyond traditional small-molecule antibiotics, including phage therapy, the engineering of antimicrobial peptides (AMPs), and the application of advanced genomic tools like CRISPR-Cas9 [14].

1.4 Italian epidemiology of multidrug-resistant microorganisms

From an epidemiological perspective, the classification of bacteria as multidrug-resistant (MDR) refers to their non-susceptibility to at least three or more antimicrobial classes, a condition that severely limits therapeutic options. Within the European landscape, Italy bears a disproportionately heavy burden of these infections, consistently reporting resistance percentages significantly higher than the EU average.

This scenario is continuously monitored by the Italian Ministry of Health in coordination with the European Antimicrobial Resistance Surveillance Network (EARS-Net), which tracks the susceptibility profiles of key pathogens, such as the ESKAPE group plus *Streptococcus pneumoniae*. Recent data from the 2024 surveillance period [24] underscore a particularly alarming situation for Gram-negative bacilli. *A. baumannii* confirms its status as a critical threat in healthcare settings, limiting the efficacy of last-resort antibiotics; specifically, carbapenem resistance rates have exceeded 74% (74.3% in 2024; Figure 3).

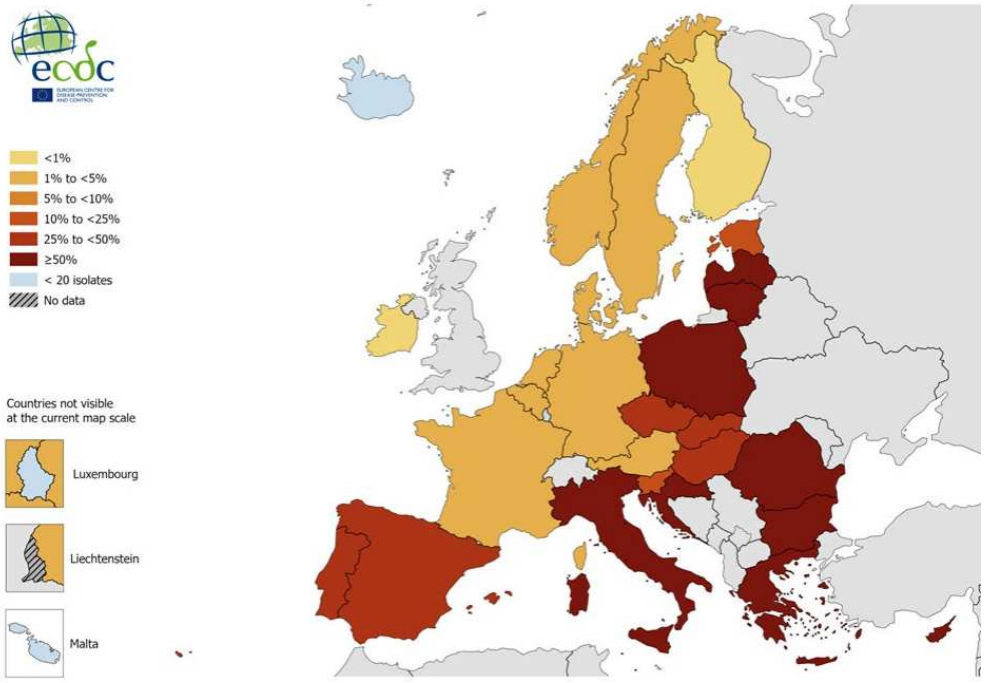


Figure 3 - Prevalence of *Acinetobacter species*. Percentage of invasive isolates with resistance to carbapenems (imipenem/meropenem) by country in EU/EEA, 2024 [24].

Similarly, *K. pneumoniae* continues to represent a major clinical challenge. Although the trend has shown a stabilization in recent years, resistance to carbapenems remained high, recorded at approximately 24% in Italy during 2024.

Regarding other Gram-negative bacteria, the 2024 landscape is heterogeneous. *P. aeruginosa* maintains a stable but clinically concerning resistance profile to carbapenems (around 13.8%). Conversely, *E. coli* presents a dual scenario: while it exhibits significant resistance to third-generation cephalosporins (approximately 27.7%) - a marker of extended-spectrum β-lactamase (ESBL) production - its resistance to carbapenems remains low (<1%), preserving these drugs as viable therapeutic options.

Finally, the epidemiology of Gram-positive bacteria is undergoing a significant shift. The historical threat of MRSA has shown a decrease over the last decade, stabilizing at around 25.7%. Conversely, *E. faecium* has exhibited a surge in vancomycin resistance, rising from 11.2% in 2015 to 34.9% in 2024.This steep increase is likely due to the host tropism, which typically colonises elderly and immunocompromised patients, who are more and more frequent inpatients in Italian hospitals.

1.5 Surveillance of HAIs: an italian experience

Historically, the surveillance of infectious diseases within healthcare settings has been based on standard epidemiological metrics, primarily the longitudinal tracking of case counts. This quantitative approach records the number of individuals diagnosed with a specific pathology across distinct time periods and geographic locations. By leveraging these data, epidemiologists can effectively identify outbreaks, monitor their trajectory, and uncover patterns that might suggest which groups are at higher risk. Thus, disease dynamics are revealed by mapping the temporal, spatial, and demographic distribution of the affected populations [25, 26].

Such monitoring remains paramount across all hospital wards, and particularly within ICUs. In these high-risk environments, the likelihood of acquiring an HAIs is significantly exacerbated - often exceeding that of general wards by 5 to 10 fold - due to both intrinsic patient vulnerabilities (e.g. immunosuppression state, increased age, and comorbidity) and extrinsic factors, such as the necessity for invasive mechanical ventilation or catheterization [27].

Effective surveillance necessitates the timely discrimination between active infection and colonization, defined as the persistence of pathogenic bacteria

in the host without an associated symptomatic immune response. This distinction is crucial when addressing pathogens characterized by high pathogenicity, rapid transmissibility, and high level of AMR.

At the IRCCS Fondazione Policlinico San Matteo in Pavia (Italy), infection prevention measures are implemented to mitigate the incidence of HAIs, paired with both passive and active surveillance protocols for sentinel microorganisms detection. Infection control measures rely on rigorous hygiene standards for medical personnel, specifically hand hygiene and the use of personal protective equipment, alongside contact precautions as well as the cohorting of colonized or infected patients.

Passive surveillance focuses on recording cases of sentinel microorganisms, for example carbapenem-resistant *Enterobacterales* (CRE), MRSA, vancomycin-resistant *Enterococcus spp.* and pandrug resistant *P. aeruginosa.* It is also extended to Tuberculous and Nontuberculous *Mycobacteria*, toxin-producers *Clostridioides difficile*, and *Candida auris.* However, this surveillance framework remains adaptive, encompassing broadly any microorganism that exhibits an incidence rate exceeding epidemiological baselines or presents a distinctive and concerning antimicrobial resistance profile.

Complementing this, active surveillance entails the systematic screening of high-risk patients, such as those admitted to ICUs or those presenting risk factors (e.g., prior hospitalization, prolonged antibiotic therapies, surgical interventions), primarily via rectal swabs to assess gastrointestinal carriage or nasal swabs for nasal colonization. Upon collection, specimens undergo microbiological processing to determine both species identification and resistance profiles. This diagnostic workflow employs culturing on selective media, followed by species identification via MALDI-TOF mass spectrometry or automated phenotypic systems (e.g., BD Phoenix), and AST performed via diffusion methods or rapid molecular assays (e.g., GeneXpert).

2. Genomic epidemiology

2.1 The evolution of genomics

Preceding the era of modern genomics with high throughput DNA sequencing, the framework for biological sequence study was initially laid in the 1960s within the domain of protein chemistry. Margaret Dayhoff was a central figure in this domain as she developed the *Atlas of Protein Sequence and Structure* [28] in 1965, thereby establishing the prototype for modern biological databases. Moreover, her formulation of PAM substitution matrices [29] provided the fundamental mathematical scaffolding for evolutionary inference. Therefore her pioneering work transformed the discipline of bioinformatics (i.e. the analysis, modeling and interpretation of complex biological data) from a descriptive practice into a rigorous analytical science [30].

In the 1970s the focus shifted from proteins to nucleic acids, driven by the revolution led by Frederick Sanger. His development of the chain-termination sequencing method (1977) marked a turning point, demonstrating its efficacy with the sequencing of the bacteriophage ΦX174 [31]. Representing the first complete DNA genome ever obtained (5.386 bp), this breakthrough necessitated a parallel evolution in computer science to manage the initial flow of nucleotide data.

Consequently, the late 1970s and 1980s were characterized by the advancing of DNA manipulation techniques and the development of computational infrastructures, empowering scientists to perform analytical tasks with an efficiency and speed previously unachievable by human effort alone [30]. In fact, determining an organism's genomic sequence provides significant advantages in data acquisition. This includes, for example, the opportunity to perform comparisons for evolutionary purposes and to identify the genes underlying adaptation to particular environments.

The 1990s marked the “genomic era” with the beginning of the Human Genome Project. Over a thirteen-year period (1990-2003), this global effort succeeded in sequencing all 3 billion base pairs (Gbp) of human DNA, identifying the approximately 20,000 to 25,000 genes that make up the human genome [32, 33]. Yet, the project accomplished much more than simply mapping the genetic code; it drove the creation of essential new bioinformatic software for genome assembly and pushed biology to integrate with the World Wide Web, making global, real-time data sharing a reality.

The completion of this era set the stage for a paradigm shift in data generation: the advent of Second-generation sequencing (also known as Next-Generation Sequencing (NGS) [34, 35]). Emerging shortly thereafter

(in 2005), this high-throughput technology made DNA sequencing widely accessible, due to a precipitous decrease in sequencing costs. NGS democratized access to genomic data by generating billions of short DNA fragments (known as "short reads"; typically 50-300 bp). Currently the market is dominated by the Illumina company and the workflow involves the fragmentation of the DNA, the attaching of adapters, the amplification and the sequencing of these fragments in a massively parallel process.

More recently, building on the success of Second-generation sequencing, Third-generation Sequencing (TGS) has further revolutionized the field by introducing long-read sequencing. Pioneered by companies such as Pacific Biosciences and Oxford Nanopore, these technologies introduced the capability to sequence native DNA strands without prior amplification. By generating reads that span from kilobases to megabases in length, this leap allows for the resolution of highly complex genomic architectures and significantly improves the detection of structural variations [36].

In the context of genomics applied to infectious diseases, the advent of TGS represents a transformative methodological shift. While short-read sequencing technologies are highly accurate at the nucleotide level, they are limited by their inability to bridge extensive repetitive elements such as insertion sequences (IS). Therefore, fragmented draft assemblies with numerous unlinked contigs are obtained. TGS circumvents this structural bottleneck by generating long reads allowing the reconstruction of complete, closed circular bacterial chromosomes. Furthermore, the ability to span long DNA fragments also facilitates the accurate assembly of mobile genetic elements, particularly plasmids which frequently serve as the primary vehicles for AMR and virulence genes [37].

2.2 Next-Generation Sequencing in clinical microbiology

Since the development of NGS, progress has been made regarding its implementation in medicine, being now an essential part of routine clinical diagnostics in many fields, including microbiology [38]. The integration of NGS into clinical microbiology marks a shift, moving the field from traditional culture-based diagnostics toward high-resolution molecular analysis. This technology is increasingly used not only to identify pathogens, but also to characterize them with a level of precision that impacts disease management and public health strategies [39]. Consequently, hospitals worldwide are progressively integrating NGS into their routine workflows [40, 41]. This trajectory is further validated by ECDC, which has established comprehensive guidelines for genomic-based surveillance of high-risk pathogens [42, 43].

At present day the dominant application of NGS in clinical microbiology is Whole Genome Sequencing (WGS: Figure 4A), a process where the organism of interest is sequenced and then its genome reconstructed. It typically relies on an isolate-based workflow [44]; in fact, for bacteria and fungi, WGS requires obtaining a pure culture prior to nucleic acid extraction, library preparation and sequencing to avoid contamination from the commensal flora naturally present in patient samples (e.g. oral or skin). However, this dependence on culture creates a bottleneck when the etiological agent is fastidious or unculturable.

To overcome this, clinical metagenomics has emerged as a complementary strategy, enabling the sequencing of potential pathogens directly from clinical specimens - a method that is already the standard for viral genome analysis (e.g. Respiratory Syncytial Virus, or SARS-CoV-2 [45]). Generally, this strategy employs two approaches.

The first approach, targeted or amplicon sequencing (Figure 4B), uses specific primers to amplify distinct genetic markers - such as the 16S or 18S ribosomal RNA (rRNA) gene for bacteria or the Internal Transcribed Spacer (ITS) for fungi - thereby enriching the sample for a desired taxonomic group. The targeted sequencing of the 16S rRNA gene constitutes also the basis of bacterial community taxonomic profiling in clinical samples [46]. This approach relies on the unique architecture of the 16S gene, which comprises highly conserved regions interspersed with nine hypervariable regions (V1-V9). These variable regions function as species-specific molecular fingerprints, accumulating mutations at a rate that allows for differentiation between distinct microbial lineages. However, while cost-effective, this method is inherently limited to taxonomic identification [47].

Conversely, the second approach, shotgun or untargeted metagenomics (Figure 4C), employs an unbiased strategy wherein all nucleic acids (DNA or RNA) present within a specimen are extracted and sequenced. By casting this "wide net", this approach facilitates the simultaneous identification of bacteria, fungi, viruses, and parasites within a single assay [48]. Then, sophisticated bioinformatics pipelines (such as DIAMOND + MEGAN [49] or DAMIAN [50]) are applied to disentangle the complex mixture, providing direct insight into the functional metabolic potential and the resistome of the microbial community. This approach has proven to be efficient in the diagnosis of infections, including meningitis or encephalitis [51], in bloodstream infections [52], and pneumonia [53].

Nevertheless, a significant challenge of this approach remains the sensitivity, which is dependent on the level of background. Tissues, for example, contain larger amounts of human host DNA compared to cell-free body fluids, resulting in a reduced number and proportion of microbial reads [54].

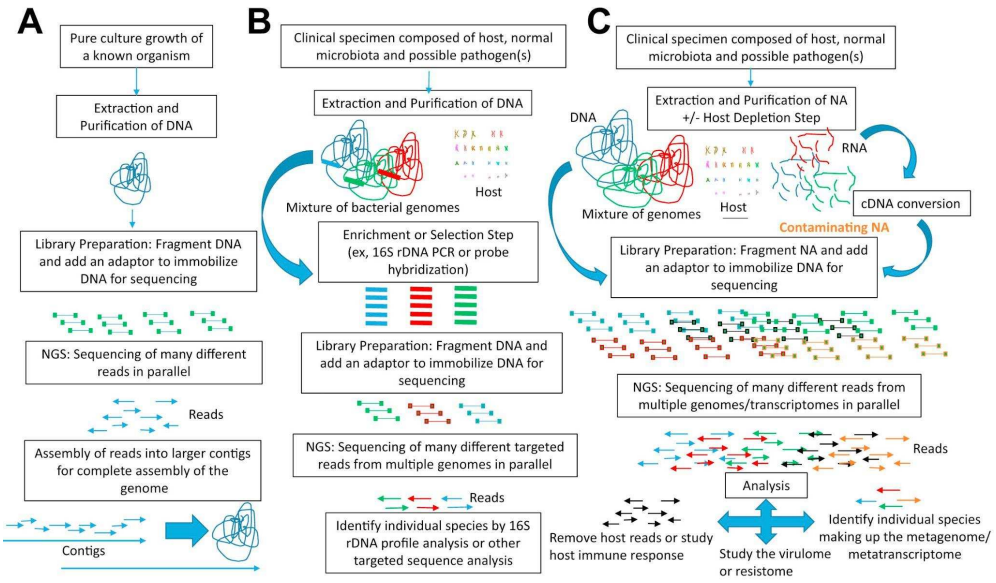


Figure 4 - Different applications of NGS in clinical microbiology. (A) Whole-genome sequencing of an organism from a pure culture. (B) Targeted (or amplicon) sequencing of conserved taxonomic markers: 16S/23S rRNA for bacteria profiling, 18S for parasite or ITS for fungi. (C) Untargeted (or shotgun) metagenomics from clinical specimens. The nucleic acid composition of the specimens includes host (black), microbiome and pathogen (blue, green, red), and contaminant (orange) sequences. The bioinformatic pipeline involves host depletion to isolate microbial reads for taxonomic profiling, resistome/virulome characterization, and transcriptome-based analysis [55].

2.3 Genomic-based surveillance

While traditional surveillance (i.e. chapter 1.5) remains indispensable for monitoring general trends, these aggregated data often lack the resolution necessary to elucidate complex transmission dynamics. Genomic epidemiology effectively addresses this analytical gap. By shifting the focus from statistical case counts to the direct inquiring of the pathogen genome coupled with patient hospitalization data, it provides more precise reconstruction of transmission pathways thus a deeper understanding of evolutionary dynamics, the detection of emerging variants and also the detection of the source.

Consequently, genomic-based surveillance has established itself as an instrument for infection prevention and control (IPC). Beyond its utility in guiding population-level interventions, this approach also advances personalized therapeutic strategies, thereby supporting a "precision public health" model. This paradigm enables the implementation of real-time,

data-driven containment measures specifically calibrated for the exigencies of healthcare settings [56].

In the modern bacterial diagnostic landscape, the WGS approach offers higher discriminatory power than conventional methods. Certainly, MALDI-TOF mass spectrometry is a gold standard for routine identification due to its rapidity and accuracy [57]. However, its diagnostic resolution is not absolute; species identification can be elusive when encountering specific anaerobic bacteria [58], and it struggles to differentiate between highly similar species. Similarly, historical molecular standards such as Pulsed-Field Gel Electrophoresis, Variable-Number Tandem Repeat profiling, and PCR-based Multilocus Sequence Typing (MLST) have long served the field due to their sensitivity and reproducibility.

2.3.1 Bacterial typing and transmission dynamics

Within the framework of genomic-based surveillance, WGS can be employed as a tool for isolates characterization, providing insights into the species, subspecies or subtypes and enabling the phylogenetic placement relative to other isolates.

Several methods have been developed for classifying and binning single isolates into bacterial clones: from *in silico* MLST that evaluates seven housekeeping genes, to the more comprehensive core genome MLST (cg-MLST), which considers more than 1500 genes on the core part of the genome [59]; the number of genes evaluated depends on the species. A further improvement in detailing the clones has been made by the introduction of Life Identification Numbers (LINs) code, a hierarchical nomenclature which assigns stable codes based on genomic similarity, thereby resolving the relatedness between strains with quantitative precision [60]. By systematically delineating clones and clonal groups (CGs), scientists can characterize their prevalence and ecological distribution across distinct environments. This approach has been widely used in understanding the population structure of different healthcare-associated pathogens, for instance of CR-Kp [61, 62] or CR-Ab [63].

In parallel with typing methods, phylogenetic analysis constitutes the core of genomic epidemiology investigation, since they are used to assess the evolutionary relationships among a set of organisms. The selection of specific genetic markers on which infer the phylogeny is guided by the resolution that is targeted: while conserved orthologous genes are well-suited for broad taxonomic comparisons, Single Nucleotide Polymorphism (SNP) analysis offers the granularity required to distinguish closely related isolates [64]. Accordingly, SNP-based methodologies have

become the gold standard for outbreak investigations in healthcare settings worldwide due to the bacterial high clonality within a short timeframe [65].

In addition, phylogenetic frameworks can be temporally calibrated to reconstruct the evolutionary history of a pathogen. By using Bayesian computational tools like BEAST, scientists can integrate sample collection dates to estimate exactly when strains diverged [66]. This approach is particularly valuable during outbreak investigations, as it maps the progression of the epidemic, allowing for the identification of its origin and the tracking of its spread over time. However, genomic data alone provides an incomplete picture when isolated from the host. In fact, for instance during outbreak investigation the genomic information must be integrated with epidemiological one, such as date of infection, hospitalisation ward, and possible patient's movement.

2.3.2 Antimicrobial resistance and virulence genes

Beyond simply identification and typing, the possibility to have fully sequenced genomes allows to predict phenotypes, such as the "resistome" (i.e., genomic traits related to antibiotic resistance) and the "virulome" (i.e., genomic traits related to virulence) which are of utmost importance in the healthcare setting.

Unlike culture based phenotypic methods which necessitate organism viability, or rapid molecular assays that are restricted to a limited panel of markers, the complete antimicrobial resistance or virulence profile of a pathogen can be defined. This predictive capacity is based on the comparison of the genome in analysis with a list of genes previously associated with antimicrobial resistance genes (ARGs) or virulence genes that are stored in curated reference databases (e.g., ResFinder [67], Comprehensive Antibiotic Resistance Database [68], and VirulenceFinder [69]). However, the accuracy of such *in silico* profiling depends strictly on the database used, which sometimes may not be fully updated. In fact, when resistance results from novel point mutations that occur independently of already recognized genes, it can elude initial bioinformatic detection. Nevertheless the subsequent identification is crucial for characterizing new resistance mechanisms and updating reference databases.

In this regard progresses have been achieved for specific high priority pathogens; for instance, Kleborate [70] represents a specialized and curated bioinformatics tool tailored for *Klebsiella spp.* and *E. coli*. Other tools such as AMRFinderPlus are now considered the gold standard for identifying ARGs, resistance-associated point mutations and also virulence genes [71].

Despite the potential power of AMR profiling, a significant challenge arises when trying to establish a perfect concordance between *in silico* predicted ARGs and phenotypic AST. In fact, the presence of a gene is not a guarantee of phenotypic resistance, as the gene expression is modulated by transcriptional regulation and environmental factors. For instance, a resistance determinant may appear genomically intact yet remain functionally "silent" due to mutation in upstream promoter regions. This is the case for the intrinsic $bla_{\text{OXA-51-like}}$ carbapenemase which is present in the chromosome of *A. baumannii* but confers resistance to carbapenems only when the insertion sequence ISAba1 locates upstream of the gene, acting as a strong promoter [72]. Moreover, phenotypic resistance is not only mediated by gene presence or absence but also it can emerge with mechanisms such as the overexpression of chromosomal efflux pumps or the downregulation of membrane porins (e.g. the loss of porins Omp35 or Omp36 in combination with ESBL production leads to carbapenem resistance in *K. pneumoniae [73]*). Furthermore, the degree of the resistance is not quantifiable with only *in silico* predictions. In fact it is not possible to predict whether an isolate with a particular genotype is likely to show high-level or low-level resistance (e.g., an MIC of >32/2 mg/L vs >16/2 mg/L for amoxicillin-clavulanate in *E. coli*), although the scientific community is making significant progress with the application of machine learning models [74].

2.3.3 Horizontal gene transfer: a focus on plasmids

While bacterial pathogens can evolve resistance and virulence traits through spontaneous mutations, these traits can also be acquired by Horizontal Gene Transfer (HGT). HGT is the process of exchanging genetic material via non-vertical transmission that can occur both between organisms within the same species or even across large evolutionary distances and ecological niches.

HGT is mediated by the three main mechanisms: transformation, transduction, and conjugation, plus “non canonical” mechanisms (e.g. extracellular membrane vesicles, tiny pilus-like nanotubes, or phage-like gene transfer agents). Transformation is defined by the uptake of DNA from the environment; while transduction is driven by bacteriophages acting as vectors; in this case the transferred genetic material can either lead to the formation of new phages or integrate into the host chromosome. Finally, conjugation necessitates direct physical contact between cells, typically mediated by a pilus, facilitating the exchange of plasmids and other mobile genetic elements from one cell to another. Once DNA successfully enters the cell, it may integrate into the chromosome via homologous recombination [75].

Plasmids are extrachromosomal self-replicative DNA elements that are independent from the chromosome. They can differ in size, copy number, GC content, replication mechanism, transmission mode, DNA topology (i.e., circular or linear), genetic cargo and host range. Typically small plasmids (1 to 10 kbp) are present in high copy numbers, while large plasmids (10 to 300 kbp) are present in few copies. Large plasmids are often conjugative and their size is positively correlated with higher numbers of ARGs. ARGs are compartmentalized in regions considered as “hotspots”, and flanked by Transposons (e.g. *Tn4401* carrying the bla_{KPC} carbapenemase gene) or IS that favour their mobilisation. Multiple ARGs in the same hotspot provide the bacterial host a selective advantage under antimicrobial pressure, conferring resistance to multiple drug classes. At the same time, harboring these large plasmids imposes a substantial metabolic cost, which typically translates to reduced replication rates and diminished competitive viability in environments without antibiotic selective pressure [76].

Due to a high density of repetitive sequences and transposable elements in plasmids, assembly algorithms used for short-read sequencing approach fail to resolve the complex structural plasmid architecture. Consequently, *in silico* classification has largely relied on categorizing the assemblies into incompatibility (Inc) groups by detecting specific replication control systems (replicons) directly within the fragmented contigs [77]. Nevertheless, this typing approach becomes particularly challenging when analyzing complex plasmids that harbor multiple replicons. In such scenarios, the implementation of long-read sequencing provides a decisive advantage: by circumventing assembly fragmentation, it enables the retrieval of the fully plasmid molecule.

Plasmid-mediated resistance has compromised the administration of widely used drugs against Gram-negative bacteria, playing a role in the diffusion of the resistance to penicillins, aminoglycosides, sulfonamides and last-resort antibiotics such as colistin or carbapenems. For instance, resistance to carbapenems in *K. pneumoniae* is mainly driven by the acquisition of genes encoding β-lactamases, including KPC-, OXA-48-, NDM-, VIM-, and IMP-type which are frequently associated with specific plasmids (e.g. KPC-2 and KPC-3 are linked to pKpQIL-like plasmids with $IncFII_k$ replicons [78]). In addition, plasmids have contributed substantially to the successful spread of MDR high-risk clones such as *K. pneumoniae* ST258 and ST512 part of the CG258 or the CG307, and *E. coli* ST131. A distinct plasmid epidemiology is observed between these lineages: whereas pKpQIL-like plasmids are a hallmark of CG258, they are rare in CG307, where replicon $IncFIB_k$ / $IncFIB_{pKP91}$, IncC and IncR are more frequent [79].

CHAPTER 2:

Aim of the research

Genomic epidemiology has increasingly become a cornerstone in healthcare settings, essential for understanding the complex transmission dynamics, in particular of multi-drug resistant (MDR) organisms associated with healthcare-associated infections (HAIs). By providing a high-resolution understanding of epidemiology, genomics directly improves patient management. Beyond tracking the circulation, it also offers the advantage of characterizing antimicrobial resistance determinants in depth, thereby serving as a tool to guide antimicrobial stewardship strategies.

In this context, the primary aim of this work was to implement and apply NGS-based approaches at the IRCCS Fondazione Policlinico San Matteo (HSM; Pavia, Italy) to enhance the surveillance of MDR bacteria across different temporal scales (Chapters 3 and 4) and to perform deep investigations of two outbreaks caused by MDR bacteria (Chapters 5 and 6). Furthermore, the work extends its scope to optimization of diagnostic pathways. This includes evaluating novel laboratory workflows for the management of bloodstream infections (BSIs) and investigating the gut microbiota in patients with Chronic Lymphocytic Leukemia (CLL) through clinical metagenomics approach (Chapters 7 and 8).

Chapter 3 reconstructs the epidemiological trajectory of carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) circulation over a nine-year period within the HSM. Given the CR-Kp prominent role in HSM, characterizing its population structure is essential to advance clinical management. Through a genomic approach, this project investigates the prevalence and persistence of circulating CR-Kp clones with their associated resistance determinants. Moreover, the analysis moves beyond traditional Inc group plasmid typing to explore the complete "plasmidome". This comprehensive characterization addresses the need to understand CR-Kp circulation, shedding light on both clonal dissemination and the horizontal transfer of genetic elements across different clones.

Chapter 4 extends the surveillance perspective, considering all carbapenem-resistant *Enterobacterales* (CRE) isolated during one year in HSM. This study addresses the necessity of monitoring the entire CRE landscape, extending beyond the most common species (such as CR-Kp) to include less frequent pathogens. The systematic sequencing of all

resistant isolates enabled the detection of rare clones and emerging antimicrobial resistance genes that might otherwise spread unnoticed.

Chapters 5 and 6 detail two high-resolution genomic outbreak investigations of CR-Kp and carbapenem-resistant *Acinetobacter baumannii* (CR-Ab). Specifically, Chapter 5 focuses on the emergence and the genomic characterization of a novel CR-Kp*,* identified as ST6668. First detected at HSM in February 2023, this clone subsequently triggered a healthcare-associated outbreak. This clone harboured the $bla_{\text{NDM-1}}$ gene, conferring extensive phenotypic resistance to carbapenems and marking a significant shift in local epidemiology from KPC- to NDM-mediated resistance. Conversely, Chapter 6 presents the investigation of an outbreak caused by CR-Ab of ST2 during the first COVID-19 pandemic wave (January-July 2020) in HSM. In this study, a hybrid sequencing approach (short- and long-read) was employed to fully characterize the transmission chain, the resistance and virulence profiles, as well as to track the plasmid dynamics.

Chapter 7 evaluates the impact of reorganizing the Clinical Microbiology Laboratory in HSM to optimize the management of BSIs. This organizational shift involved the implementation of a 24/7 workflow to ensure continuous sample processing and introduced a "fast-track" diagnostic pathway specifically for critically ill patients. This project demonstrates how this model enhances early sepsis management and supports antimicrobial stewardship through rapid diagnostics.

Finally, Chapter 8 is focused on the application of targeted metagenomics in clinical microbiology, exploring the gut microbial community in a cohort of patients with CLL. This project stratifies patients based on their response to targeted therapies (B-cell receptor inhibitors), investigating potential correlations between gut microbiota composition and therapeutic outcomes.

CHAPTER 3:

Longitudinal phylogenomic surveillance of carbapenem-resistant *Klebsiella pneumoniae* reveals clonal replacement and segregated plasmid dissemination

Greta Petazzoni[1,2], Stefano Gaiarsa[2], Marta Corbella[2], Cristina Merla[2], Irene Mileto[2], Angela Kuka[2], Alba Muzzi[3], Gherard Batisti Biffignandi[2], Davide Sassera[4,5], Vincenzo Di Pilato[6,7], Gian Maria Rossolini[8,9], Fausto Baldanti[1,2], Patrizia Cambieri[2]

1. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia. Pavia, Italy
2. Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo. Pavia, Italy
3. Medical Direction, Fondazione IRCCS Policlinico San Matteo. Pavia, Italy
4. Department of Biology and Biotechnology, University of Pavia. Pavia, Italy
5. Fondazione IRCCS Policlinico San Matteo. Pavia, Italy
6. Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genoa, Genoa, Italy
7. Microbiology Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy
8. Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy
9. Microbiology and Virology Unit, Careggi University Hospital, Florence, Italy

The manuscript is **under review by Genome Medicine,** December 2025.

My contribution to this work encompassed leading and executing all the bioinformatic analyses, from the beginning to the conclusion. Furthermore, I was responsible for structuring and drafting the manuscript.

Abstract

Carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) represents a critical healthcare-associated pathogen frequently linked to severe clinical outcomes. In the Italian epidemiological landscape, the dissemination of CR-Kp is primarily sustained by high-risk lineages, specifically CG258, sequence type (ST) 101, and ST307, which harbor plasmid-mediated carbapenemase genes such as bla_{KPC}, bla_{NDM}, and bla_{VIM}. Plasmids also carry other antimicrobial resistance genes and play a key role in the persistence and spread of specific lineages. This study investigates the population structure, transmission patterns, and plasmid diversity of CR-Kp within a hospital setting over a nine-year timeframe.

A total of 556 unique CR-Kp isolates were retrospectively selected between 2015 and 2023 from hospitalized adults at Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy), ensuring balanced representation across time periods and clinical units. Whole-genome sequencing (WGS) was performed using Illumina, followed by assembly and characterization with Kleborate. Phylogenetic relationships were inferred using core-genome SNP analysis, identifying epidemic clusters. To explore plasmid content, draft genomes were screened via BLAST against a custom plasmid database. Then, the unitig presence/absence matrix was dimensionally reduced using UMAP and clustered via HDBSCAN to identify plasmidome clusters.

The most prevalent clones retrieved were ST307 (32.4%) and ST512 (25.5%), followed by ST258 (11.5%) and ST6668 (9.0%). The dominant resistance mechanisms were bla_{KPC-3} (307/556) and bla_{KPC-2} (130/556), with sporadic detection of bla_{NDM}, bla_{VIM}, and rare bla_{KPC} variants. Phylogenetic analysis identified 24 epidemic clusters (ECs), encompassing 73.7% of all CR-Kp. ST307 and ST512 exhibited persistent and evolving ECs, often associated with specific plasmidome clusters (PCs). Plasmidome mapping revealed 23 distinct PCs, which were strongly segregated by clonal groups (CG258, CG307, CG147), suggesting limited horizontal plasmid transfer across CGs.

Our study reveals long-term coexistence and diversification of high-risk CR-Kp clones in a single hospital over nine years and clonal replacement. The expansion and persistence of clones such as ST307 and ST512 are likely driven by the plasticity in their plasmidic content. Integrated genomic and plasmidome surveillance is crucial to detect emerging threats, understand transmission dynamics, and guide infection control strategies.

Introduction

Klebsiella pneumoniae is a major opportunistic pathogen frequently causing hospital-acquired infections and associated with high mortality rates (1). In Europe its prevalence is among the highest worldwide (2) and many outbreaks have occurred in Southern states (Italy, Greece, Spain) driven by a small number of high-risk clones (3). The groundbreaker lineage is clonal group 258 (CG258), mainly comprising sequence types (ST) 258 and its single-locus variant ST512, which is currently endemic in Italy. Other lineages frequently circulate in Italy, such as STs 11 and 101, while others are emerging to date, as STs 307 and 147 (4-6).

These healthcare-associated clones exhibit resistance to a broad spectrum of antibiotics, including carbapenems, which are considered as last resort agents. Carbapenem resistance may be multifactorial but is primarily driven by the acquisition of mobile genetic elements such as plasmids and transposons, which harbor genes encoding carbapenemases (7). In carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp), plasmids frequently carry multiple antimicrobial resistance (AMR) genes - averaging approximately nine per plasmid (8) - as well as systems for plasmid transfer and stability, facilitating horizontal gene transfer within the *Enterobacterales*. Consequently, plasmid content plays a critical role in shaping the evolutionary success of specific lineages and constitutes a major driving force in the dissemination of resistance genes among high-risk clones.
Notably, carbapenemase genes such as *bla*$_{KPC}$, *bla*$_{NDM}$, *bla*$_{VIM}$, and *bla*$_{OXA-48}$ are routinely identified in successful nosocomial clones, often revealing a strong association between specific lineages and distinct resistance determinants. In the Italian context, *bla*$_{KPC}$ remains the most prevalent resistance gene among CR-Kp isolates, with the *bla*$_{KPC-2}$ and *bla*$_{KPC-3}$ variants being predominant and strongly linked to CG258 and ST307 (4,5). Concurrently, bla$_{NDM-1}$ is emerging as a result of the recent local expansion of additional lineages, including ST147 (in Tuscany) (6) and the novel ST6668 (in the Pavia area) (9).

Given the importance and complexity of this issue, hospital-based genomic surveillance is essential for monitoring the emergence and the dissemination of CR-Kp clones. Indeed WGS enables high-resolution phylogenetic analysis to identify transmission chains and epidemic lineages, together with AMR and virulence genes and give an insight into plasmid investigation (10,11). In this retrospective study, covering the period from January 2015 to September 2023 at Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy), we applied genomic and phylogenomic approaches to characterise clonal diversity, map transmission clusters, and investigate the dissemination of plasmid content within the CR-Kp population.

Materials and Methods

Study design, isolate selection, and sequencing

A retrospective analysis was conducted on carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) isolates recovered at the Fondazione IRCCS Policlinico San Matteo (HSM) in Pavia, Italy, spanning the period from January 1, 2015, to September 30, 2023. Clinical and microbiological data were retrieved from the hospital database. Surveillance specimens (rectal swabs or faeces) were plated onto selective chromogenic agar CHROMID® Carba medium (BioMérieux, Marcy-l'Etoile, France) and incubated; while other clinical specimens (blood, urine, respiratory, and other non-surveillance specimens) were processed following internal procedures. Isolates from all specimens were identified with Matrix Assisted Laser Desorption/Ionisation Time-Of-Flight Mass Spectrometry (MALDI-TOF MS; Bruker Daltonics, Bremen, Germany) equipped with Bruker Biotyper 3.1 software. Antibiotic susceptibility was tested using the BD Phoenix NMIC-474 panel loaded on a BD Phoenix M50 automated instrument (Becton Dickinson, Franklin Lakes, NJ, USA) and Sensititre EUMDRXXF (Thermo Scientific Diagnostics B.V., The Netherlands). Minimum inhibitory concentration (MIC) values were interpreted according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints; versions of breakpoints were chosen according to the year of sampling. Carbapenemase production was confirmed using the NG-Test® CARBA-5 lateral flow immunoassay (NG Biotech laboratories) or Xpert Carba-R (Cepheid, Italy).

For each patient (age >=18) with at least one phenotypically confirmed CR-Kp (Meropenem MIC > 0.125 mg/L) and hospital admission, only the first CR-Kp was included. To achieve a balanced temporal and spatial sampling, we divided each calendar year into four trimesters and grouped hospital wards into six areas: Intensive care units (ICUs), General Surgery Units, Surgical Specialty Units, General Medicine Units, Specialist Medicine Units, and the Emergency Room (E.R.). We thus obtained a total of 210 bins. Subsequently, we established a predefined sampling threshold, aiming to get the genomic data of a minimum of 20% of the total available clinical isolates within each respective bin. The dataset integrated genomes sequenced in previous HSM surveillance initiatives, specifically incorporating 22 isolates from an ICU outbreak (12) (BioProject PRJEB32609); 82 from the SpARK One-Health campaign (13) (BioProject PRJEB27342); 38 from an ST6668 outbreak (9) (BioProject PRJNA1019273); 47 from the MUSE project (BioProject PRJNA1019247); and 208 from a one-year genomic surveillance program (14) (BioProject PRJNA1266098).

In addition, genomic data were generated for other 159 CR-Kp isolates. These supplementary isolates were drawn from the repository of each targeted bin using a computer-generated simple random sampling algorithm, carefully ensuring the exclusion of any clinical samples from the same patient (BioProject PRJNA1284063).

The final cohort comprised 556 unique CR-Kp isolates (Table S1). For the 159 newly sequenced genomes, genomic DNA was isolated using the QIAsymphony DSP Virus/Pathogen Midi Kit (QIAGEN, Germany). Then, libraries were prepared with the Nextera XT protocol and sequenced on an Illumina MiSeq platform (Illumina, San Diego, CA, USA). De novo assembly of raw reads was performed using Shovill (github.com/tseemann/shovill, v1.0.4) with default parameters. Assemblies were retained only if they met quality criteria of ≥ 30× average coverage and an N50 of at least 20 kb.

Genomic characterization, phylogeny, and statistical analysis

Each genome was characterized with Kleborate (15) for multilocus sequence typing (MLST), identification of antibiotic resistance genes, virulence factors, and capsule locus typing. Any putative novel carbapenemase variants flagged by Kleborate were manually validated by BLAST against the NCBI nucleotide database. PlasmidFinder was also used to retrieve Incompatibility groups (16).

To infer phylogenetic relationships among CR-kp having the same Sequence Type (ST, n ≥ 2), we applied the P-DOR pipeline (17). This approach facilitated the contextualization of local isolates within the global diversity by incorporating closely related genomes from the BV-BRC database (www.bv-brc.org). We then identified phylogenetic clusters which were defined as monophyletic clades with bootstrap support > 80/100. Clades having more than five genomes within were classified as epidemic clusters (ECs), those with three to four genomes as low-transmission clusters, and singletons were interpreted as unique introductions.

Continuous variables were reported as mean with standard deviation. Associations between STs and resistance determinants identified with Kleborate, as well as between hospital areas and the ECs, were evaluated using chi-square tests. A p-value < 0.05 was considered statistically significant.

Plasmid landscape analysis

To characterize the plasmidic content in our dataset, we first retrieved all publicly available *K. pneumoniae* complete genomes and extracted their plasmid sequences, resulting in 8565 plasmidic reference sequences (NCBI RefSeq database, accessed in March 2025). These reference sequences were then used to screen our draft genomes to identify putative plasmid-derived sequences using BLAST, obtaining the plasmidome.
The screening criteria for BLAST were set to ensure high-confidence matches: query length > 1,000 bp, sequence identity ≥ 95%, alignment length ≥ 1,000 bp, and alignment coverage ≥ 90% of the query length. This strategy allowed us to focus on substantial and highly similar plasmid content, minimizing the inclusion of fragmented or ambiguous matches.
Then, unitig presence-abscence matrix of plasmid-derived sequences in each plasmidome were built using uniting-caller (https://github.com/bacpop/unitig-caller) and threshold proposed by Arredondo-Alonso and colleagues in mge-cluster (18). Then umap was applied to reduce dimensionality (with Jaccard distance; 15 neighbours) using the uwot package in R (19). The HDBSCAN algorithm (with minPts = 10) (20) was used on the coordinates identified by the umap to identify plasmidome clusters (PCs).
Finally, functional annotation of PCs was carried out using Bakta (21), while ResFinder (22) was employed to screen for antimicrobial resistance genes. The comparison of the distribution of genes number in each PC, as well as the distribution of the proportion of antimicrobial resistance genes on total number of genes in each EC was performed using the Wilcoxon-Mann-Whitney test.

Results

Carbapenem-resistant *K. pneumoniae* population

Between January 1, 2015 and September 30, 2023 a total of 4027 carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) were isolated at Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy). We filtered for the first CR-Kp per patient and hospitalization event, including at least 20% of the total number of CR-Kp per hospital area and trimester. We selected 556 of them. We screened previous HSM projects (n=397) to retrieve already available genomes and further sequences were produced (n=159) to cover undersampled periods and areas (see Materials and Methods). All CR-Kp had an average genome size of 5.6 ± 0.1 Mbp (Figure S1, Table S2).

Selected CR-Kp were collected from 541 hospitalized adult patients, with a mean age of 64 years and predominantly composed by male (358/541, 66.17%). Clinical specimens were mainly rectal swabs (391/556, 70.32%) and urine (67/556, 12.05%), with other 98 samples obtained from infection-related materials (Table S3).

Nearly half of all CR-Kp (265/556, 47.7%) were isolated in Intensive Care Units (ICUs); the remaining were recovered from Specialist Medicine Units (118/556, 21.22%), Surgical Specialty Units (62/556, 11.15%), General Surgery Units (61/556, 10.97%), General Medicine Units (36/556, 6.47%), and the Emergency Room (E.R.; 14/556, 2.52%) (Figure 1 a,b and Table S3).

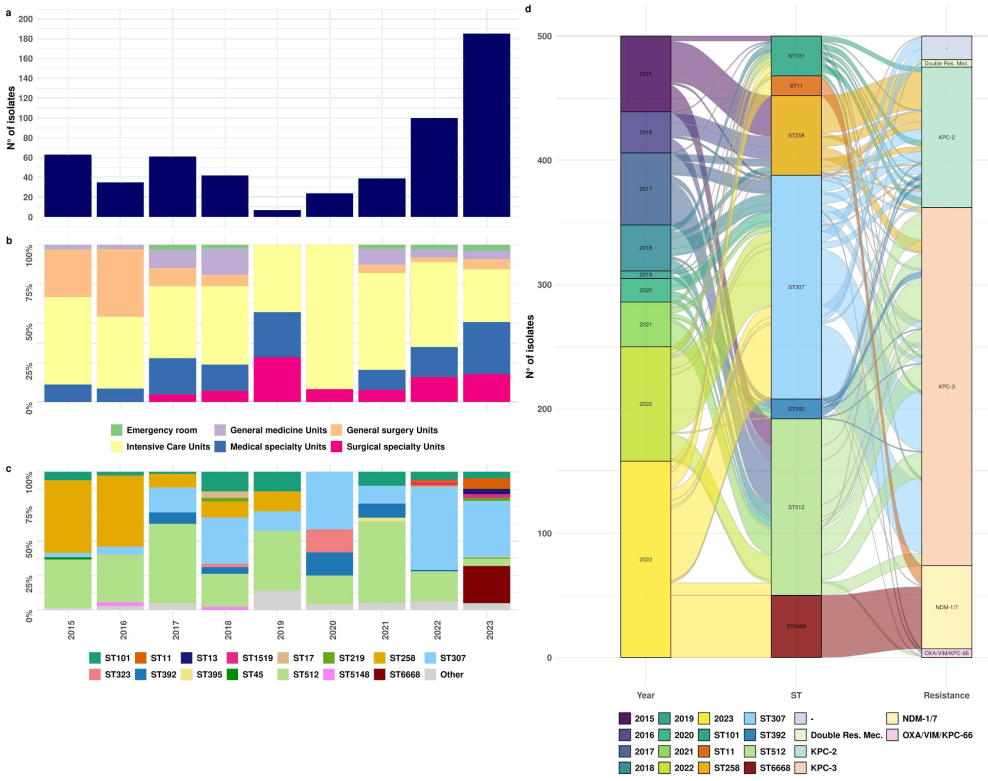


Figure 1 - Characterisation of 556 CR-Kp in the dataset. (a) Number of isolates collected per year. (b) Distribution of isolates in different hospital areas: Emergency Room, Intensive Care Units, General Medicine Units, Specialist Medicine Units, General Surgery Units, and Surgical Specialty Units. (c) Distribution of sequence types (STs) per year with at least 2 occurrences. (d) Temporal distribution of principal STs with associated resistance mechanisms; the dash symbol ("-") indicates the absence of a detected carbapenem resistance mechanism.

Genomic epidemiology of CR-Kp

The CR-kp genomes were characterized using Kleborate in order to define Sequence types (STs), resistance and virulence determinants. Genomic characterization revealed that ST307 (180/556, 32.37%) and ST512 (142/556, 25.54%) were the most prevalent clones in HSM, followed by the ST258 (64/556, 11.54%) and ST6668 (50/556, 8.99%) (Figure 1c). Additional lineages included ST101 (32/556, 5.76%), ST11 (16/556, 2.88%), ST392 (16/556, 2.88%), and different low-frequency STs (Table S4 and S5).

From 2015 until its decline, ST258 had two overlapping populations: one associated with *bla*$_{KPC-2}$ and capsule type KL106 (wzi29; n=50) and the second with *bla*$_{KPC-3}$ and capsule type KL107 (wzi154; n=14), historically

defined as ST258-clade I and ST258-Clade II (23,24). Its single-locus variant (SLV) ST512, was first found in 2015 and remained present until 2023 in HSM with strains mainly carrying bla_{KPC-3} and KL107 (n=138). Other two CR-Kp ST512 co-harboured bla_{KPC-3} plus bla_{VIM-1} while the remaining two had a sporadic KPC-66 carbapenemase (Figure 1c and d, Table S4).

ST307 steadily increased in frequency after 2018, co-circulating with ST512 and eventually predominating in 2022–2023 (Figure 1c). It was mainly associated with bla_{KPC-3} (120/180, $p<0.05$) but occurrences of bla_{KPC-2} (33/180) were found. Moreover, four CR-Kp ST307 co-harboured bla_{KPC-3} + bla_{NDM-1}, two bla_{OXA-48}, one bla_{NDM-7}, and one bla_{KPC-66} (Figure 1d).

Moreover, ST6668, a SLV of ST147, was first detected in March 2023 (Figure 1c). It drove a large HSM outbreak during 2023 (9) and it was significantly associated with bla_{NDM-1} spreading. In addition, ST101 was mainly associated with bla_{KPC-2} (22/32, $p<0.05$), while others had bla_{KPC-3} (n=8) coupled with genes for Aerobactin (n=6/8 *iuc1*), and bla_{VIM-1} (n=2). All ST101 also harboured genes encoding Yersinabactin (*ybt9*) on ICEKp3 (Table S4).

In summary, carbapenemase profiling showed a dominance of KPC-type enzymes in CR-Kp (444/556 79.86%), primarily encoded by bla_{KPC-3} (307/556) and bla_{KPC-2} (130/556) genes (Figure S2). Sporadic variants of bla_{KPC-} included bla_{KPC-66} (n=3), $bla_{KPC-239}$ (n=3, all ST1519 in which conferred resistance to ceftazidime/avibactam), and $bla_{KPC-231}$ (n=1). Class B carbapenemases were also found, mainly of bla_{NDM-1} (n=66), and bla_{VIM-1} (n=8). Finally, twenty-eight CR-Kp were phenotypically resistant to Meropenem but lacked detectable carbapenemase genes (Table S4).

Phylogeny and transmission dynamics

In order to understand the relationship of CR-Kp within each clone (at least 2 CR-Kp per ST) and elucidate the possible transmission chain, core SNPs-based phylogenetic analyses were performed. We therefore excluded the 24 STs that we found once as they were considered as single-entry events. Within the phylogeny, we identified clusters as well-supported (bootstrap > 80) monophyletic clades. Eighty-nine such clusters emerged: 37 (41.57%) represented single introductions with no onward spread, 28 (31.46%) were “low-transmission” clusters (2-4 genomes), and 24 (26.97%) were called epidemic clusters (ECs; ≥ 5 genomes). Overall, 73.7% of CR-Kp belonged to ECs, indicating high clonal expansion within our hospital.

The overall ECs diversity varies across the study period in HSM. Within the CG258 lineage (ST258/ST512), 12 distinct ECs were identified (Figure 2). In ST258, ECs A to D expanded between 2015 and 2018, before disappearing or remaining as sporadic occurrences (e.g. EC B); in detail EC A was part of the ST258-Clade II while other ECs were part of ST258-Clade I. In ST512, ECs were more large in number; some were sporadic with limited diffusion in time (ECs ST512_A, B, C, and D) with partial overlapping while two persisted through 2023 (ECs ST512_E and G).

ST307 ECs were absent in 2015–2016, first appearing in 2017 (ECs ST307_A, and C) and persisting thereafter. In 2023, ST307 diversification increased with novel clusters D and E, while cluster A disappeared. Moreover, ECs of ST392 (part of CG147) were firstly detected in 2017 and had an intermittent presence/absence pattern until 2022, suggesting the existence of a reservoir within HSM. Additionally, other less persistent ECs were identified from 2022 in our hospital, such as those of ST11, ST101, ST13, and ST6668.

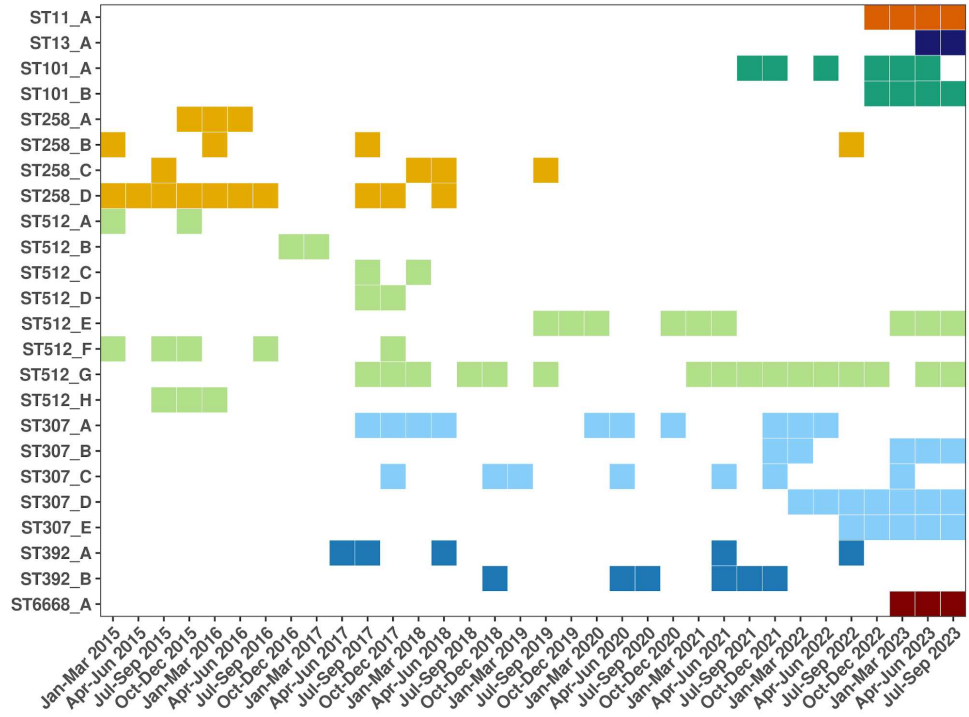


Figure 2 - Temporal dynamics per trimester of major epidemic clusters (ECs) within HSM. Sequence types (STs) are color-coded, and letters indicate ECs within each ST, defined as clades in the core-SNP phylogeny supported by high bootstrap (> 80) and comprising more than five CR-Kp.

ECs exhibited significant areas-specific patterns of occurrence (Figure S3). The ST258 lineage was strongly enriched in ICUs ($p < 0.05$), particularly within clusters A and B, whereas EC ST258_D predominated in the General Surgery Units ($p < 0.05$).

The ST512 disseminated broadly throughout the HSM (e.g. ST512_E, F and G), but ST512_H showed a significant overrepresentation in Surgical Specialty Units ($p < 0.05$). Similarly, ECs of ST307 (ST307_A, B, and C) were detected across a wide range of areas and, notably, EC ST307_E recurred with high frequency in both Surgical Specialty Units and ICUs ($p < 0.05$). The ST11_A was most commonly isolated from General Medicine Units and the E.R., while ST13_A cases were almost exclusively associated with Surgical Specialty Units ($p < 0.05$). In contrast, ST6668 isolates were predominantly associated with Specialist Medicine Units ($p < 0.05$).

Plasmidome landscape

To characterize the plasmid landscape within our collection, we searched for plasmid-derived sequences in draft genomes using BLAST against a curated database of 8,565 complete plasmids, thereby defining plasmidomes. Unitigs originating from plasmidomes were extracted and converted into a presence/absence matrix, which was subjected to dimensional reduction and hierarchical clustering. This analysis delineated 23 plasmidome clusters (PCs), encompassing 91.91% (n = 511) of CR-Kp plasmidomes, while the remaining 45 were unassigned (Table S6).

The large majority of PCs occupied three well-separated regions in the ordination space (Figure 3a). In the left region, PCs 2–9 comprised plasmidomes of CR-Kp predominantly belonging to CG258. ST512 (121/199 CR-Kp plasmidomes) spanned all PCs in this region, whereas ST258 (63/199) was concentrated in PCs 2-5 and ST1519 (6/199) in PCs 6-7; seven additional sporadic STs were also detected. In the top-right region, plasmidomes from CG307 (PCs 11-17) were located, consisting exclusively of ST307 plus three sporadic STs (ST140, ST3227, ST348). In the bottom-right region, plasmidic clusters corresponded to CG147 (PCs 21-23), with single-locus variants of ST147. PC 21 encompassed only ST392, whereas PCs 22-23 comprised ST6668. Additional PCs were identified: PC 1 (ST307 plus ST45 and ST253), PC 10 (ST11 only), PCs 18-19 (mixed STs), and PC 20 (predominantly ST101). Overall, the plasmidome map segregated into three macro-areas mirroring CG divisions, suggesting limited plasmid exchange between STs over the years in our hospital.

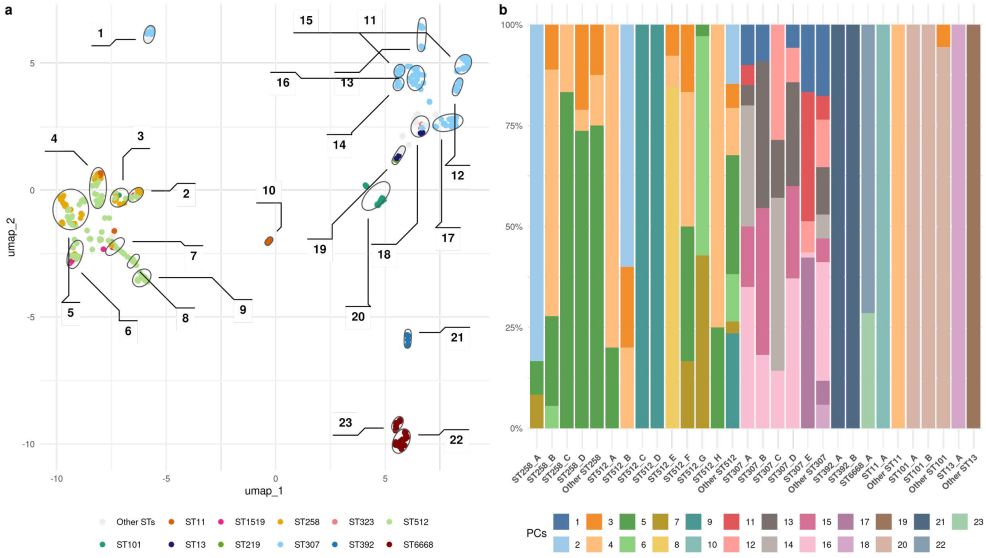


Figure 3 - Plasmidome clusters (PC). (a) Two dimensional umap representation of plasmid-derived sequences colored according to chromosome ST. Ellipses represent Plasmidome clusters (PCs) with their respective numbers. (b) Barplot indicating the proportion of CR-Kp in each epidemic cluster (EC) that had a specific PC, colored by PCs.

PCs were subsequently correlated with ECs defined by the core-genome phylogeny to investigate potential plasmidome transmission among CR-Kp. Extensive PC diversity within STs was observed (Figure 3b), particularly in lineages that have circulated for years in our hospital (e.g. ST258, ST512, and ST307). In ST258, six PCs were distributed across four ECs (ST258_A-D: Figure 4a and S4a). CR-Kp of ST258_A (part of the ST258-clade II) harboured almost exclusively PC 2 (n=10/12), while those in ST258_B harboured PC 4 in 2015 while PC 5 late in time (2023). Also, PC 5 was the most persistent and abundant PC in CR-Kp of ST258_C and D (n=19/25) and in other ST258 not belonging to any EC (Figure S4a).

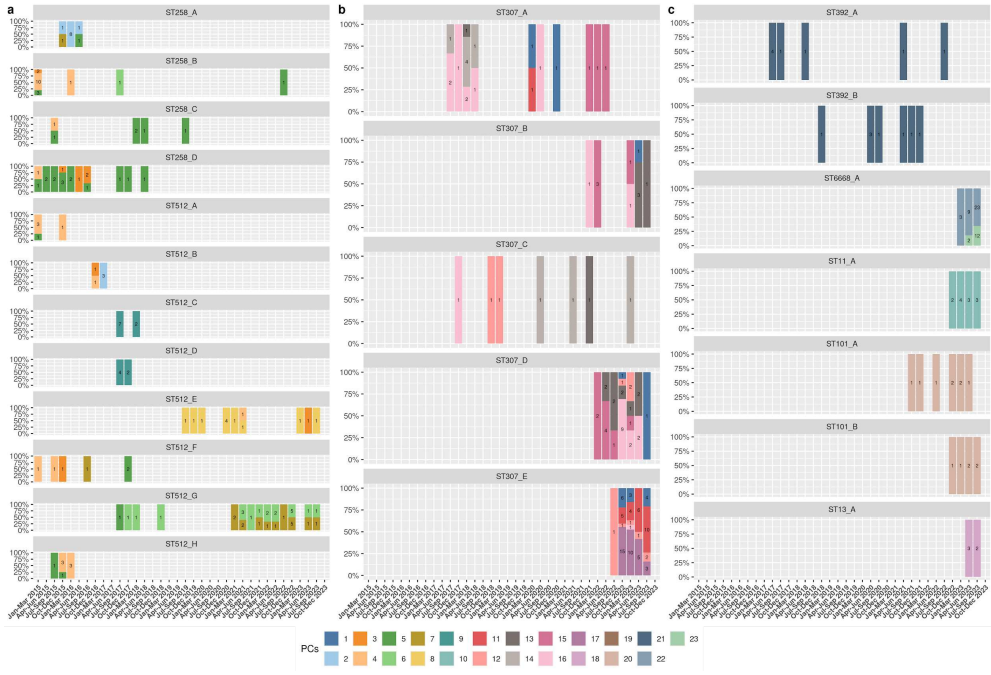


Figure 4 - Plasmidome cluster (PC) distribution across epidemic clusters (ECs). Proportion and total number of PCs within ECs during each trimester, stratified by sequence type (ST). Data are shown for ST258 and ST512 (a), ST307 (b), and other STs (c). PCs were defined by analyzing the full plasmid content in each carbapenem-resistant Klebsiella pneumoniae genome, while ECs were defined as highly supported clades in the core-SNP phylogeny (bootstrap > 80; ≥5 genomes).

PC composition in ST512 ECs (eight PCs in 8 ECs) was uniform over time, with few or single PCs per EC (e.g. PC 9 in EC ST512_C and D, PCs 4-5 in EC ST512_H; Figure 3b and 4a) and some of them were in common with PCs of ST258 (i.e. mainly PC 4 and 5). In long-standing ECs of ST512 (ST512_E and ST512_G), most CR-Kp carried PCs almost exclusively to their EC (PC 8 and PCs 6-7, respectively). For PCs 6 and 7, the comparison with the phylogeny exhibited a random pattern (Figure S4b), and no large difference was detected in the number of genes or resistance genes within plasmidomes, suggesting the absence of events of gain/loss of entire large plasmids (Figures S5a and S5b).

In ST307, eight PCs (11-17) were distributed across five ECs throughout the study period (Figure 3b and 4b), indicating marked plasmid plasticity. Some PCs were temporally restricted yet shared across ECs, such as PC 15 (October 2021-March 2023 in ECs ST307_A, B, D) and PC 16. In EC ST307_E, a novel PC (PC 17) emerged in late 2022 and was present in most members (n = 33/79) of the clade, with no evidence of vertical

transmission (Figure S6). Despite core-genome relatedness, CR-Kp in this clade also carried PCs 1 (n = 13), 11 (n = 25), and 12 (n = 6); PC1 and 11 had a number of genes similar to PC 17 (median [Q1-Q3] number of genes per PC: 313 [304-317] in PC 1; 312 [309-315] in PC 11 vs. 315 [312-318] in PC 17; Figure S5a), while PC 12 had a lower number of genes (median [Q1-Q3]: 244 [241-247], Wilcoxon test on PC 12 vs. PC 17 $p < 0.05$), suggesting substantial plasmidic content loss. No significant difference was also observed in resistance-gene counts in these PCs (Figure S5b), beside PC 12 ($p < 0.05$). Overall, the proportion of resistance genes relative to total genes was higher in ECs of ST307 than in ECs of ST258 and ST512 (Wilcoxon test $p < 0.05$; Figure S5c and d).

Conversely, in ST392 ECs (ST392_A, B) a single PC was identified, not shared with the other CG147 member observed (i.e. ST6668_A; Figure 3b and 4c). Within ST6668 two distinct PCs (22, 23) were detected (Figures 3b and 4c) despite CR-Kp high clonality (Figure S7). Substantial plasmidic content difference between the two PC was supported by a higher gene count in PC 22 compared with PC 23 (median [Q1-Q3] number of genes: 362 [358-476] vs. 159 [100-214]; Figure S5a) and distinct incompatibility-group profiles (Figure S8). In detail, CR-Kp in PC 23 had lost IncHI1B and IncFIB-Mar (n=13/14) compared with PC 22, while IncFIB(pQil) was present in all CR-Kp of PC 23 but absent in six of 35 plasmidomes of PC 22. Loss of plasmidic content in PC 23 was evident but appeared random, with no clear association to specific phylogenetic subclades (Figure S7).

Discussion

Carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) has emerged as a significant global public health threat due to its rapid dissemination and ability to evade last-resort β-lactams. Comprehensive genomic characterization of resistant isolates is essential for infection-control efforts and informs clinical decision-making by guiding antibiotic stewardship and targeted interventions (25). Here we inferred the epidemiology of 556 CR-Kp using the genome data obtained from clinical samples over a nine year period. Most (70.32%) CR-Kp were collected from routine rectal swab screening according to established protocols in high-risk patients for detecting asymptomatic carriage of carbapenem-resistant *Enterobacteriaceae*.

Genomic surveillance revealed that ST258 initially dominated among CR-Kp clones; however by 2018 it was supplanted by its SLV ST512, mirroring pan-European patterns of CG258 evolution (4,5). ST307 appeared later in HSM and, similarly to ST512, replaced ST258 over the years becoming, at present day, endemic in Italy (26). From 2021 onwards, we observed a marked increase in clonal diversity, with the emergence of novel STs and subclusters of established lineages. This result could have been influenced by the availability of a higher amount of genomes from an earlier HSM study (14), where a deeper sampling was carried out, which allowed us to also detect less prevalent clones.

The majority (73.7%) of CR-Kp belonged to epidemic clusters (ECs, monophyletic clades with ≥ 5 genomes), indicating sustained intra-hospital transmission. This hypothesis is further supported by the observation that singleton STs constituted merely 6.65% of all CR-Kp in our population. The identified ECs were both widespread or area-specific, reflecting in some cases limited cross-area spread. Certain ECs persisted for years and were widespread in HSM, such as ST512_E and G, ST307_A and C, and ST392_A and B.

Interestingly, despite the global impact of the COVID-19 pandemic between 2020 and 2022, which significantly shaped hospital management and led to high patient turnover, we observed a notable persistence of specific CR-Kp ECs following the initial pandemic waves. This clonal stability, regardless of the increased flux of patients and potential opportunities for external introduction or replacement, suggests the presence of stable reservoirs within the local healthcare network. These reservoirs may include long-term care facilities, chronically colonized patients, or environmental niches that facilitated the reintroduction and maintenance of ECs over time (27).

Plasmidome profiling largely recapitulated the phylogeny, with plasmidomes from CR-Kp belonging to the same ST clustering closely together in the

two-dimensional ordination space. This observation suggested limited horizontal transfer of plasmidic content between different CGs in HSM (28). When ECs were correlated with plasmidome clusters (PCs) we observed opposite trends. In general, ECs of CG258 showed a lower PC variability than those of ST307, both in terms of numerosity and stability over time. This stability could have fostered the expansion of the CG worldwide in the past decades (4,5,29). Notably, we observed that EC E and G of ST512, which persisted over the years until the end of the study period, carried almost exclusive PCs (PC8, PC6, and PC7) which could have had a positive role in the epidemiological success of the two ECs.

The ECs of ST307 appeared to have high diversity in their PCs and a larger proportion of antimicrobial resistance genes relative to the total number of plasmidic genes. This enhanced genomic plasticity may result from both small rearrangements in gene order - likely driven by mobile elements such as insertion sequences and transposons that may have caused PC exchange, as well as from more substantial events of gene loss and/or acquisition, which were further reflected in the differential presence or absence of replicons. Moreover, the greater plasticity in all ST307 ECs may have facilitated its dissemination and persistence, despite high fitness cost (30), and may have contributed to the displacement of CG258. A notable exception to this observation is ST512_G, which shows a high PC diversity (as for ST307 ECs; Figure S4b) and persisted in HSM throughout the whole study period. A high plasticity level has been already observed by multiple groups as a possible reason for the ST307 worldwide success (31,32). Conversely, in emerging STs, PC diversity is less evident (e.g. ST11 and ST13), except in ST6668 which shows a greater plasmidic plasticity beside high core-genome relatedness, as already observed (9).

Even though CR-Kp are typical heal-care associated pathogens, also non carbapenem resistant CR-Kp can cause infections in hospitalized patients. A limitation of our approach is the exclusive focus on CR-Kp, omitting the susceptible *K. pneumoniae* population. Because antimicrobial resistance frequently emerges via horizontal gene transfer, resistant and susceptible lineages are often intermixed rather than forming distinct monophyletic clades in the species phylogeny. Consequently, analyzing only resistant isolates can hide broader evolutionary dynamics, making it difficult to distinguish true clonal expansion from the independent dissemination of plasmidic content across the susceptible population. Future studies incorporating both phenotypes would provide more context, enabling the precise mapping of independent resistance acquisition events, the identification of susceptible ancestors, and a more robust evaluation of lineage-specific evolutionary success.

Conclusion

The genomic investigation of CR-Kp conducted on a large cohort of isolates collected over a nine-year period revealed the coexistence of multiple high-risk clones and the clonal replacement during years. Beside the limited transfer of plasmidic content between clones, the plasmidic content diversification within certain epidemic clusters appears to have favoured their persistence and spreading.

Acknowledgment

This study was funded by NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project PE00000007, INF-ACT), Ricerca Corrente 1004-rcr2022-43 (Italian Ministry of Health), and Ricerca Finalizzata 2016 - grant no. RF-2016-02364584 (Italian Ministry of Health).

Data availability

Genome sequence data analyzed in this paper are listed in Table S1, with individual accession identifiers for reads and assemblies and deposited in the National Center for Biotechnology Information.

Supplementary material

Supplementary material can be found in the Chapter 3 folder at this link: https://drive.google.com/drive/folders/1PWOT8cuMTOMLwZhP2MymwC_UzTy0iyOZ.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105). The work described herein is a study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

References

1. Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, et al. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. Lancet Infect Dis. 2013 Sep;13(9):785–96.
2. European Centre for Disease Prevention and Control. Point prevalence survey of healthcareassociated infections and antimicrobial use in European acute care hospitals. Stockholm: ECDC; 2024.
3. David S, Reuter S, Harris SR, Glasner C, Feltwell T, Argimon S, et al. Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. Nat Microbiol. 2019 Nov;4(11):1919–29.
4. Budia-Silva M, Kostyanev T, Ayala-Montaño S, Bravo-Ferrer Acosta J, Garcia-Castillo M, Cantón R, et al. International and regional spread of carbapenem-resistant Klebsiella pneumoniae in Europe. Nat Commun. 2024 Jun 14;15(1):5092.
5. Di Pilato V, Errico G, Monaco M, Giani T, Del Grosso M, Antonelli A, et al. The changing epidemiology of carbapenemase-producing Klebsiella pneumoniae in Italy: toward polyclonal evolution with emergence of high-risk lineages. J Antimicrob Chemother. 2021 Jan 19;76(2):355–61.
6. Di Pilato V, Henrici De Angelis L, Aiezza N, Baccani I, Niccolai C, Parisio EM, et al. Resistome and virulome accretion in an NDM-1-producing ST147 sublineage of Klebsiella pneumoniae associated with an outbreak in Tuscany, Italy: a genotypic and phenotypic characterisation. Lancet Microbe. 2022 Mar;3(3):e224–34.
7. Navon-Venezia S, Kondratyeva K, Carattoli A. Klebsiella pneumoniae: a major worldwide source and shuttle for antibiotic resistance. FEMS Microbiol Rev. 2017 May 1;41(3):252–75.
8. Hendrickx APA, Landman F, de Haan A, Borst D, Witteveen S, van Santen-Verheuvel MG, et al. Plasmid diversity among genetically related Klebsiella pneumoniae blaKPC-2 and blaKPC-3 isolates collected in the Dutch national surveillance. Sci Rep. 2020 Oct 8;10(1):16778.
9. Mileto I, Petazzoni G, Corbella M, Gaiarsa S, Merla C, Kuka A, et al. Rapid spread of a novel NDM-producing clone of Klebsiella pneumoniae CC147, Northern Italy, February to August 2023. Euro Surveill. 2023 Oct;28(42).
10. Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. Microbiology (Reading, Engl). 2018 Oct;164(10):1213–9.
11. Teo JQ-M, Tang CY, Tan SH, Chang HY, Ong SM, Lee SJ-Y, et al. Genomic Surveillance of Carbapenem-Resistant Klebsiella

pneumoniae from a Major Public Health Hospital in Singapore. Microbiol Spectr. 2022 Oct 26;10(5):e0095722.
12. Ferrari C, Corbella M, Gaiarsa S, Comandatore F, Scaltriti E, Bandi C, et al. Multiple Klebsiella pneumoniae KPC Clones Contribute to an Extended Hospital Outbreak. Front Microbiol. 2019 Nov 29;10:2767.
13. Thorpe HA, Booton R, Kallonen T, Gibbon MJ, Couto N, Passet V, et al. A large-scale genomic snapshot of Klebsiella spp. isolates in Northern Italy reveals limited transmission between clinical and non-clinical settings. Nat Microbiol. 2022 Dec;7(12):2054–67.
14. Petazzoni, G, Gaiarsa, S, Mileto, I, Kuka, A, Corbella, M, Merla, C et al. Empowering active surveillance of Carbapenem-resistant Enterobacterales: Insights from one year of genomic surveillance. Journal of infection and public health. J Infect Public Health. 2026; 19(4):103187
15. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun. 2021 Jul 7;12(1):4188.
16. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014 Jul;58(7):3895–903.
17. Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics. 2023 Sep 2;39(9).
18. Arredondo-Alonso S, Gladstone RA, Pöntinen AK, Gama JA, Schürch AC, Lanza VF, et al. Mge-cluster: a reference-free approach for typing bacterial plasmids. NAR Genom Bioinform. 2023 Sep;5(3):lqad066.
19. Melville J. uwot: The Uniform Manifold Approximation and Projection (UMAP) Method for Dimensionality Reduction. 2024;
20. Hahsler M, Piekenbrock M, Doran D. dbscan : Fast Density-Based Clustering with *R*. J Stat Softw. 2019;91(1).
21. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021 Nov;7(11).
22. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom. 2022 Jan;8(1).
23. Deleo FR, Chen L, Porcella SF, Martens CA, Kobayashi SD, Porter AR, et al. Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 Klebsiella pneumoniae. Proc Natl Acad Sci USA. 2014 Apr 1;111(13):4988–93.
24. Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D,

et al. Genomic Analysis of the Emergence and Rapid Global Dissemination of the Clonal Group 258 Klebsiella pneumoniae Pandemic. PLoS ONE. 2015 Jul 21;10(7):e0133727.
25. Hernández-García M, Castillo-Polo JA, Cordero DG, Pérez-Viso B, García-Castillo M, Saez de la Fuente J, et al. Impact of Ceftazidime-Avibactam Treatment in the Emergence of Novel KPC Variants in the ST307-Klebsiella pneumoniae High-Risk Clone and Consequences for Their Routine Detection. J Clin Microbiol. 2022 Mar 16;60(3):e0224521.
26. Peirano G, Chen L, Kreiswirth BN, Pitout JDD. Emerging Antimicrobial-Resistant High-Risk Klebsiella pneumoniae Clones ST307 and ST147. Antimicrob Agents Chemother. 2020 Sep 21;64(10).
27. Neidhöfer C, Sib E, Neuenhoff M, Schwengers O, Dummin T, Buechler C, et al. Hospital sanitary facilities on wards with high antibiotic exposure play an important role in maintaining a reservoir of resistant pathogens, even over many years. Antimicrob Resist Infect Control. 2023 Apr 15;12(1):33.
28. Comandatore F, Sassera D, Bayliss SC, Scaltriti E, Gaiarsa S, Cao X, et al. Gene Composition as a Potential Barrier to Large Recombinations in the Bacterial Pathogen Klebsiella pneumoniae. Genome Biol Evol. 2019 Nov 1;11(11):3240–51.
29. Villa L, Feudi C, Fortini D, Brisse S, Passet V, Bonura C, et al. Diversity, virulence, and antimicrobial resistance of the KPC-producing Klebsiella pneumoniae ST307 clone. Microb Genom. 2017 Apr 26;3(4):e000110.
30. Brockhurst MA, Harrison E. Ecological and evolutionary solutions to the plasmid paradox. Trends Microbiol. 2022 Jun;30(6):534–43.
31. Wyres KL, Wick RR, Judd LM, Froumine R, Tokolyi A, Gorrie CL, et al. Distinct evolutionary dynamics of horizontal gene transfer in drug resistant and virulent clones of *Klebsiella pneumoniae*. PLoS Genet. 2019 Apr 15;15(4):e1008114.
32. Bellinzona G, Biffignandi GB, Brilli M, Baldanti F, Sassera D, Gaiarsa S. How to measure bacterial genome plasticity? A novel index helps gather insights on pathogens. Microb Genom. 2025 Aug;11(8).

CHAPTER 4:

Empowering active surveillance of carbapenem-resistant *Enterobacterales*: insights from one year of genomic surveillance

Greta Petazzoni [1,2], Stefano Gaiarsa[1], Irene Mileto[1] , Angela Kuka[1], Marta Corbella[1], Cristina Merla[1], Antonio Piralla[1], Marina Ramus[1], Aurora Piazza[2,3], Vittoria Mattioni Marchetti[2], Roberta Migliavacca[2], Alba Muzzi[3], Patrizia Cambieri[1], Fausto Baldanti[1,2]

1. Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo. Pavia, Italy
2. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia. Pavia, Italy
3. Fondazione IRCCS Policlinico San Matteo. Pavia, Italy

The manuscript has been **published in Journal of Infection and Public Health**, March 2026, Volume 19, Issue 4.
https://doi.org/10.1016/j.jiph.2026.103187

I was responsible for all *in silico* analyses presented in this manuscript, overseeing the study from preliminary data inspection to the final results. In addition to the analytical work, I also conceptualized the manuscript's structure and drafted it.

Abstract

Healthcare-associated infections (HAIs) in Europe affect millions of patients annually, driving antimicrobial resistance and imposing significant economic burdens. Carbapenem-resistant *Enterobacterales* (CRE) are particularly concerning due to their ability to spread, making genomic surveillance essential for tracking their transmission.

From October 2022 to September 2023, we sequenced and analysed the genome of 340 CRE obtained from patients admitted to Fondazione IRCCS Policlinico San Matteo (Italy). Sequencing was performed using a short-read approach (Illumina technology) and genomes were *in silico* characterized to identify Sequence Types (STs), antimicrobial resistance and virulence determinants, and plasmid content. Moreover, a core SNP-based phylogeny was performed to identify potential epidemic clusters (ECs; ≥ 5 strains in each clade) within the same ST.

Among the 340 CRE isolated from surveillance (N=234) and other clinical specimens (N=106), 87.7% belonged to the *Klebsiella* genus, primarily *K. pneumoniae* (CR-Kp, N=284). *Escherichia coli* (CR-Ec) accounted for 8.5% (N=29); while the remaining species were *Enterobacter spp.* (N=13) and *Citrobacter spp.* (N=9). bla_{KPC} and bla_{NDM} gene variants were detected in 55.9% and 25.6% of CRE, respectively. The most prevalent CR-Kp clones were ST307, ST6668, ST258, and ST512, whereas ST131 was dominant among CR-Ec. Large epidemic clusters (ECs; ≥ 5 strains in each clade) were mainly found in CR-Kp associated with ST307 and ST6668, with 80.63% of CR-Kp grouped into nine ECs. Conversely, other species were observed as sporadic cases or within transmission clusters with limited spread.

Our comprehensive analysis of CRE circulation at HSM revealed that their dissemination was largely driven by highly successful clones, shedding light on several of them.

Introduction

Healthcare-associated infections (HAIs) are among the most common complications in healthcare, affecting one in ten inpatients, with a higher prevalence in low-/middle-income countries and high-risk patients (1). Europe records 4.5 million HAIs annually, leading to over 90,000 deaths and 2.5 million disability-adjusted life years. HAIs are strongly linked to antimicrobial resistance, with 71% of resistant infections occurring in healthcare settings (2). Globally, 136 million antibiotic-resistant HAIs occur every year (3), imposing significant economic burdens of billion euros.

Carbapenem-resistant *Enterobacterales* (CRE) pose a major global threat due to their ability to spread undetected as colonisers of the human gut. Italy is among the countries in south Europe with the highest percentage of invasive isolates resistant to carbapenem; indeed a recent report of European Centre for Disease Prevention and Control (ECDC) pointed out that 24% of invasive *Klebsiella pneumoniae* isolates in Italy were resistant to carbapenems. In this context, Italy has witnessed a rapid spread of resistance primarily driven by Klebsiella pneumoniae producing KPC carbapenemases (4).

Surveillance plays a crucial role in limiting the transmission of HAI-causing bacteria and thereby identifying epidemic events. Many healthcare settings worldwide have incorporated genomic characterisation into surveillance programs, improving strain typing and outbreak tracing. Genomics has been estimated to reduce HAIs by 18-37% and to prevent thousands of related deaths (5,6).

At Fondazione IRCCS Policlinico San Matteo (HSM), a 900-bed teaching hospital in Pavia (Italy), active surveillance for CRE is conducted on patients admitted to Intensive Care Units (ICUs) or those in other wards presenting risk factors (e.g., prior hospitalization, prolonged antibiotic therapies, surgical intervention). Surveillance specimens (rectal swabs or faeces) are collected on admission and weekly, identifying patients with CRE colonisation. As a part of infection prevention and control measures for multi-drug resistant (MDR) diffusion, CRE isolated from other specimens are also monitored.

From October 2022 to September 2023, whole-genome sequencing of CRE was integrated into traditional surveillance to provide a comprehensive overview of bacterial circulation. Here, the results of one year of genomic surveillance of CRE at HSM are presented.

Materials and methods

Criteria of inclusion

The study was conducted from 1st October 2022 to 30th September 2023, at Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy). Surveillance specimens (rectal swabs or faeces) were plated onto selective chromogenic agar CHROMID® Carba medium (BioMérieux, Marcy-l'Etoile, France). Plates were incubated at 37°C for 16 h into WASPLab® (Copan Italia). Other clinical specimens (blood, respiratory, urine samples and non-rectal swabs) were processed following internal procedures (7). Isolates from all specimens were identified with Matrix Assisted Laser Desorption/Ionisation Time-Of-Flight Mass Spectrometry (MALDI-TOF MS; Bruker Daltonics, Bremen, Germany) equipped with Bruker Biotyper 3.1 software. Antibiotic susceptibility was tested using the BD Phoenix NMIC-474 panel loaded on a BD Phoenix M50 automated instrument (Becton Dickinson, Franklin Lakes, NJ, USA) and Sensititre EUMDRXXF (Thermo Scientific Diagnostics B.V., The Netherlands). Minimum inhibitory concentration (MIC) values were interpreted according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints (version 12.0 (8) and 13.0 (9)) and carbapenemase production confirmed using the NG-Test® CARBA-5 lateral flow immunoassay (NG Biotech laboratories) or Xpert Carba-R (Cepheid, Italy). CRE colonisation was assessed via rectal swabs or faeces collection.

For this study, we considered all *Enterobacterales* isolates with MIC of Meropenem > 0.125 mg/L obtained processed during diagnostic routine. The Meropenem MIC cut-off was chosen since above this value, carbapenemase production should be suspected (8,9). For each patient (aged ≥ 18 years), only the first isolate for each bacterial species and antibiotic susceptibility profile were included.

DNA extraction and sequencing

The genomic DNA was extracted using QIAsymphony DSP Virus/Pathogen Midi Kit (QIAGEN, Germany) and sequenced using Illumina MiSeq (San Diego, California, U.S.) after Nextera XT library preparation.

Short reads were assembled using Shovill (github.com/tseemann/shovill). Forty-four genomes were retrieved from previous work (BioProject: PRJNA1019273). Kleborate (10,11) was used to assign Sequence Types (ST) and predict resistance and virulence genes for *Klebsiella spp.*.

Whether a new carbapenem gene variant was predicted with Kleborate, we manually checked using BLAST.

Regarding other CRE, the Pasteur multilocus sequence typing (MLST) scheme and Achtman scheme for *Escherichia coli* were used to type strains whenever feasible using mlst (https://github.com/tseemann/mlst). For all CRE with no assigned MLST, sequences were submitted to the PubMLST database (https://pubmlst.org/). Finally, ResFinder (12) and VirulenceFinder (13) were used to search for resistance and virulence genes, while SerotypeFinder was used to assess the *E. coli* serotypes (14). PlasmidFinder (15) was performed on the dataset to identify plasmidic replicon sequences.

The P-DOR pipeline (16) was used to assess the phylogenetic relationships among strains of the same ST ($n \geq 2$) based on core-SNP analysis. This approach allowed the contextualization of local isolates within the global diversity, by incorporating closely related genomes from the BV-BRC database (www.bv-brc.org). We then identified transmission clusters which were defined as monophyletic clades with bootstrap support > 80/100.

To assess the association between species and wards, as well as ST and wards, the chi-square test was performed, followed by post hoc analysis with Bonferroni correction to adjust for multiple comparisons. The significance threshold was set at 0.05.

In vitro evaluation of carbapenem resistance retention capability

We evaluated the capability of retaining the carbapenem resistance factors in two clades of *K. pneumoniae*, selecting a subset of strains for each clade. These strains were initially cultured on Columbia agar supplemented with 5% sheep blood (COS; bioMérieux, France) and incubated at 37°C for 48 hours. Subsequently, colonies were transferred every 24 hours from COS to two different media: fresh COS and CHROMID® Carba (bioMérieux, France), to evaluate their carbapenemase activity each time. Each transfer cycle was defined as a "round". In cases in which no growth was observed on CHROMID® Carba in a round, a phenotypic test was performed using the KPC/MBL and OXA-48 Confirm Kit (Rosco, Denmark) on Mueller-Hinton II agar (Liofilchem, Italy) to confirm the absence of the resistance gene. The number of rounds required to observe the loss was then recorded. The Mann-Whitney U test was performed to compare the number of rounds in the two clades and between cases and controls.

Results

From October 2022 to September 2023, 4069 Enterobacterales were isolated from samples processed during routine diagnostic (patients ≥ 18 years old) in Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy). Of these, 48.5% (N=1975) were beta-lactamases susceptible, 21.7% (N=881) produced extended-spectrum beta-lactamases, and 23.34% (N=950) were carbapenem-resistant (CRE).

Dataset characterisation

We selected the first CRE per patient, obtaining 340 CRE from 294 patients (Figure 1a and Table S1). Two hundred thirty-four (68.5%) CRE were isolated from surveillance specimens, 57 from urine, 20 from bloodstream, 17 from respiratory tract, while the remaining from other specimens. One hundred ninety-five patients were male (66.3%) with a median age of 67 years (Q1-Q3: 55-76), further details are present in Table S1 and Supplementary results 1. The majority of CRE were obtained from patients in Intensive Care Units (ICUs; N=130; 38.2%), Infectious diseases (N=28; 8.2%), and Pneumology (N=22; 6.5%).

Forty-seven out of 294 patients developed invasive infections, leading to 46 bacteremia and two meningitis cases (one patient developed both). The median time from colonisation to infection was 11 days (range: 0-60). WGS was performed on the blood culture isolates of 20 patients; eight of them were colonised by the same pathogen (seven *K. pneumoniae* KPC, one VIM), while 12 were not colonised. The isolates obtained from the 26 remaining patients were not sequenced since they did not fulfil the inclusion criteria (first isolate per patient). Eighteen of 26 patients were colonised by the same bacteremia-causing pathogen (15 *K. pneumoniae* KPC and three *K. pneumoniae* NDM), while one had different CRE (colonised by *K. pneumoniae* KPC+NDM and infected by *K. pneumoniae* NDM). Eight patients carried strains with multiple resistance mechanisms: five had *K. pneumoniae* KPC+NDM, one had *K. pneumoniae* KPC+VIM, one had *E. coli* OXA+NDM, and the last had *E.coli* NDM+VIM.

The 340 CRE isolates comprised 11 species: the majority were within the *Klebsiella* genus (N=298; 87.7%), including *K. pneumoniae* (N=284), *K. quasipneumoniae* (N=3), *K. variicola* (N=2), *K. aerogenes* (N=7), and *K. oxytoca complex* (N=2). Other less frequently isolated CRE were *Escherichia coli* (N=29), *Enterobacter spp.* (eight *E. hormaechei*, one *E. asburiae*), and *Citrobacter spp*. (two *C. freundii*, one *C. amalonaticus*, and one *C. telavivensis*). *K. pneumoniae* was significantly associated with ICUs ($p < 0.05$), while other species were evenly distributed (Figure 1a and b and Table S1).

High-quality genomes of the 340 CRE were obtained ((mean sequencing depth: 55x; range: 12x-88x; Table S2). Genomic characterization revealed that 190 (55.9%) CRE carried bla_{KPC} variants, while 87 (25.6%) bla_{NDM} variants, and 22 (6.5%) bla_{VIM-1}. Six were found to harbour bla_{OXA-48}-like genes. The co-occurrence of multiple carbapenem resistance genes was observed in six cases of K. pneumoniae and one case of E. coli (Figure 1c). In 32 CRE, no acquired gene associated with carbapenem resistance was found (Table S3).

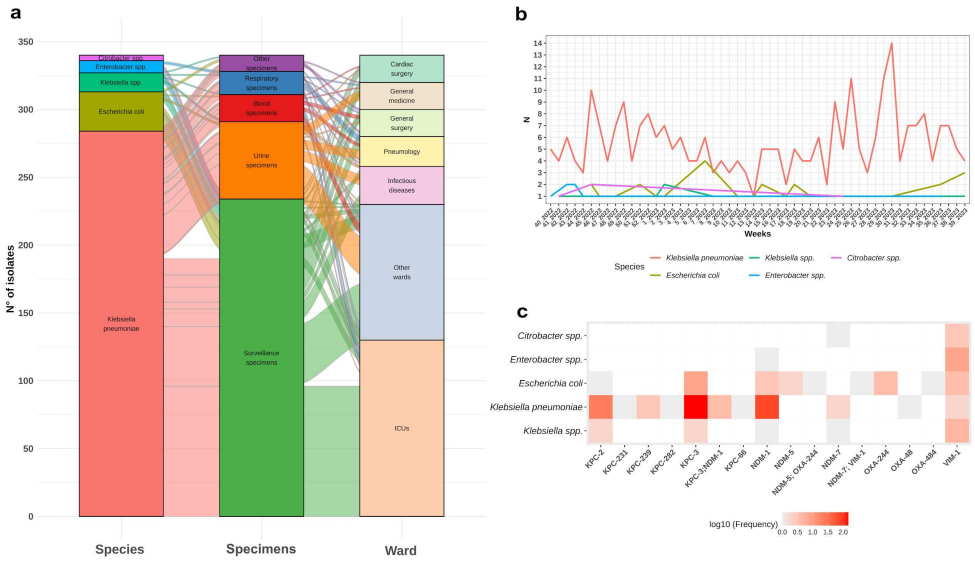


Figure 1 - Epidemiological and genomic landscape of the 340 carbapenem-resistant *Enterobacterales* (CRE) isolates. (**a**) Alluvial diagram shows the dataset characterization, illustrating the relationship between species, specimens, and wards (see Table S1 for further details). (**b**) Weekly incidence of CRE from October 1, 2022 (Week 40 of 2022), to September 31, 2023 (Week 39 of 2023), stratified by species. (**c**) Heatmap illustrates the distribution of resistance mechanisms across CRE species; color intensity reflects the log_{10}-transformed frequency.

Transmission clusters

For species with at least two genomes assigned to the same Sequence Type (ST), we performed core SNP-based phylogenetic analysis, integrating BV-BRC database genomes for evolutionary contextualisation. We identified four types of transmission: epidemic clusters (ECs; monophyletic clades with at least five CRE); moderate transmission clusters (3-4 study CRE in the monophyletic clade), one-to-one transmission clusters (2 study CRE), and singletons, representing independent hospital entries with no further spread.

Klebsiella pneumoniae

Regarding carbapenem-resistant *K. pneumoniae* (CR-Kp), clones detected were ST307 (N=130), ST6668 (N=56), ST512 (N=26) with four ST258, ST101 (N=16), ST11 (N=15), and also other less prevalent clones (Table S4). In general, the diffusion of CR-Kp was particularly prominent: 19 transmission clusters were identified. Of these, nine consisted of ECs, five were moderate transmission clusters while the other five were one-to-one transmission clusters. Additionally, 28 CR-Kp were singletons. The nine CR-Kp ECs identified comprised most of the CR-Kp strains (N= 229, 80.63 %). See Figure 2.

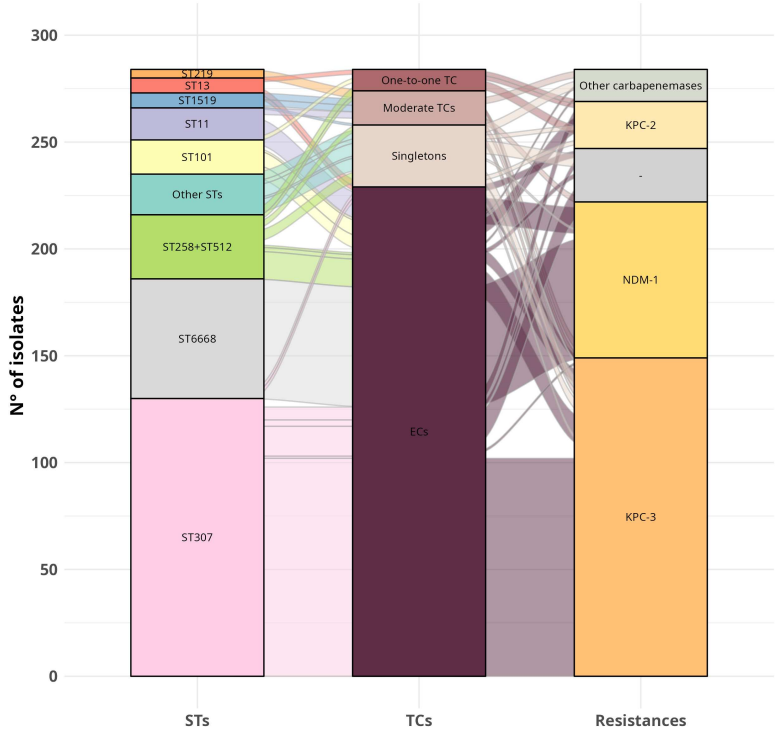


Figure 2 - Distribution of carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) strains. The alluvional plot depicts majoritarian STs (n ≥ 2 within the same ST); the category "Other STs" encompass STs detected only once. The second stratum indicates the distribution of STs across transmission clusters (TC), identified by inspecting the core-SNP phylogeny: Epidemic clusters (ECs; n ≥ 5 CR-Kp within the monophyletic clade), Moderate TCs (n = 3 to 4 CR-Kp within the monophyletic clade), and One-to-one TC (n=2 in the monophyletic clade). Resistance profiles correspond to the associated carbapenemase enzymes. The category "Other carbapenemases" includes less prevalent enzymes (see Tables S3 and S4 for details).

K. pneumoniae ST307

The ST307 core SNP phylogeny identified three ECs with strong support (bootstrap > 95; ST307_A, ST307_B, and ST307_C), one moderate transmission cluster (ST307_D, n = 3), and two singletons (Figure 3).

In ST307_D, two CR-Kp strains (8447_U390 and 8404_U406) harbored bla_{NDM-7} within large genomic islands (36.309 bp and 38.246 bp), identical to a region of a complete IncX3 plasmid from E. coli 8501_U437 (hybrid assembly available at (17); 100% of identity and query cover). This E. coli was isolated later in time from the same patient harbouring CR-Kp 8404_U406; therefore the genomic island with bla_{NDM-7} appears to have undergone intra-patient horizontal transfer between CR-Kp and E. coli, likely through plasmid mediation. Conversely, the high genomic similarity (13 SNPs) between the NDM-7 CR-Kp isolated from two different patients hospitalized in the same ICU within the same month suggests an inter-patient transmission event (Figure 3).

EC ST307_A (SNPs range: 20-61) comprised nine genomes. Eight harboured bla_{KPC-3}, four of which also carried bla_{NDM-1}. One strain (8813_U449) had only bla_{NDM-1}, suggesting KPC loss. In this strain, bla_{NDM-1} was located in a ~14Kbp genomic island (14,402 bp) that was identical to a NDM-1-harbouring genomic island present in the three CR-Kp with KPC-3+NDM-1 (8603_U455, 8494_U410, 8377_U307), while the remaining CR-Kp (8504_U391) had bla_{NDM-1} on a 5969-bp contig. These CR-Kp were isolated from patients in Pneumology (n = 3), Infectious Diseases, and Nephrology across different months (Figure 3).

EC ST307_B (n = 30, SNPs range: 4-45) was mainly bla_{KPC-3}-positive (n = 21), whereas nine CR-Kp lacked carbapenem resistance genes. One strain exhibited an OmpK35 truncation (88% amino acid length) with $bla_{CTX-M-15}$, suggesting that phenotypic carbapenem resistance was due to a combination of factors. The absence of genomic determinants correlated with the lack of the IncFIB(pQil) plasmid. ST307_A and ST307_B formed a highly supported, Italy-exclusive clade (bootstrap = 100), comprising a singleton (strain 7815_U174) and three Italian BV-BRC genomes (strains

BD_573.18714, BD_573.49061, and BD_573.49056). ST307_A isolates were mainly from June - October 2023, whereas ST307_B spanned October 2022 - May 2023 (Figure 3).

ST307_C was the largest EC (n = 84, SNPs: 6-277), and it was predominantly *bla*$_{KPC-3}$-positive (n = 76), with three *bla*$_{KPC-2}$ strains. Five CR-Kp strains lacked carbapenem resistance genes, correlating with IncFIB(pQil) plasmid absence. However, one of these (8626_U520) had an OmpK36 truncated. ST307_C was widely distributed across wards, particularly ICUs ($p < 0.05$), suggesting extensive ICU dissemination and potential reservoir. None of the three ECs carried hypervirulence genes (Figure 3).

For ST307 CR-Kp exhibiting phenotypic carbapenem resistance but no genomic resistance determinants, we hypothesised an *in vitro* loss during sample processing before sequencing. Notably, plasmid/gene loss was more frequent in ST307_B (9/30, 30%) than in ST307_C (5/84, 6.0%), suggesting ST307_B may be more prone to resistance loss. To investigate this hypothesis, we randomly selected five ST307_B and five ST307_C strains lacking genomic resistance determinants, along with nine control strains (five from ST307_B and four from ST307_C). All were carbapenemase-positive after cold storage. We then forced *in vitro* gene loss and measured the number of plating rounds before carbapenemases were no longer detected. No significant difference in resistance loss between ECs or case/control strains (range: 1-6 rounds).

Figure 3 - Core SNP phylogeny of 310 CR-Kp ST307: 180 from the HSM (black labels) and 258 from the BV-BRC database (grey labels). Left: phylogenetic tree with branch length with date of collection, wards, carbapenemases and Omp mutations mapped. Heatmaps display the presence/absence of virulence factors (blue) and plasmid incompatibility groups (magenta). **Right**:

tree topology on which bootstrap values (ranging from 80 to 100) are represented by light purple circles proportional to their support value. This figure was obtained using iTol (https://itol.embl.de/).

Other highly prevalent CR-Kp clones

K. pneumoniae ST6668 (n = 56) is a single-locus variant of ST147 (*phoE*: C414T) and a member of the Clonal Group (CG) 147. ST6668 was responsible for a large NDM-1 outbreak in the HSM and other nearby hospitals (18). All ST6668 strains were monophyletic (bootstrap = 100; SNPs range: 3-202; see Figure S1) and carbapenem- and ceftazidime-avibactam-resistant, but retained resistance to gentamicin.

The *K. pneumoniae* part of the CG258, in particular ST258 and ST512, were divided into one EC (ST258_ST512_A; n =1 6; SNP range: 8-103), one moderate transmission cluster, three one-to-one transmission clusters, and five singletons (Figure S2). Most (14/16) CR-Kp strains carried bla_{KPC-3}, whereas two Nephrology isolates lacked carbapenem resistance genes. Three CR-Kp (7642_U98, 33663_U302 and the singleton 8975_U560) were part of a highly supported larger cluster of only Italian strains. Within this cluster, there were strains involved in a previous HSM outbreak during 2015-2016 (7) (i.e. BD_573.27635, BD_573.27631, and BD_573.27630), and one CR-Kp isolated in Genoa (Italy) during 2021.

We also detected 15 *K. pneumoniae* of ST11, harbouring bla_{NDM-1} and the gene *mgrB* mutated (conferring resistance to colistin). In HSM, most were part of a large EC (ST11_A; n = 12; SNPs range: 8-41) and carried *ybt15* for Yersiniabactin, while three formed a moderate cluster with *ybt10* (Figure S3).

Finally, we also found 7 CR-Kp ST1519, predominantly isolated in ICUs and 4/7 were bla_{KPC-3}-positive forming a moderate transmission cluster and one singleton. . The remaining three harboured the $bla_{KPC-239}$ in a moderate transmission cluster and were resistant to ceftazidime-avibactam, while they remained susceptible at increased exposure to Meropenem (Figure S4).

K. pneumoniae ST101 (n = 16) was divided into two ECs (ST101_A and ST101_B) and one-to-one transmission cluster (Figure S5). Notably, ST101_A CR-Kp (n = 8; SNPs range: 5-27), were mainly collected in ICUs (December 2022-August 2023). Most of them (7/8) were bla_{KPC-3}-positive except for 8817_U532 that presented a novel variant. This variant was identical at 99.7% to $bla_{KPC-162}$ (806_808delGAC) and it was found for the first time in Argentina from a ceftazidime-avibactam resistant strain; also in our case, the strain was resistant to the molecule. All CR-Kp within the EC harboured the *ybt9* for Yersiniabactin.

The one-to-one transmission cluster was separated from ST101_A by a genome from the BV-BRC database (BD_573.36604 isolated in the Czech Republic) and it was composed of two CR-Kp isolated in General surgery and General medicine in September and December 2022.
CR-Kp in EC ST101_B (n = 6; SNPs range: 15-255) carried bla_{KPC-2} and were isolated from multiple wards. These CR-Kp were associated with multi-drug and hypervirulence traits, such as *iuc1* for Aerobactin (except for 8453_U401 with a partial *iuc1*), *ybt9* for Yersiniabactin and *rmpA2* gene in two CR-Kp. The sister clade of ST101_B (n = 6) encompassed only Italian strains from the BV-BRC database, carrying bla_{KPC}-variants, *iuc1* for Aerobactin, and *ybt9* for Yersiniabactin. The aforementioned CR-Kp, together with ST101_B, were part of a broader Italian cluster of 2013-2021 CR-Kp (n = 25), supporting the Italian origin of ST101_B.

ST13 (n = 7) and ST219 (n = 4) were also identified. ST13 were collected exclusively from April to September 2023 in multiple wards and were mainly included into EC ST13_A (n = 5; SNPs range: 12-26). ST13_A had two bla_{KPC-3}-positive CR-Kp concordant with the presence of IncFIB(pQil), while three lacked carbapenem resistance genes. All ST13_A had hypervirulence markers as *clb3* and *ybt1*. The remaining two formed a one-to-one transmission cluster (21 SNPs) and were bla_{KPC-2} and *ybt1*-positive. Both were isolated from the same patient, in June and September 2023 (Figure S6). The latter also presented a mutated *mgrB* which resulted in the phenotypic resistance to colistin.
ST219 genomes formed a moderate cluster (n = 4; SNPs range:17-47); 3/4 were collected in June 2023 from mainly ICUs (n = 2). All harboured bla_{KPC-3}, *ybt14* for Yersiniabactin and multiple incompatibility groups, such as IncFIB(K), IncFII(K), IncR, IncFIA(HI1), Col440I, and Col440II (Figure S7).

In general, ST307 was mainly associated with the ICUs, whereas ST101 and ST11 were associated with General surgery and General medicine wards, respectively ($p < 0.05$). For the most prevalent clones, we identified nine ECs that persisted over the year and spread across HSM, particularly in ICUs (Figure 4 a and b).

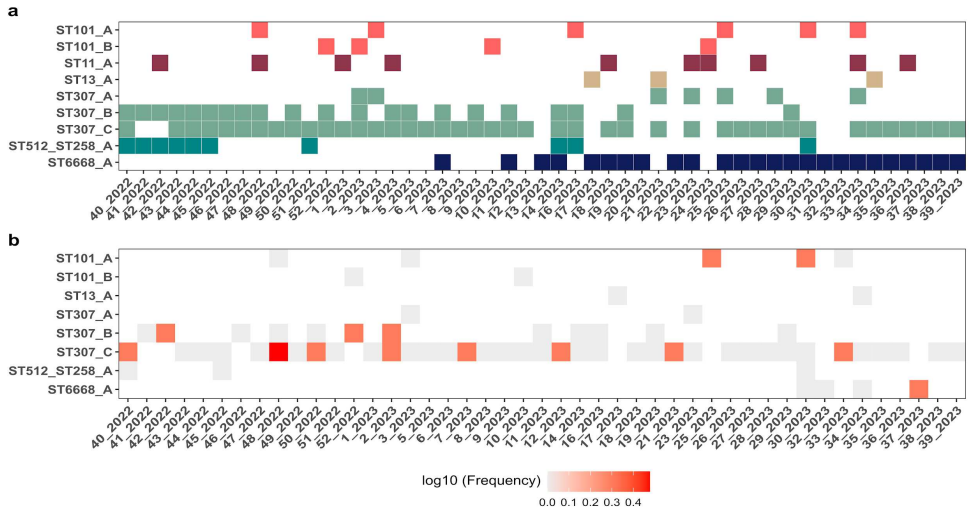


Figure 4 - Temporal distribution of Epidemic Clusters (ECs) in carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) clones. (a) Weekly presence/absence of CR-Kp ECs from October 1, 2022 (Week: 40_2022), to September 31, 2023 (Week: 39_2023), coloured by ST. (b) Heatmap illustrates the temporal distribution of each CR-Kp ECs in ICUs; color intensity reflects the log10-transformed CR-Kp frequency observed in each bin.

Virulence factors and Inc groups in *K. pneumoniae*

Siderophores were detected at varying prevalence: Yersiniabactin in 43.7% (124/284), Aerobactin in 3.2% (9/284), and Colibactin in 2.11% (6/284) of CR-Kp, mainly in ECs ST101_A and B, ST11_A, ST13_A, ST512_ST258_A, and ST6668_A (see Table S5). A truncated *rmpA2* gene for the hypermucoid phenotype was also found in two ST101.

All KPC-Kp strains carried at least one Inc group, with a median of 4 (Q1-Q3: 3-5). ST219 and ST6668 had the highest Inc counts (median: 6 [Q1-Q3: 6-6] and 5.5 [Q1-Q3: 4-6]), while ST307 and ST258/ST512 had median of 4 (IQR: 3-4) and 4 (IQR: 4-5), respectively.

Klebsiella spp.

We identified 14 additional CRE within the *Klebsiella* genus. Seven were *K. aerogenes* of diverse STs, including two novel ones (ST731, ST732; see Table S5). These clones, unrelated to highly prevalent human-associated lineages (e.g., ST4, ST93 [18]), were sporadic in HSM. Two carried *bla*$_{KPC-2}$, one *bla*$_{KPC-3}$, one *bla*$_{NDM-1}$, and one *bla*$_{NDM-7}$. Notably, strain 8457_U388 harboured *bla*$_{NDM-7}$ on a ~45kb contig identical to the IncX3 plasmid sequence found in *E. coli* 8501_U437. Additionally, we detected three *K. quasipneumoniae* subsp*. quasipneumoniae* (with *bla*$_{VIM-1}$), two *K. variicola* subsp*. variicola* (with *bla*$_{VIM-1}$ and *bla*$_{KPC-3}$), and two *Klebsiella oxytoca* complex, one carrying *bla*$_{VIM-1}$ and one lacking carbapenem resistance determinants.

Escherichia coli

We identified 29 carbapenem-resistant *E. coli* (CR-Ec) strains spanning multiple STs, with ST131 (n = 10) as the most prevalent. Other detected STs included ST69, ST410, ST401, and ST58, with additional sporadic clones (see Table S6). Carbapenemase genes were found in 25/29 CR-Ec (83.3%), with *bla*$_{KPC-3}$ (n = 8 out of 9) being the most frequent among KPC variants, followed by *bla*$_{NDM-1}$ (in 3 out of 7 CR-Ec), *bla*$_{OXA-244}$ (n=5/6), and *bla*$_{VIM-1}$ (n = 5). See Table S3 for further details. Co-occurring resistance was observed in two CR-Ec: ST10 (strain 8234M_U285) with *bla*$_{NDM-5}$ and *bla*$_{OXA-244}$ and ST3564 (strain 8501_U437) carrying *bla*$_{NDM-7}$ plus *bla*$_{VIM-1}$, recently described by Mattioni Marchetti and colleagues (17). No association was found between clones and HSM wards.

E. coli ST131

The 10 ST131 CR-Ec strains were isolated from December 2022 to July 2023 across multiple wards. All belonged to phylogroup B2, with 9/10 showing serotype O25b:H4. However, the first CR-Ec belonged to the less common O16:H5. The most common carbapenemase was the chromosomally encoded *bla*$_{OXA-244}$ (n = 4), with the detection of two *bla*$_{KPC-3}$ and one *bla*$_{NDM-1}$. Three strains lacked carbapenemase genes and were susceptible to Imipenem and Meropenem, with MIC = 0.25 mg/L. In the core SNP phylogeny, 7/10 isolates were singletons, while three consisted of a moderate transmission cluster (Figure S8). The most common plasmid Inc were of Col-type (i.e. MG828 and 156) and IncFIB, with IncHI1B_pNDM-MAR and IncI1 in *bla*$_{NDM-1}$-positive, and IncN and IncX3 in the *bla*$_{KPC-3}$ strain (further details in (19)).

Other prevalent CR-Ec clones

Among other prevalent CR-Ec clones, ST69 (n = 3) was detected in different wards from November 2022 to March 2023. These strains harboured one *bla*$_{KPC-2}$, *bla*$_{KPC-3}$, and *bla*$_{VIM-1.}$ All belonged to phylogroup D and exhibited distinct serotypes (O15:H18, OGp9:H18, and O117:H4). The core SNP phylogeny confirmed them as singletons (Figure S9). IncFIB(pQil) was found in KPC-strains, while IncA/C in the VIM-1-positive.

ST58, ST401, and ST410 were found in two strains each. ST58 CR-Ec was identified in November 2022 and September 2023, both belonging to phylogroup B1 but carrying different serotypes. The September isolate was NDM-1-positive and carried three distinct Inc groups (IncA/C2, IncFIB/IncHI1B_pNDM-MAR, and IncI1). Phylogenetic analysis confirmed that these strains were unrelated (Figure S10).

The ST401 clone was found in General Surgery (November 2022) and Hematology (September 2023), with both strains belonging to phylogroup A but differing in serotypes and carbapenemase genes (*bla*$_{KPC-3}$ *vs*. *bla*$_{VIM-1}$). Phylogenetic analysis confirmed no relation between them (Figure S11).

Two ST410 strains, isolated in February and May 2023, were both part of serogroup C but carried distinct carbapenemases. The first one harboured *bla*$_{OXA-484}$, located in the ColKP3 plasmid, along with *bla*$_{CMY-42}$, located in the contig with IncI group, and *bla*$_{TEM-1}$. In contrast, the other carried *bla*$_{NDM-5}$, along with *bla*$_{CMY-2}$ β-lactamases and multiple plasmid Inc groups (IncFIIpAMA1167-NDM-5, IncFIA, and IncIB). These two strains were also phylogenetically unrelated (Figure S12).

Other species

Among other CRE, nine *Enterobacter* spp. were carbapenem-resistant: eight *E. hormaechei* and one *E. asburiae*. All but one carried *bla*$_{VIM-1}$; the remaining carried *bla*$_{NDM-1}$ (Table S3). Most *E. hormaechei* clones were sporadic (Table S8), except two ST190 consisting of a one-to-one transmission cluster (41 SNPs, see Figure S13). Similarly, two ST66 formed a one-to-one transmission cluster, but they were 257 SNPs apart (Figure S14). Additionally, four carbapenem-resistant *Citrobacter* spp. were detected: two *C. freundii* (with *bla*$_{VIM-1}$), one *C. telavivensis* of newly assigned ST1266, carrying *bla*$_{VIM-1}$ and one *C. amalonaticus* with *bla*$_{NDM-7}$ (Table S8). The latter was isolated in June 2023 from the same patient with *K. aerogenes* 8457_U388, CR-Kp 8404_U406, and CR-Ec 8501_U437, all *bla*$_{NDM-7}$-positive. As in other NDM-7-producing CRE, in this case, four contigs covered 99% of the entire sequence (100% identity) of the IncX3 plasmid found in *E. coli* 8501_U437.

Discussion

Real-time genomics is increasingly integrated into traditional surveillance practices to reduce the spread of CRE in healthcare settings (20). In fact, robust screening campaigns coupled with reliable sequencing technologies were essential for identifying Epidemic clusters (ECs) and singletons that could potentially evolve into larger outbreaks.

During one year of surveillance at the HSM (October 2022 to September 2023), the prevalence of CRE was 23.4% of the total number of *Enterobacterales* isolates. Genomic characterisation was performed on the first CRE for each patient considering the bacterial species and antimicrobial susceptibility profile, obtaining 85.3% CRE from colonisation specimens. 294 patients were included in this study, 47 of them experienced invasive infections and nearly half were previously colonised, confirming colonisation as a key driver of transmission and infection risk (21,22).

In this study *K. pneumoniae* accounted for more than 80% of CRE and it has been confirmed as the dominant healthcare-associated pathogen in Italy. Twenty-six clones were identified, with the majority of strains grouped into nine ECs, indicating significant intra-hospital spread. ST307 was the most prevalent clone, consistent with its status as the actual dominant lineage in Italy, alongside ST258/ST512, which has circulated in HSM since the early 2010s (7, 23).

In contrast, ST6668 emerged in February 2023, causing a large NDM-producing outbreak in HSM and nearby hospitals (18), further narrowing patients' treatment options (24). The other NDM-producing CR-Kp was colistin-resistant ST11 bearing $bla_{\text{NDM-1}}$. These strains were not linked to the ST11 clone circulating in China, which also harbours colistin resistance determinants associated with KPC-2, NDM-1, or OXA-48 carbapenemases (25).

ECs ST101_A, ST307_B, and ST307_C were ICU-associated, confirming that the ICU wards are reservoirs for resistant pathogens. Although the sampling bias of screening rectal swabs may enlarge the percentage of isolations, these results align with evidence that ICU environments, with high antibiotic usage and fragile patients, create ideal conditions for the survival and spread of multidrug-resistant pathogens (26). In contrast, ST6668_A spread across multiple wards, suggesting diverse transmission routes within the HSM.

In addition to multidrug resistance, 3.5% CR-Kp exhibited hypervirulent traits, indicating a convergence between these two characteristics. The Yersiniabactin siderophore was found in 43.5% of the patients, reflecting

the overall prevalence of the CR-Kp population. Other siderophores were detected at lower prevalence in specific ECs, such as *iro1* (3.2%) for Aerobactin in ST101_B and *clb3* (2.1%) for Colibactin in ST13_A. This evidence supports the general trend of largely non-overlapping resistant and hypervirulent populations (27).

Finally, other less prevalent species were *E. coli*, *Enterobacter* spp., and *Citrobacter* spp.*;* these were mainly (58.9% of cases) found as singletons, with no further HSM spread. However, the IncX3 plasmid encoding NDM-7 was detected across multiple species, despite not being associated with any major worldwide spread clones in which the plasmid has been detected (24). Specifically, the IncX3 plasmid was observed to have likely transferred from *K. pneumoniae* ST307 to *K. aerogenes*, *C. amalonaticus*, and *E. coli* ST3564 within a single patient hospitalised in the ICU during June 2023 (strains 8404_U406, 8457_U388, 8456_U381, and 8501_U437, respectively). On the contrary, the other NDM-7 CR-Kp (strain 8447_U390) was isolated from a second patient hospitalised in the same ICU during the same month, suggesting inter-patient plasmid transmission.

This study has some limitations. First we sequenced and analysed the first CRE isolate per patient, selecting distinct bacterial species and antibiotic susceptibility profiles. However, even among isolates with the same susceptibility profiles, the underlying mechanisms of carbapenem resistance may differ, potentially introducing a selection bias in strain representation. Second, we observed 32 dataset genomes lacking carbapenem resistance genes. Generally, we performed bacterial DNA extraction from colonies grown in pure culture, thus requiring one or more isolation steps (multiple generations) from the original sample culture. We suppose that the subsequent plating could have influenced the loss of some resistance factors. The hypothesis was fostered by the observation of *in vitro* plasmid loss even after one single round of plating. This finding highlights the importance of minimising sample handling and isolation rounds before DNA extraction for sequencing.

Conclusion

Our findings highlight the role of sequencing as an integral component of clinical management and epidemiological surveillance. The effective management of CRE demands a dual approach: controlling the proliferation of dominant clones while curtailing the dissemination of sporadic, emerging strains. The integration of advanced genomic tools into routine surveillance enables real-time monitoring of clonal and antimicrobial determinant spreading, allowing timely implementation of targeted infection control measures (e.g. tailored cohorting strategies) and enhancing overall antimicrobial resistance management.

Acknowledgement

This work was supported by NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project PE00000007, INF-ACT), by Ricerca Corrente n° 08075723 (1034-rcr2023–43), and by JPIAMR 2024_ IMPACT 197.

Data availability

All genomic data were available under BioProject PRJNA1266098 and PRJNA1019273.

Additional material can be found in the “Supplemental material” of the manuscript.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105). The work described herein is a study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

References

1. Key facts and figures on World Hand Hygiene Day. https://www.who.int/campaigns/world-hand-hygiene-day/key-facts-and-figures (accessed Jan 30, 2025).
2. Healthcare-associated infections. https://www.ecdc.europa.eu/en/healthcare-associated-infections (accessed Nov 4, 2024).
3. Balasubramanian R, Van Boeckel TP, Carmeli Y, Cosgrove S, Laxminarayan R. Global incidence in hospital-associated infections resistant to antibiotics: An analysis of point prevalence surveys from 99 countries. PLoS Med 2023; 20: e1004178.
4. European Centre for Disease Prevention and Control. ECDC country visit to Italy to discuss antimicrobial resistance issues. Stockholm: ECDC; 2017.
5. Fox JM, Saunders NJ, Jerwood SH. Economic and health impact modelling of a whole-genome sequencing-led intervention strategy for bacterial healthcare-associated infections for England and for the USA. Microb Genom2023; 9. DOI:10.1099/mgen.0.001087.
6. Gordon LG, Elliott TM, Forde B, et al. Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant

bacterial pathogens in Queensland, Australia. BMJ Open 2021; 11: e041968.
7. Ferrari C, Corbella M, Gaiarsa S, et al. Multiple Klebsiella pneumoniae KPC Clones Contribute to an Extended Hospital Outbreak. Front Microbiol 2019; 10: 2767.
8. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 12.0, 2022. . http://www.eucast.org (accessed Oct 15, 2024).
9. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 13.1, 2023. http://www.eucast.org (accessed Oct 15, 2024).
10. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun 2021; 12: 4188.
11. Wyres KL, Wick RR, Gorrie C, et al.Identification of Klebsiella capsule synthesis loci from whole-genome data. Microb Genom2016; 2: e000102.
12. Bortolaia V, Kaas RS, Ruppe E, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 2020; 75: 3491–500.
13. Joensen KG, Scheutz F, Lund O, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 2014; 52: 1501–10.
14. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and Easy In Silico Serotyping of Escherichia coli Isolates by Use of Whole-Genome Sequencing Data. J Clin Microbiol 2015; 53: 2410–26.
15. Carattoli A, Zankari E, García-Fernández A, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 2014; 58: 3895–903.
16. Batisti Biffignandi G, Bellinzona G, Petazzoni G, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics 2023; 39. DOI:10.1093/bioinformatics/btad571.
17. Mattioni Marchetti V, Corbella M, Piazza A, et al. Polyclonal carbapenemase-producing Escherichia coli in Northern Italy: the emergence of NDM-7. Front Cell Infect Microbiol 2025; 15.
18. Mileto I, Petazzoni G, Corbella M, et al. Rapid spread of a novel NDM-producing clone of Klebsiella pneumoniae CC147, Northern Italy, February to August 2023. Euro Surveill 2023; 28. DOI:10.2807/1560-7917.ES.2023.28.42.2300522.
19. Piazza A, Corbella M, Mattioni Marchetti V, et al. Clinical isolates of ST131 blaOXA-244-positive Escherichia coli, Italy, December 2022

to July 2023. Euro Surveill 2024; 29. DOI:10.2807/1560-7917.ES.2024.29.8.2400073.
20. Baker KS, Jauneikaite E, Hopkins KL, et al. Genomics for public health and international surveillance of antimicrobial resistance. Lancet Microbe 2023; 4: e1047–55.
21. Xiao Y, Duan J, Tan C, et al.Correlation between intestinal CRE colonisation and consequent systemic infection in hospitalised patients in China. Sci Rep2024; 14: 26017.
22. David S, Reuter S, Harris SR, et al. Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. Nat Microbiol 2019; 4: 1919–29.
23. Rossi M, Chatenoud L, Gona F, et al.Characteristics and Clinical Implications of Carbapenemase-Producing Klebsiella pneumoniae colonisation and Infection, Italy. Emerging Infect Dis2021; 27: 1416–26.
24. Wu W, Feng Y, Tang G, Qiao F, McNally A, Zong Z. NDM Metallo-β-Lactamases and Their Bacterial Producers in Health Care Settings. Clin Microbiol Rev 2019; 32. DOI:10.1128/CMR.00115-18.
25. Liao W, Liu Y, Zhang W. Virulence evolution, molecular mechanisms of resistance and prevalence of ST11 carbapenem-resistant Klebsiella pneumoniae in China: A review over the last 10 years. J Glob Antimicrob Resist 2020; 23: 174–80.
26. Lee C-R, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global Dissemination of Carbapenemase-Producing Klebsiella pneumoniae: Epidemiology, Genetic Context, Treatment Options, and Detection Methods. Front Microbiol 2016; 7: 895.
27. Wyres KL, Lam MMC, Holt KE. Population genomics of Klebsiella pneumoniae. Nat Rev Microbiol 2020; 18: 344–59.

CHAPTER 5:

Rapid spread of a novel NDM-producing clone of *Klebsiella pneumoniae* CC147, Northern Italy, February to August 2023

Irene Mileto[1,2] *, **Greta Petazzoni**[1,3] * , Marta Corbella[1], Stefano Gaiarsa[1], Cristina Merla[1], Angela Kuka[1], Marina Ramus[1,2], Cristina Terulla[4] , Micaela Brandolini[5] , Antonio Piralla[1], Patrizia Cambieri[1], Fausto Baldanti[1,3]

1. Microbiology and Virology Unit, IRCCS Fondazione Policlinico San Matteo, Pavia, Italy
2. Specialization School of Microbiology and Virology, University of Pavia, Pavia, Italy
3. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia, Pavia, Italy
4. ASST Pavia, Ospedale Civile di Voghera, Voghera, Italy
5. ASST Pavia, Ospedale Unificato Broni-Stradella, Internal Medicine Unit, Stradella, Italy

* These two authors contributed equally

The manuscript has been **published on Eurosurveillance**, October 2023, Volume 28, Issue 42,
https://doi.org/10.2807/1560-7917.ES.2023.28.42.2300522

In this work I took part in the conceptualization and carried out all the *in silico* analyses. I also contributed substantially to drafting and revising the manuscript.

Introduction

The diffusion of New Delhi metallo-beta-lactamase (NDM)-producing *Klebsiella pneumoniae* (Kp) poses a significant public health challenge due to its transmission within healthcare facilities and its extensive resistance to most antimicrobial agents. Current therapeutic strategies for NDM-producing *Enterobacterales* primarily rely on cefiderocol or a combination of ceftazidime-avibactam and aztreonam (1). Globally, NDM production has been linked to various *Kp* sequence types (ST), including ST11 and ST147. Notably, ST147 and its related clonal complex (CC) represent a high-risk international clone (2), previously responsible for a major outbreak in the Tuscany region of Italy between 2018 and 2021 (3,4). This study reports the isolation of a novel NDM-producing Kp sequence type, identified as ST6668 (belonging to CC147), which was associated with a multi-center outbreak in the province of Pavia (Northern Italy) from February to August 2023.

Event description

Prior to 2022, *K. pneumoniae* carbapenemase (KPC) was the predominant carbapenemase identified in Kp strains at the Fondazione IRCCS Policlinico San Matteo (HSM), a 900-bed teaching hospital located in Pavia. While the incidence of KPC remained stable between August 2022 and August 2023 (averaging 16.3 ± 5.6 isolates per month), a distinct surge in NDM incidence was observed starting in April 2023 (Figure 1A). Consequently, the prevalence of NDM among carbapenem-resistant Kp isolates increased significantly, rising from 10% in April to 61.8% by August 2023 (Figure 1B).

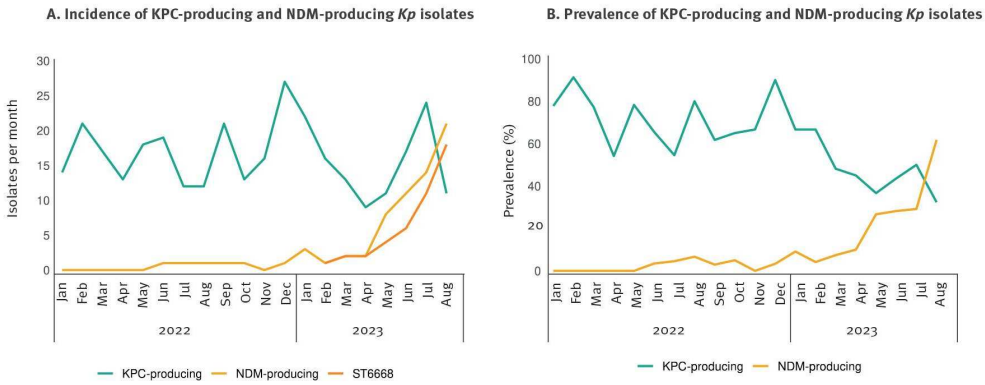


Figure 1 - **Incidence and prevalence of *K. pneumoniae*.** Incidence (A) and prevalence (B) of *Klebsiella pneumoniae* carbapenemase (KPC) and New Delhi metallo-beta-lactamase (NDM)-producing *Klebsiella pneumoniae* isolates at Fondazione IRCCS Policlinico San Matteo (HSM) from January 2022 to August 2023 (n = 394).

Between January and August 2023, 117 NDM-producing Kp isolates were collected at HSM from a cohort of 66 patients. Whole genome sequencing (WGS) was performed on 66 isolates (representing the first isolate for each patient) using Illumina technology, following the protocol described in (5). Furthermore, we also sequenced the genomes of eight NDM-producing Kp strains isolated from patients at Ospedale Civile di Voghera, a 200-bed facility within the Pavia Health Authority Unit (Pavia-HAU), situated approximately 30 km from Pavia. Genomic characterization of the 74 total isolates revealed that 52 belonged to a single clone exhibiting a unique MLST allelic profile. Upon submission to the BIGSdb database (https://bigsdb.pasteur.fr), this profile was validated as a novel clone, designated ST6668. This specific sequence type has been identified as the primary driver of the rapid increase in NDM-producing Kp observed since February 2023 (Figure 1A).

Among the 52 identified Kp ST6668 strains, 44 had been collected from patients in HSM and eight from the Voghera hospital. Regarding the isolation source, 44 strains were obtained from rectal swabs and urine samples, five from blood cultures, and three from other sites (bronchoalveolar lavage, pus, and drainage fluid). The patient cohort (n=52) included 31 males, with a mean age of 70 years (SD: ± 15.4; range: 31–90 years). At HSM, patient distribution included admissions to specialist medical units (pneumology, nephrology/rheumatology, infectious diseases; n=13), oncology (n=10), Intensive Care Units (ICUs; n=7), surgery units (n=5), and general medicine (n=3), alongside six outpatients.

In accordance with internal protocols for high-risk patients and ICUs admissions, 42 patients at HSM underwent rectal swab screening for multidrug-resistant (MDR) *Enterobacterales* colonization. Results indicated that 36 patients were colonized upon admission, whereas six acquired colonization during hospitalization (mean time: 15 days post-admission). Seven patients progressed to develop bloodstream or pulmonary infections. Additionally, three patients initially admitted to Voghera were subsequently transferred to a smaller Pavia-HAU hospital located 20 km from both Pavia and Voghera.

Clone characterisation

Antimicrobial susceptibility profiling

Antimicrobial susceptibility profiles were determined utilizing the Sensititre DKMGN broth microdilution system (ThermoFisher), with results interpreted in adherence to the clinical breakpoints established by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (6). Resistance to the majority of tested antibiotics - including carbapenems, ceftazidime-avibactam, and amikacin - was observed across all 52 isolates. Conversely, resistance to cefiderocol was detected in only two strains. Susceptibility to gentamicin was retained in all cases. Detailed minimum inhibitory concentrations (MIC) are provided in the Supplementary Table. Furthermore, the combination of ceftazidime-avibactam and aztreonam demonstrated synergistic activity against all screened strains (7).

Typing and phylogenetic analysis

De novo assembly of Illumina short reads was performed using Shovill (github.com/tseemann/shovill). Then, we used P-DOR (8) to obtain the single nucleotide polymorphism (SNP)-based maximum-likelihood phylogeny. The analysis included the 52 isolates from this study alongside 38 high-quality similar genomes retrieved from the BV-BRC database (https://www.bv-brc.org/) (Figure 2).

The 52 genomes originating from the Pavia region formed a monophyletic clade (SNP distance range: 4–51), indicating that these strains represent a single epidemic event. All outbreak isolates belonged to ST6668, which differs from ST147 by a point mutation (C414T) in the *phoE* locus; whereas the 38 genomes used as background belonged to ST147.

The genomes most closely related to the outbreak clade were recovered from samples in Lebanon (blood, 2020) and the United States (urine and rectal swabs, 2022). Moreover, none of the closest strains retrieved in the phylogeny was the ST147 clone responsible for the large outbreak in Tuscany (4). This evidence suggests that the ST6668 clone likely resulted from an independent importation event of the CC147 lineage.

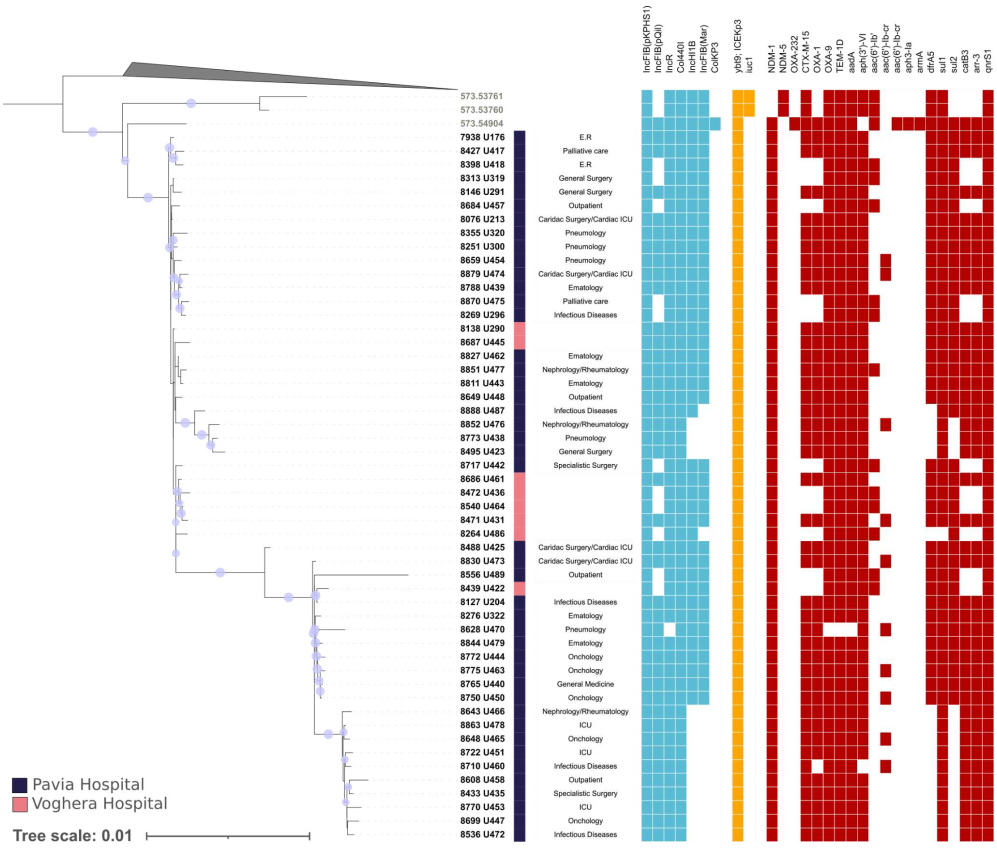


Figure 2 - **Phylogenetic analysis of carbapenemase-producing *Klebsiella pneumoniae* ST6668.** The phylogeny comprises isolates from Pavia hospital (Fondazione IRCCS Policlinico San Matteo; HSM) and Voghera hospital (Ospedale Civile di Voghera) from February to August 2023 (n = 52). The maximum-likelihood phylogeny of 90 Kp strains is inferred on coreSNPs with RAxML, including 52 from the outbreak (black labels) and 38 from the BV-BRC database (grey labels – 35 leaves have been collapsed). Bootstrap values above 70/100 are indicated with proportional violet dots. Isolate information is reported alongside the tree, with the hospital of isolation (HSM/Voghera hospital), and wards for HSM. Presence/absence matrix shows the pattern of plasmids (blue), virulence determinants (yellow) and resistance determinants (red).

Strain characterisation

The identification of antibiotic resistance determinants and virulence factors was conducted using Kleborate (9), while plasmid replicon typing was performed via PlasmidFinder (10). Detailed profiles concerning genomic resistance, virulence attributes, and plasmid content are provided in the Supplementary Table. The bla_{NDM-1} gene, responsible for carbapenem resistance, was ubiquitously detected across all outbreak strains. Additionally, every isolate carried the *ybt* gene (allele 9) - encoding yersiniabactin - located within the ICEKp3 integrative element.

Regarding the plasmid landscape, six distinct sequences were identified. Two of these, IncFIB(pKPHS1) and Col440I, were conserved in all 52 isolates. The remaining four displayed a variable presence/absence pattern that demonstrated a high degree of concordance with specific antimicrobial resistance determinants (Figure 2). Specifically, the IncFIB(pQil) plasmid was linked to the carriage of *bla*CTX-M-15, *bla*OXA-1, *catB3*, and *arr-3*. In contrast, *dfrA5* and *sul2* genes were associated with the IncHI1B and IncFIB(Mar) plasmid sequences. These findings point to a scenario characterized by the frequent exchange of genetic material. This hypothesis is further supported by phylogenetic evidence: isolates that lost the IncHI1B and IncFIB(Mar) plasmids (along with their associated resistance genes) clustered into two distinct, high-confidence monophyletic subclades (Figure 2).

Discussion

Here we report the isolation of a new Kp clone, ST6668, which disseminated rapidly across a significant patient cohort in two hospitals within the Pavia province. The detection of the strain from 46 inpatients and six outpatients indicates that circulation is currently occurring beyond the strict boundaries of the healthcare environment.

Phylogenomic analysis of the 52 epidemic isolates revealed that the ST6668 cluster presents distinct high-confidence monophyletic subclades, although internal resolution remains unsolved. Notably, the eight isolates originating from the Voghera hospital did not group into a single unique cluster but were dispersed, suggesting multiple independent transmission events between healthcare facilities in the Pavia area.

The introduction of this new clone caused a shift in the local carbapenemase epidemiology. Historically, KPC was the predominant enzyme; however, NDM detection rates have surged since April 2023, establishing NDM as the dominant carbapenemase by August 2023.
The presence in the new clone of a large number of resistance genes, including $bla_{\text{NDM-1}}$, could have facilitated the rapid spread of this clone (11), while no particular virulence factor was observed. Moreover, the high mean age of the patients involved suggests that they may have a history of frequent admissions to hospitals or long-term care facilities. This could have favoured the spread of the ST6668 over the Pavia area. Indeed, three patients were transferred from Voghera to another nearby hospital, belonging to the Pavia-HAU.

This scenario underlines the need to establish local, regional and international surveillance and infection prevention strategies (12) in any healthcare setting. In particular, real-time genomic surveillance has proven essential in the timely detection of novel fast-spreading clones and variants of clinical importance (13,14).

Although most of the patients of this study were colonised by ST6668 Kp without developing infections, hospitalised patients tend to remain colonised with MDR strains for the duration of their stay, increasing the risk of transmission of MDR strains to other patients (15).

Conclusion

The fast spread of ST6668 and other NDM-producing clones should be controlled, since their reduced susceptibility to most antibiotics allows few therapeutic options. In light of how rapidly the new clone has spread within the healthcare setting in the Pavia area, we underline the importance of prompt communication between microbiologists and clinicians.

Acknowledgement

This research was financially supported by EU funding within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. P E00000007, INF-ACT) and by Ricerca Corrente n 08022218 “Real-Time bioinformatic surveillance of nosocomial infections and multidrug resistant bacteria to improve infection control”.

Data availability

Sequence data of 52 ST6668 K. pneumoniae are available under BioProject N PRJNA1019273.

Additional material can be found in the “Supplementary Data” section of the manuscript.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105). The work described herein is a retrospective study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

References

1. Tamma PD, Aitken SL, Bonomo RA, Mathers AJ, van Duin D, Clancy CJ. Infectious Diseases Society of America 2023 guidance on the treatment of antimicrobial resistant Gram-negative infections. Clin Infect Dis. 2023;ciad428. https://doi.org/10.1093/cid/ciad428
2. Peirano G, Chen L, Kreiswirth BN, Pitout JDD. Emerging antimicrobial-resistant high-risk Klebsiella pneumoniae clones ST307

and ST147. Antimicrob Agents Chemother. 2020;64(10):e01148-20. https://doi.org/10.1128/AAC.01148-20

3. Di Pilato V, Henrici De Angelis L, Aiezza N, Baccani I, Niccolai C, Parisio EM, et al. Resistome and virulome accretion in an NDM-1-producing ST147 sublineage of Klebsiella pneumoniae associated with an outbreak in Tuscany, Italy: a genotypic and phenotypic characterisation. Lancet Microbe. 2022;3(3):e224-34. https://doi.org/10.1016/S2666-5247(21)00268-8
4. Martin MJ, Corey BW, Sannio F, Hall LR, MacDonald U, Jones BT, et al. Anatomy of an extensively drug-resistant Klebsiella pneumoniae outbreak in Tuscany, Italy. Proc Natl Acad Sci USA. 2021;118(48):e2110227118. https://doi.org/10.1073/pnas.2110227118
5. Merla C, Kuka A, Petazzoni G, Postiglione U, Zatelli M, Gaiarsa S, et al. Livestock-associated methicillin-resistant Staphylococcus aureus in inpatients: a snapshot from an Italian hospital. J Glob Antimicrob Resist. 2022;30:10-5. https://doi.org/10.1016/j.jgar.2022.05.013
6. The European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters. Version 13.0. Växjö: EUCAST; 2023. Available from: https://www.eucast.org
7. Wenzler E, Deraedt MF, Harrington AT, Danizger LH. Synergistic activity of ceftazidime-avibactam and aztreonam against serine and metallo-β-lactamase-producing gram-negative pathogens. Diagn Microbiol Infect Dis. 2017;88(4):352-4. https://doi.org/10.1016/j.diagmicrobio.2017.05.009
8. Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics. 2023;39(9):btad571. https://doi.org/10.1093/bioinformatics/btad571
9. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun. 2021;12(1):4188. https://doi.org/10.1038/s41467-021-24448-3
10. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895-903. https://doi.org/10.1128/AAC.02412-14
11. Gullberg E, Albrecht LM, Karlsson C, Sandegren L, Andersson DI. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. MBio. 2014;5(5):e01918-14. https://doi.org/10.1128/mBio.01918-14
12. Ciccolini M, Donker T, Köck R, Mielke M, Hendrix R, Jurke A, et al. Infection prevention in a connected world: the case for a regional approach. Int J Med Microbiol. 2013;303(6-7):380-7. https://doi.org/10.1016/j.ijmm.2013.02.003

13. Forde BM, Bergh H, Cuddihy T, Hajkowicz K, Hurst T, Playford EG, et al. Clinical implementation of routine whole-genome sequencing for hospital infection control of multi-drug resistant pathogens. Clin Infect Dis. 2023;76(3):e1277-84. https://doi.org/10.1093/cid/ciac726
14. Kwong JC, Lane CR, Romanes F, Gonçalves da Silva A, Easton M, Cronin K, et al. Translating genomics into practice for real-time surveillance and response to carbapenemase-producing Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak. PeerJ. 2018;6:e4210. https://doi.org/10.7717/peerj.4210
15. Vink JP, Otter JA, Edgeworth JD. Carbapenemase-producing Enterobacteriaceae - once positive always positive? Curr Opin Gastroenterol. 2020;36(1):9-16.

CHAPTER 6:

The COVID-19 pandemic sparked off a large-scale outbreak of carbapenem-resistant *Acinetobacter baumannii* from the endemic strains at an italian hospital

Greta Petazzoni[1] *, Greta Bellinzona[2] *, Cristina Merla[1], Marta Corbella[1], Vincenzina Monzillo[1,3], Ørjan Samuelsen[4,5], Jukka Corander[6,7], Davide Sassera[2], Stefano Gaiarsa[1], Patrizia Cambieri[1]

1. Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
2. Department of Biology and Biotechnology, University of Pavia, Pavia, Italy
3. Department of Internal Medicine and Medical Therapy, University of Pavia, Pavia, Italy
4. Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway
5. Institute of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway
6. Department of Biostatistics, University of Oslo, Oslo, Norway
7. Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

* These two authors contributed equally

The manuscript has been **published on Microbiology Spectrum**, March/April 2023, Volume 11, Issue 2.
https://doi.org/10.1128/spectrum.04505-22

In this work I retrieved the metadata and took part in the *in silico* analyses. I also contributed substantially to drafting and revising the manuscript.

Abstract

Acinetobacter baumannii represents a critical nosocomial pathogen, posing a growing public health threat driven by the escalating incidence of multidrug-resistant (MDR) lineages. Throughout the COVID-19 pandemic, numerous outbreaks attributed to MDR *A. baumannii* clones were documented globally. This study details a comprehensive investigation of an MDR *A. baumannii* outbreak that affected the Fondazione IRCCS Policlinico San Matteo (Pavia, Italy).

The dataset comprised 96 strains isolated between January and July 2020 from 41 hospitalized patients (stratified by SARS-CoV-2 status) across various wards. Characterization was performed via phenotypic assays and genomic analysis, utilizing a hybrid sequencing approach (Illumina and Nanopore). Antimicrobial susceptibility testing demonstrated universal resistance to carbapenems, a phenotype genomically linked to the presence of the $bla_{\text{OXA-23}}$ carbapenemase gene. Virulence profiling revealed that all isolates harbored determinants for biofilm formation. Furthermore, plasmidome analysis identified two distinct plasmids, one of which was approximately 100 kbp in size and carried a phage sequence.

Phylogenetic reconstruction based on the core genome integrated outbreak strains with background genomes from public repositories and local surveillance data. All isolates were assigned to the globally disseminated Sequence Type 2 (ST2) and segregated primarily into two distinct clades. Notably, outbreak isolates clustered with surveillance strains collected in 2019, indicating that the event was driven by two lineages already endemic to the hospital prior to the pandemic onset. The rapid intra-hospital dissemination of *A. baumannii* was likely exacerbated by the extreme operational pressure of the first COVID-19 wave, which inadvertently compromised adherence to standard infection prevention and control practices.

Importance

The COVID-19 pandemic, particularly during its initial surge, placed unprecedented strain on healthcare management, creating conditions favorable for the dissemination of nosocomial pathogens. Due to its ability to persist on environmental surfaces for extended periods, MDR *A. baumannii* is highly prone to causing outbreaks in healthcare facilities. Consequently, robust infection prevention and control (IPC) measures, coupled with epidemiological surveillance and microbiological screening, remain fundamental for containment.

In this work, we leveraged both short- and long-read sequencing technologies to analyze 96 genomes, allowing for a high-resolution reconstruction of the outbreak dynamics. The genomic data confirmed that two endemic MDR *A. baumannii* clones served as the source of this major hospital-wide outbreak. These findings underscore how the disruption of standard prevention protocols during the COVID-19 emergency facilitated the unchecked spread of resistant pathogens.

Introduction

Acinetobacter baumannii is a Gram-negative opportunistic bacterium frequently implicated in healthcare-associated infections (HAIs). Although it accounts for a relatively minor proportion of total HAIs (approximately 2%) (1), it is classified within the high-priority ESKAPE group due to its significant virulence and extensive antibiotic resistance profile. This pathogen possesses a remarkable capacity for rapid evolution, readily acquiring resistance determinants against a broad spectrum of antimicrobials, including last-line agents such as carbapenems. The isolation frequency of multidrug-resistant (MDR) strains is rising globally; this trend is particularly alarming in Southern Europe, and specifically in Italy, where carbapenem-resistant *A. baumannii* (CR-Ab) rates reached 80.8% in 2020 (2). Beyond resistance, the clinical concern regarding *A. baumannii* is amplified by its exceptional ability to persist for extended durations in environmental reservoirs-facilitated by biofilm production, even within clinical settings. This environmental resilience promotes transmission and chronic colonization (3), particularly in intensive Care Units (ICUs) (4). Typically, *A. baumannii* infections are device-associated (e.g., ventilator-associated pneumonia and catheter-related bloodstream infections), whereas soft tissue, surgical site, and urinary tract infections occur less frequently (5). Consequently, the World Health Organization has designated MDR *A. baumannii* as a "critical" priority pathogen, highlighting the urgent need for new therapeutic development (6).

The COVID-19 pandemic has been posing an arduous challenge for hospitals worldwide, especially in the early period. The increasing COVID-19-related hospitalizations resulted in critical shortages of staff, personal protective equipment (PPE), and medical devices. These deficits frequently compromised adherence to strict infection prevention and control (IPC) protocols, precipitating a rise in secondary bacterial and fungal co-infections. Indeed, numerous *A. baumannii* outbreaks were documented worldwide during the first pandemic wave (7,8).

The Fondazione IRCCS Policlinico San Matteo (Pavia, Italy), a major ~900-bed teaching hospital in the Lombardy region, served as a pivotal hub for COVID-19 management in Northern Italy. Prior to the pandemic, the institution had established an active genomic surveillance framework and IPC strategies targeting ESKAPE pathogens. These measures effectively maintained *A. baumannii* prevalence at low levels during the 2018-2019 biennium (<1 isolate per 1,000 patient-days). However, coinciding with the first COVID-19 wave, a significant outbreak of MDR *A. baumannii* occurred, with incidence rising to approximately 2.8 isolates per 1,000 patient-days. This study presents a comprehensive genomic characterization of that outbreak event.

Results

Dataset characterisation

During the first COVID-19 wave, an increase in incidence of A. baumannii strains was registered in the study hospital; the cases raised from a baseline of less than one isolate per 1,000 days of hospitalization in the 2018-2019 biennium to 2.8 isolates during the outbreak. This study characterized 96 strains recovered between January and July 2020 from a cohort of 41 patients. The majority of individuals were admitted to ICUs (n=37), with limited cases from pneumology (n=2) and otorhinolaryngology (n=1) wards. Notably, one patient provided three distinct samples: two collected at an external facility in Pavia and one following admission to the emergency department of the study center.

The patient cohort was predominantly male (87.8%; n=36) with a mean age of 61 years (standard deviation [SD] ±10; range: 30-83). The median length of stay (LOS) was 47 days (IQR: 25-74). The mortality rate was 43.9% (n=18), with non-survivors exhibiting a median LOS of 52 days (IQR: 37-85). Regarding SARS-CoV-2 status, 30 patients were positive, seven tested negative, and four were hospitalized prior to the pandemic onset. Clinical isolation sources included lower respiratory tract (33.3%; n=32), bloodstream (22.9%; n=22), and urinary tract (3.1%; n=3). The remaining 40.6% of strains were recovered from screening samples, specifically rectal (n=38) and nasal swabs (n=1). Detailed patient and isolate metadata are provided in Table S1.

All 96 isolates underwent hybrid sequencing utilizing both long- and short-read technologies. This approach yielded 8 complete genomes and 87

high-quality draft genomes (mean size: 4 Mbp; mean contig count: 11.3; Table S2). One isolate (sample 4964_2020) failed quality control and was excluded from downstream analysis. MLST typing confirmed that all 95 retained genomes belonged to Sequence Type 2 (ST2), a member of the globally disseminated International Clone II (IC2) lineage (9).

Phylogenetic analyses

To elucidate the outbreak dynamics, a phylogenetic analysis was performed including the 95 outbreak genomes, 23 isolates from the local hospital surveillance program, and 367 background genomes retrieved from the BV-BRC database. The resulting phylogeny (Figure 1) suggests a polyclonal etiology for the hospital outbreak.

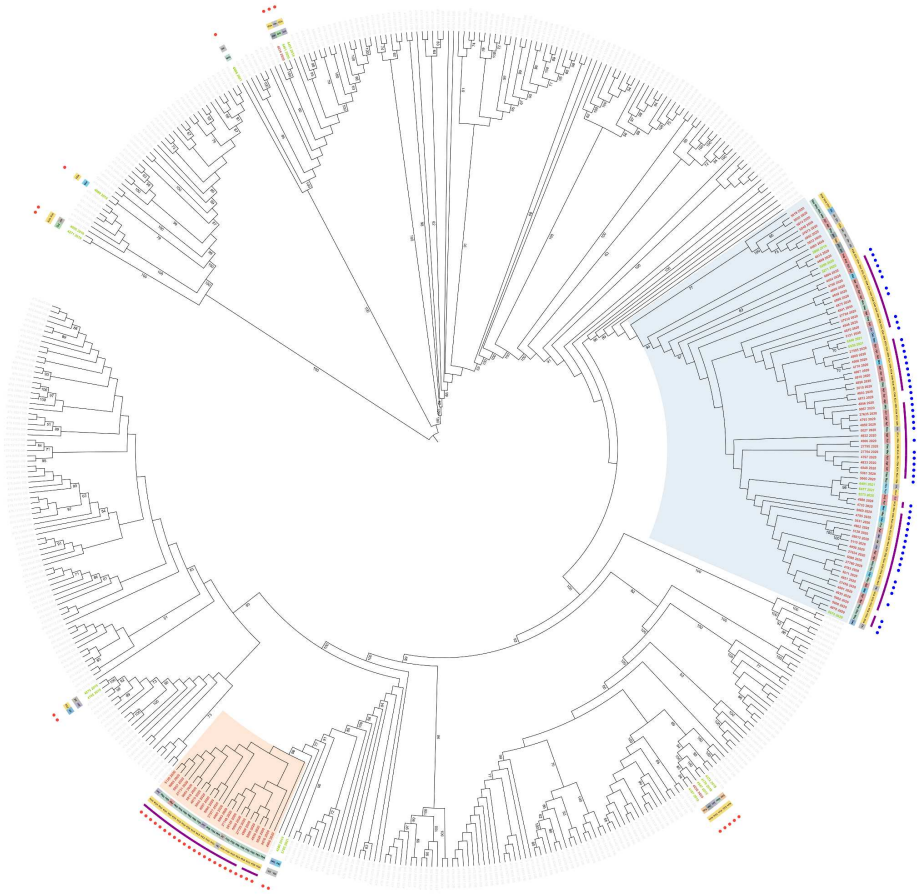


Figure 1 - Maximum-likelihood phylogeny of 485 *A. baumannii* strains. The phylogeny includes 95 outbreak genomes (red labels), 23 surveillance genomes (green labels), and 367 BV-BRC genomes (gray labels) inferred on coreSNPs with RAxML. Date of isolation (innermost crown), ward (second crown), and SARS-CoV-2 positivity (third crown) were retrieved from the hospital and microbiological database. The presence of 100-kbp plasmid (blue dots) and pgaC gene (red dots) were determined by genomic analyses. The two main outbreak clusters are highlighted in blue (cluster 1) and red (cluster 2). Bootstrap values above 50/100 are indicated on tree branches. This figure was obtained using iTol (https://itol.embl.de/).

The vast majority of outbreak isolates (n=93) segregated into two distinct monophyletic clades. The larger group (Cluster 1) contained 71 outbreak strains within a broader clade of 80 genomes, exhibiting a single nucleotide polymorphism (SNP) distance of 0-100. This cluster also encompassed nine local surveillance isolates (one from 2019, four from 2020, and four from 2021) and one external genome from the BV-BRC database.The smaller group (Cluster 2) comprised 23 outbreak strains with an internal SNP distance of 0-43. While Cluster 2 consisted exclusively of isolates from 2020, a surveillance strain from 2019 (isolate 4200_2019) was positioned basally to this group. A background isolate from Belgium (BV-BRC ID 470.13595) was located between the 2019 surveillance strain and Cluster 2, although the branch supporting this relationship showed limited bootstrap support (29/100). The genetic divergence between the two main outbreak clusters (Cluster 1 vs. Cluster 2) ranged from 274 to 356 SNPs.
Interestingly, the genome of one of the two isolates from January 2020 not included in the main clusters (4636_2020) formed a monophyletic group with four 2019 surveillance genomes, while the second one (4614_2020) clustered with one surveillance genome from late 2020.

Antimicrobial resistance profile

Phenotypic testing demonstrated universal carbapenem resistance among the 95 outbreak isolates (100% resistance to Imipenem and Meropenem). Resistance rates were similarly high for aminoglycosides, with 97.9% (n=93) and 98.9% (n=94) of isolates exhibiting non-susceptibility to gentamicin and amikacin, respectively. Conversely, colistin retained full efficacy across the cohort. Regarding tigecycline, only 8.9% (n=7) of tested isolates were resistant, while eight strains were classified as susceptible with increased exposure. Additionally, the totality of isolates tested for ciprofloxacin (n = 90) were resistant.

Genomic predictions showed high concordance with phenotypic data. Carbapenem resistance was genetically substantiated by the ubiquitous presence of $bla_{\text{OXA-23}}$. Although the intrinsic $bla_{\text{OXA-51-like}}$ variant ($bla_{\text{OXA-66}}$) was consistently detected, no upstream insertion of ISAba1 sequences was

identified. Additionally, all strains harbored an AmpC-type enzyme from the *bla*$_{ADC-25}$ family: specifically, the majority carried *bla*$_{ADC-73}$ (n=89), followed by *bla*$_{ADC-56}$ (n=5) and *bla*$_{ADC-30}$ (n=1) (Table S3).

Virulence determinants and biofilm formation

Genomic screening identified several virulence factors, with a marked predominance of determinants linked to biofilm development. The ompA gene, crucial for biofilm stability on abiotic surfaces (10), and the bap gene, essential for robust structural formation (11), were conserved across all isolates. The Chaperone-Usher secretion (CUS) system (encoded by the csuA/BABCDE operon), which mediates initial bacterial adhesion (12), was present in 94 of 95 isolates; the sole exception was strain 4614_2020, isolated in January 2020. Furthermore, all strains possessed the bfmS/bfmR two-component regulatory system, which modulates both CUS expression and ompA (13).Regarding the poly-\beta-(1,6)-N-acetylglucosamine (PNAG) synthesis cluster (pgaABCD), the complete functional operon was identified in 23.3% (n=25) of isolates (12). This group included two early isolates from January 2020 and the entirety of Cluster 2. In contrast, within Cluster 1 genomes (n=70), the pgaC gene was interrupted by an ISAba1 insertion sequence (Table S4).

Phenotypic biofilm assays conducted on outbreak and surveillance strains revealed that the majority were weak producers (n=64), followed by moderate (n=40) and strong producers (n=13), with a single non-producer (Figure S1). Notably, isolates belonging to the two major phylogenetic outbreak clusters exhibited reduced biofilm capacity compared to non-cluster strains. In Cluster 1, phenotypes were distributed as 67.50% weak, 28.75% moderate, and 3.75% strong. Cluster 2 displayed a shift toward higher production (43.48% weak, 47.82% moderate, 8.70% strong). Conversely, sporadic isolates outside these clusters showed the highest capacity (53.33% strong producers).

Plasmidome analysis

Plasmid profiling revealed that all outbreak genomes shared a ubiquitous ~8.7 kbp plasmid (100\% identity to GenBank: NZ_CP008708), carrying 12 coding sequences (CDS), ten of which encode hypothetical proteins. A second, larger plasmid of ~106 kbp (highly homologous to GenBank: CP081146) was detected in the majority of Cluster 1 genomes (51/71). Neither plasmid harbored known resistance or virulence determinants.

Deep characterization of the ~106 kbp plasmid showed that 83 of its 131 CDSs were hypothetical, with 20 associated with phage functions. Structurally, the plasmid was dominated by a substantial prophage region spanning approximately 85 kbp (positions 9261–95595). Phylogenetic mapping (Figure 1) indicates that this large plasmid is exclusive to Cluster 1. Its absence in the basal genomes of the cluster suggests a single acquisition event occurring early in the outbreak, followed by multiple subsequent loss events as indicated by the presence/absence pattern within the lineage. Global population analysis (n=8969 genomes) demonstrated differential distribution of these plasmids: the ~8 kbp plasmid was identified in 32.1% (n=2882) of the general population, whereas the ~106 kbp plasmid was rarer, occurring in only 7% (n=636) (Figure S2).

Discussion

This work presents a significant hospital-wide outbreak of MDR A. baumannii coinciding with the initial wave of the COVID-19 pandemic. The escalation of healthcare-associated infections (HAIs) during pandemic waves was a global phenomenon (14-16), with A. baumannii emerging as a predominant pathogen alongside methicillin-resistant Staphylococcus aureus and fungal infections (17-19). Notable examples include a report by Perez et al. regarding a bla_{OXA-23}-positive CR-Ab outbreak affecting 34 ICUs patients in New Jersey (7), and a study by Gottesman et al. describing an IC2-linked outbreak of bla_{OXA-24}-carrying strains in Israel (8).

Phylogenomic reconstruction identified two major clusters (Figure 1) encompassing the majority of outbreak isolates. By integrating genomic data from the hospital surveillance program spanning pre- and post-outbreak periods, we demonstrated that the outbreak was not caused by a novel introduction. Rather, it resulted from the expansion of lineages already circulating within the facility, specifically linked to ancestral strains (e.g., 3996_2019 and 4200_2019) detected prior to the pandemic.

Plasmidome analysis revealed two distinct elements (~8 kbp and ~100 kbp). The larger ~100 kbp plasmid, which harbors a substantial prophage region with largely uncharacterized genes, was exclusive to a specific subclade within Cluster 1. Its absence in basal strains isolated in 2019 and early 2020 indicates a secondary acquisition event, likely occurring within the hospital environment as the outbreak progressed. This hypothesis is supported by the existence of a separate, plasmid-free subclade (n=8). Moreover, this phage-encoding plasmid is supposedly specific to A. baumannii, as it is not

present in other species (BLASTN on the nucleotide [nt] database performed on 4 October 2022). This observation suggests that the source of acquisition could have been another strain of A. baumannii, not detected by our surveillance program. Despite its fitness advantage not being clear, we can observe a tendency to preserve the plasmid in the genomes analyzed. This trend can be observed both in the local phylogeny (Figure 1) and in the global one (see Figure S2 in the supplemental material).

Metadata mapping indicated that SARS-CoV-2 coinfection rates were balanced between the two phylogenetic clusters, ruling out a direct association between specific A. baumannii lineages and viral status. The observed cross-transmission between COVID-19 positive and negative patients - independent of viral transmission pathways - is likely attributable to A. baumannii's exceptional environmental persistence and biofilm-forming capacity (20). This environmental resilience likely facilitated the pathogen's proliferation during the emergency. Furthermore, no significant correlations were established between phylogenetic clusters and specific wards, clinical outcomes, or patient age.

The phylogenetic tree unveiled eight distinct genomic clusters, four of which spanned the pre- and post-epidemic timeframe. This confirms the presence of at least four endemic clones within the hospital, two of which fueled the outbreak. Remarkably, genomic screening yielded no specific virulence factors known to enhance transmissibility. In fact, biofilm assays revealed that outbreak strains (Clusters 1 and 2) generally produced weaker biofilms compared to sporadic isolates (Figure S1). Specifically, Cluster 1 genomes exhibited an ISAba1-mediated interruption of the pgaC gene, a key component for PNAG synthesis (21) and adhesion (22). Despite this reduced biofilm potential, Cluster 1 was the most expansive lineage.

Collectively, these findings suggest that the outbreak was not driven by intrinsic bacterial hyper-virulence, but rather by the extrinsic pressure of the COVID-19 emergency. Factors such as critical staffing shortages, bed capacity saturation, and the prioritization of healthcare worker protection (e.g., reduced frequency of glove changes) inadvertently compromised standard patient-handling hygiene. This interpretation is corroborated by epidemiological data: following the first wave, A. baumannii incidence regressed to baseline levels (<1 isolate per 1,000 patient-days) and remained controlled during subsequent waves, reflecting improved COVID-19 management protocols.

Materials and methods

Criteria of inclusion

This retrospective investigation was carried out at the IRCCS Fondazione Policlinico San Matteo (Pavia, Italy), encompassing a collection period from January 6, 2020, to July 30, 2020. From an initial pool of 151 *A. baumannii* isolates, a subset of 96 was selected. Selection criteria prioritized the inclusion of paired samples (colonization and infection sites) for each patient, where available, and of multiple samples in case of long hospitalizations (1 sample/7 days/material). Species identification was confirmed via matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) utilizing a Bruker Daltonik system (Bremen, Germany) and Bruker Biotyper 3.1 software. Antimicrobial susceptibility profiling was conducted using a dual approach: the Sensititre Gram-negative EUMDROXF plate (Thermo Fisher Scientific, Rodano, Italy) and the Phoenix NMIC-402 panel on the BD Phoenix M50 platform. Minimum Inhibitory Concentration (MIC) values were interpreted in accordance with European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints (23), with the exception of tigecycline, for which Clinical and Laboratory Standards Institute (CLSI) criteria were applied (24). Pertinent clinical data were extracted from hospital electronic health records. Patients were classified as COVID-19 positive based on real-time reverse transcriptase PCR (RT-PCR) detection of SARS-CoV-2 RNA at the time of admission.

Genomic DNA extraction and sequencing

For the 96 selected strains, genomic DNA was isolated using distinct protocols based on the sequencing platform: the DNeasy Blood and Tissue Kit (Qiagen) was employed for short-read Illumina sequencing, while the MagAttract High-Molecular-Weight (HMW) DNA Kit (Qiagen) was utilized for long-read Nanopore sequencing. Hybrid assembly of short and long reads was executed using Unicycler (25). *In silico* multilocus sequence typing (MLST) was performed on the assembled genomes using an *in-house* script based on the Pasteur scheme (26). The resistome and virulome were characterized by screening genomes with ABRicate (https://github.com/tseemann/abricate) against the NCBI AMRFinderPlus (27), CARD (28), ResFinder (29), and VirulenceFinder (30) databases. Automated outputs underwent manual curation and validation via BLAST and raw long-read alignment. Insertion sequences were identified using MobileElementFinder (31).

Plasmidome characterisation

Initial plasmid screening was conducted using PlasmidFinder (32). A comprehensive structural analysis was subsequently performed on the least fragmented assemblies using Bandage (33) to visualize assembly graphs, verifying circularity and differential depth coverage. Once reference plasmid sequences were established, their presence was systematically assessed in the remaining dataset via BLASTN. Functional annotation of plasmid sequences integrated results from Prokka (34) and the NCBI Conserved Domain Database (35), supplemented by BLASTP searches against the non-redundant (nr) database where necessary. Prophage regions were identified and annotated using PHASTER (36). To evaluate the prevalence of the identified plasmids within the broader *A. baumannii* population, a comprehensive genomic collection was curated. This dataset combined the study genomes with all high-quality *A. baumannii* genomes available in the BV-BRC database (retrieved July 2022 via the P-DOR pipeline's makepdordb.py script [https://github.com/SteMIDIfactory/P-DOR]). The entire collection was screened for plasmid gene content; a genome was classified as plasmid-positive if it harbored the majority of the plasmid-specific coding sequences (≥11 genes for the ~8 kbp; ≥120 genes for the ~100 kbp). Global phylogeny was inferred using FastTree (37) to map plasmid distribution across the species population structure.

Hospital surveillance program

The study incorporated additional genomes processed under the hospital's routine genomic surveillance program. Standard workflows for these isolates involved DNA extraction via DNeasy Blood and Tissue Kit (Qiagen), Illumina sequencing, and assembly using Unicycler (25) with default settings. Specifically, all ST2 genomes (n=23) isolated in the timeframe flanking the outbreak (12 from 2019, 5 from 2020, and 6 from 2021) were retrieved and integrated into the study dataset.

Biofilm production assay

The biofilm-forming capacity of all *A. baumannii* isolates (n=119) was evaluated following the protocol established by Kumari et al. (38).
Pseudomonas aeruginosa ATCC 27853 and a characterized strong biofilm-producing *Staphylococcus epidermidis* strain (30678, internal collection) served as positive controls, with sterile media acting as the negative control. Optical density (OD) readings were interpreted as described previously (39). In detail, the cutoff value (c) was calculated as

three standard deviations (SDs) above the average optical density (OD) of the negative control. Strains were classified into the following four categories based upon mean OD values: “no producer” when mean OD is ≤c, “weak producer” when c is less than the mean OD, which is ≤2c, “moderate producer” when 2c is less than the mean OD, which is ≤4c, and “strong producer” when mean OD is >4c.

Phylogenetic analyses

To contextualize the outbreak, a genomic background dataset was constructed by retrieving the 20 high-quality genomes from the BV-BRC database (https://www.bv-brc.org) that exhibited the highest k-mer similarity (calculated via MASH (40)) to the study isolates. The P-DOR pipeline (version beta1.5; https://github.com/SteMIDIfactory/P-DOR) was used to align the complete dataset against an internal reference genome (strain 4946) and extract core single nucleotide polymorphisms (SNPs). Maximum-likelihood phylogeny was inferred from the core SNP alignment using RAxML (41) with 100 bootstrap replicates. The General Time Reversible (GTR) substitution model with ascertainment bias correction (43) was applied, as determined by ModelTest-NG (42).

Acknowledgment

We acknowledge Anna K. Pöntinen and Francois Cleon for experimental assistance and the Genomics Support Centre Tromsø, UiT The Arctic University of Norway, for the short- and long-read sequencing.
J.C. was supported by ERC grant 742158. S.G. was supported by Ricerca Corrente grant 924-rcr2018i-34 (Italian Ministry of Health).

Data availability

Assembled genomes of outbreak and surveillance *A. baumannii* isolates are deposited in the National Center for Biotechnology Information (NCBI) under the BioProject accession no. PRJNA844013.Additional material is present in the “Supplemental material” of the manuscript.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (internal project code, 08022298/13; project number 733-rcr2013-34). The work described herein is a retrospective study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

References

1. Lob SH, Hoban DJ, Sahm DF, Badal RE. 2016. Regional differences and trends in antimicrobial susceptibility of Acinetobacter baumannii. Int J Antimicrob Agents 47:317–323. Crossref. PubMed. Web of Science.
2. Yang C-H. 2022. WHO regional office for Europe/European centre for disease prevention and control. antimicrobial resistance surveillance in Europe 2022 to 2020 data. WHO Regional Office for Europe, Copenhagen, Denmark.
3. Yang C-H, Su P-W, Moi S-H, Chuang L-Y. 2019. Biofilm formation in Acinetobacter baumannii: genotype-phenotype correlation. Molecules 24:1849. Crossref. PubMed. Web of Science.
4. Antunes LCS, Visca P, Towner KJ. 2014. Acinetobacter baumannii: evolution of a global pathogen. Pathog Dis 71:292–301. Crossref. PubMed. Web of Science.
5. Wong D, Nielsen TB, Bonomo RA, Pantapalangkoor P, Luna B, Spellberg B. 2017. Clinical and pathophysiological overview of

Acinetobacter infections: a century of challenges. Clin Microbiol Rev 30:409–447. Crossref. PubMed. Web of Science.

6. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outterson K, Patel J, Cavaleri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18:318–327. Crossref. PubMed. Web of Science.
7. Perez S, Innes GK, Walters MS, Mehr J, Arias J, Greeley R, Chew D. 2020. Increase in hospital-acquired carbapenem-resistant Acinetobacter baumannii infection and colonization in an acute care hospital during a surge in COVID-19 admissions—New Jersey, February-July 2020. MMWR Morb Mortal Wkly Rep 69:1827–1831. Crossref. PubMed. Web of Science.
8. Gottesman T, Fedorowsky R, Yerushalmi R, Lellouche J, Nutman A. 2021. An outbreak of carbapenem-resistant Acinetobacter baumannii in a COVID-19 dedicated hospital. Infect Prev Pract 3:100113. Crossref. PubMed. Web of Science.
9. Dijkshoorn L, Aucken H, Gerner-Smidt P, Janssen P, Kaufmann ME, Garaizar J, Ursing J, Pitt TL. 1996. Comparison of outbreak and nonoutbreak Acinetobacter baumannii strains by genotypic and phenotypic methods. J Clin Microbiol 34:1519–1525. Crossref. PubMed. Web of Science.
10. Gaddy JA, Tomaras AP, Actis LA. 2009. The Acinetobacter baumannii 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. Infect Immun 77:3150–3160. Crossref. PubMed. Web of Science.
11. Azizi O, Shahcheraghi F, Salimizand H, Modarresi F, Shakibaie MR, Mansouri S, Ramazanzadeh R, Badmasti F, Nikbin V. 2016. Molecular analysis and expression of bap gene in biofilm-forming multi-drug-resistant Acinetobacter baumannii. Rep Biochem Mol Biol 5:62–72. PubMed.
12. Tomaras AP, Dorsey CW, Edelmann RE, Actis LA. 2003. Attachment to and biofilm formation on abiotic surfaces by Acinetobacter baumannii: involvement of a novel chaperone-usher pili assembly system. Microbiology (Reading) 149:3473–3484. Crossref. PubMed. Web of Science.
13. Kishii K, Hamada M, Aoki K, Ito K, Onodera J, Ishii Y, Tateda K. 2020. Differences in biofilm formation and transcription of biofilm-associated genes among Acinetobacter baumannii clinical strains belonging to the international clone II lineage. J Infect Chemother 26:693–698. Crossref. PubMed.

14. Garcia-Vidal C, Sanjuan G, Moreno-García E, Puerta-Alcalde P, Garcia-Pouton N, Chumbita M, Fernandez-Pittol M, Pitart C, Inciarte A, Bodro M, Morata L, Ambrosioni J, Grafia I, Meira F, Macaya I, Cardozo C, Casals C, Tellez A, Castro P, Marco F, García F, Mensa J, Martínez JA, Soriano A, COVID-19 Researchers Group. 2021. Incidence of co-infections and superinfections in hospitalized patients with COVID-19: a retrospective cohort study. Clin Microbiol Infect 27:83–88. Crossref. PubMed. Web of Science.
15. Cusumano JA, Dupper AC, Malik Y, Gavioli EM, Banga J, Berbel Caban A, Nadkarni D, Obla A, Vasa CV, Mazo D, Altman DR. 2020. Staphylococcus aureus bacteremia in patients infected with COVID-19: a case series. Open Forum Infect Dis 7:ofaa518. Crossref. PubMed.
16. Contou D, Claudinon A, Pajot O, Micaëlo M, Longuet Flandre P, Dubert M, Cally R, Logre E, Fraissé M, Mentec H, Plantefève G. 2020. Bacterial and viral co-infections in patients with severe SARS-CoV-2 pneumonia admitted to a French ICU. Ann Intensive Care 10:119. Crossref. PubMed. Web of Science.
17. Pascale R, Bussini L, Gaibani P, Bovo F, Fornaro G, Lombardo D, Ambretti S, Pensalfine G, Appolloni L, Bartoletti M, Tedeschi S, Tumietto F, Lewis R, Viale P, Giannella M. 2022. Carbapenem-resistant bacteria in an intensive care unit during the coronavirus disease 2019 (COVID-19) pandemic: a multicenter before-and-after cross-sectional study. Infect Control Hosp Epidemiol 43:461–466. Crossref. PubMed.
18. O'Toole RF. 2021. The interface between COVID-19 and bacterial healthcare-associated infections. Clin Microbiol Infect 27:1772–1776. Crossref. PubMed. Web of Science.
19. Kundu R, Singla N. 2022. COVID-19 and plethora of fungal infections. Curr Fungal Infect Rep 16:47–54. Crossref. PubMed.
20. Eze EC, Chenia HY, El Zowalaty ME. 2018. Acinetobacter baumannii biofilms: effects of physicochemical factors, virulence, antibiotic resistance determinants, gene regulation, and future antimicrobial treatments. Infect Drug Resist 11:2277–2299. Crossref. PubMed. Web of Science.
21. Choi AHK, Slamti L, Avci FY, Pier GB, Maira-Litrán T. 2009. The pgaABCD locus of Acinetobacter baumannii encodes the production of poly-beta-1–6-N-acetylglucosamine, which is critical for biofilm formation. J Bacteriol 191:5953–5963. Crossref. PubMed. Web of Science.
22. Chen K-M, Chiang M-K, Wang M, Ho H-C, Lu M-C, Lai Y-C. 2014. The role of pgaC in Klebsiella pneumoniae virulence and biofilm formation. Microb Pathog 77:89–99. Crossref. PubMed. Web of Science.

23. The European Committee on Antimicrobial Susceptibility Testing. 2020. Breakpoint tables for interpretation of MICs and zone diameters. version 10.0.
24. Clinical and Laboratory Standards Institute. 2020. Performance Standards for Antimicrobial Susceptibility Testing. CLSI, CLSI supplement M100, 30th ed, Wayne, PA.
25. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. Crossref. PubMed. Web of Science.
26. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S. 2010. The population structure of Acinetobacter baumannii: expanding multiresistant clones from an ancestral susceptible genetic pool. PLoS One 5:e10034. Crossref. PubMed. Web of Science.
27. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 63:e00483-19. Crossref. PubMed. Web of Science.
28. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 45:D566–D573. Crossref. PubMed. Web of Science.
29. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. Crossref. PubMed. Web of Science.
30. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510. Crossref. PubMed. Web of Science.
31. Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. 2021. Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. J Antimicrob Chemother 76:101–109. Crossref. PubMed. Web of Science.
32. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus

sequence typing. Antimicrob Agents Chemother 58:3895–3903. Crossref. PubMed. Web of Science.
33. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. Crossref. PubMed.
34. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. Crossref. PubMed. Web of Science.
35. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39:D225–D229. Crossref. PubMed. Web of Science.
36. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–21. Crossref. PubMed. Web of Science.
37. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. Crossref. PubMed. Web of Science.
38. Kumari M, Bhattarai NR, Rai K, Pandit TK, Khanal B. 2022. Multidrug-resistant Acinetobacter: detection of ESBL, MBL, blaNDM-1 genotype, and biofilm formation at a tertiary care hospital in eastern Nepal. Int J Microbiol 2022:8168000. Crossref. PubMed.
39. Stepanović S, Vuković D, Hola V, Di Bonaventura G, Djukić S, Cirković I, Ruzicka F. 2007. Quantification of biofilm in microtiter plates: overview of testing conditions and practical recommendations for assessment of biofilm production by staphylococci. APMIS 115:891–899. Crossref. PubMed. Web of Science.
40. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17:132. Crossref. PubMed. Web of Science.
41. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. Crossref. PubMed. Web of Science.
42. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291–294. Crossref. PubMed. Web of Science.
43. Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925. Crossref. PubMed. Web of Science.

CHAPTER 7:

Time is of the essence: a new 24/7 laboratory organization and a fast-track diagnostic pathway for blood culture of critical patients

Corbella Marta [1*], **Petazzoni Greta** [1,2*], Seminari Elena[3], Merla Cristina[1], De Vitis Debora[1], Iskandar Elizabeth[1], Muzzi Alba[4], Rettani Marco[4], Raffaele Bruno[2,3], Baldanti Fausto[1,2], Cambieri Patrizia[1] on behalf of the San Matteo Pavia Microbiology and Virology Working Group

1. Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
2. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia. Pavia, Italy
3. Infectious Diseases, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
4. Fondazione IRCCS Policlinico San Matteo. Pavia, Italy

* These two authors contributed equally

The manuscript has been **submitted at Journal of Clinical Microbiology,** March 2026

In this work I contributed to the study design and performed all statistical analyses. I also contributed substantially to drafting and revising the manuscript.

Abstract

Bloodstream infections are among the most severe infectious diseases, with mortality rates of up to 25%. Delays as short as one hour in diagnosis or initiation of appropriate antimicrobial therapy can significantly worsen patient outcomes.

This retrospective study, conducted in a 900-bed Italian hospital from January 2019 to December 2024, evaluates the impact of a 24/7 reorganization of the Clinical Microbiology Laboratory. This reorganization included a night shift to ensure around-the-clock processing and the introduction of a fast-track diagnostic pathway for blood cultures from critically ill patients (designated as "urgent blood cultures"). We analyzed the impact on turnaround times (TATs) for Gram stain, microorganism identification, and resistance marker detection.

A total of 194,171 blood cultures were processed. Following the implementation of the 24/7 model, median Gram stain TAT decreased from 4.46 to 1.40 hours, microorganism identification TAT from 5.75 to 2.35 hours, and resistance marker TAT from 6.97 to 2.68 hours. Significant reductions were observed, particularly during night shifts. Urgent blood cultures yielded a higher positivity rate (16.22% vs. 13.04%) and included the isolation of time-critical bacteria capable of causing meningitis, such as *Streptococcus pneumoniae*.

A continuous, prioritized blood culture diagnostic system significantly reduced reporting times, particularly overnight. This model enhances early sepsis management, supports antimicrobial stewardship, and demonstrates the value of tailored diagnostic pathways for critically ill patients.

Introduction

Bloodstream infections (BSIs) represent one of the most serious infectious diseases, with associated mortality rates ranging from 11% to 25% (1). Even a one-hour delay in reporting results and subsequently in initiating or adjusting antibiotic therapy might adversely affect patient survival (1,2). Sepsis guidelines advocate the initiation of antimicrobial therapy within one hour of patient admission in case of septic shock, and within three hours for possible sepsis without shock (3). Early identification of critically ill patients with septic shock and/or at high risk of developing septic shock is essential, and guidelines (3) recommend the use of early warning score screening systems, such as MEWS, NEWS, NEWS-2, and qSOFA (4). Therefore, uninterrupted blood culture processing, performed 24 hours a day, 7 days a week, is critical to ensure that no delay in pathogen detection or reporting compromises patient outcomes.

Blood cultures (BCs) remain the gold standard for microbiological diagnosis of BSIs (5-6). Timely and accurate identification of the causative bacterial species, along with detection of key resistance mechanisms, is essential for improving both patient outcomes and the appropriate use of antimicrobials (1). The use of continuous-monitoring blood culture systems, combined with advances in rapid diagnostic tests, has significantly shortened the time required to obtain results (6). Indeed, the overall turnaround time for positive samples relies on the efficiency of laboratory operations. Specifically, the delay between BCs bottle positivity and the final microbiological diagnosis is influenced by timely and appropriate sample processing (6,7). This highlights the necessity of round-the-clock BC processing, since any interruption in workflow, even overnight or during weekends, can extend turnaround times, delaying critical therapeutic decisions.

Since 2012, a project aimed at improving sepsis management has been implemented at our hospital, Fondazione IRCCS Policlinico San Matteo in Pavia (Italy), with progressive steps introduced over the years (8,9). In 2015, BC incubators were placed in the Emergency Room (E.R.), in the Infectious Diseases ward, and in the Intensive Care Units (ICUs) to reduce pre-incubation time (8,9). Since 2019, an alert system to notify Infectious Diseases consultants has enabled real-time consultation of BC preliminary results, improving Antimicrobial Stewardship programs. In 2020, a fast-track system was also introduced for BCs collected from critically ill patients with suspected sepsis and/or septic shock. These samples were flagged with a specific code: Urgent Blood Cultures (Urg-BC) to differentiate and prioritize them for rapid processing. BCs were therefore managed 24 hours a day, 7 days a week, through this organizational model specifically designed to provide a fast and preferential route for positive samples identified as ‘urgent’. This system was sustained by a dedicated team working on a structured shift schedule, including night coverage, to ensure an

uninterrupted around-the-clock workflow and real-time response at any time of day or week.

Based on these organizational interventions, the present study aimed to evaluate turnaround times (TATs) for Gram staining, microorganism identification, and resistance marker detection in BCs over a six-year period, following the implementation of a 24/7 laboratory organization and a fast-track, preferential diagnostic pathway dedicated to BCs from critically ill patients.

Materials and methods

The Fondazione IRCCS Policlinico San Matteo (Pavia, Italy) is a 900-bed tertiary-care teaching hospital recording approximately 28,000 hospitalizations per year. Within the hospital, when sepsis or bacteremia is suspected in adult patients, two or three sets of blood culture bottles (BCbs) are collected, as recommended by CLSI guidelines (2,10). Each specimen is required to contain at least 10 mL of blood, resulting in a total recommended volume of 40–60 mL.

Definitions

- *Blood culture* (BC): Two or three sets of blood culture bottles (BCbs), including both aerobic and anaerobic bottles, collected from the same patient on the same day.
- *Urgent blood culture* (Urg-BC): BCbs collected from patients, identified as critical through MEWS>=3 or NEWS-2>=5 (11,12). Software connected to the Laboratory Information System (LIS) stored this information and alerted the personnel on duty through an automated phone call upon BC positivity.
- *Non urgent blood culture* (Non Urg-BC): BCbs collected from non-critical patients.
- *Time to detection* (TTD): the time elapsed between BCb incubation and the positivity.
- *First encounter*: the BCb with the shortest TTD among the BCbs collected from each patient within the same day, considering the Gram stain result.

- *Gram stain turnaround time* (Gram-TAT): the time elapsed between BCb positivity and the Gram staining result.
- *Microorganism identification turnaround time* (ID-TAT): the time elapsed between BCb positivity and microorganism identification.
- *Resistance markers turnaround time* (Marker-TAT): the time elapsed between BCb positivity and the identification of antibiotic resistance markers.
- *BCb contamination*: Coagulase-negative staphylococci, aerobic and anaerobic diphtheroids, *Micrococcus spp.*, *Bacillus spp.*, and viridans group streptococci were considered contaminants when a single BCb was positive. The isolate was reported as a probable contaminant, and susceptibility testing was not performed (2,13).

Study design

This retrospective study considered BCbs collected from January 1, 2019, to December 31, 2024, excluding those from the first COVID-19 pandemic wave (January 1 to April 30, 2020), as routine laboratory workflows were significantly disrupted during this period. Starting from January 2020, a fast-track and preferential diagnostic pathway dedicated to Urg-BC was introduced. In the same period, all personnel, including those from the Virology, Mycology, and Parasitology Units, were trained to handle positive BCbs during both day and night shifts. Throughout the entire study period, BCbs were incubated 24/7. Positive BCbs were processed during the operating hours of the Clinical Microbiology Laboratory (CML), which changed over the course of the years as detailed in Figure 1 and below.

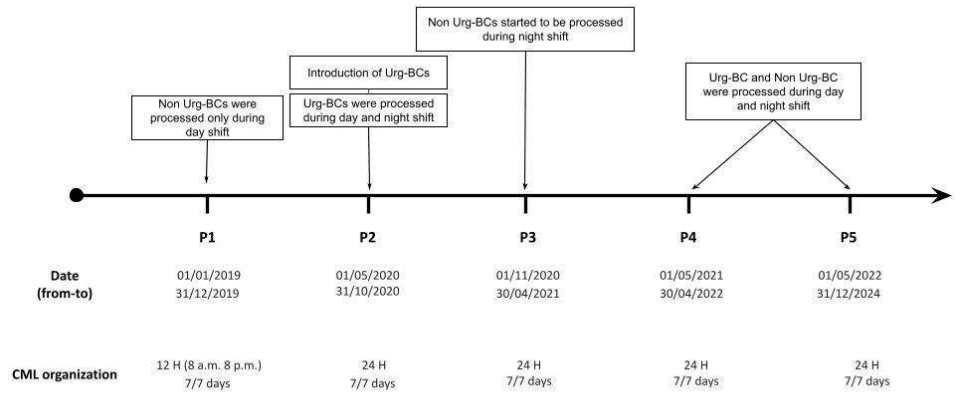


Figure 1 - Timeline of CML organisation. Schematic timeline of CML organisation in relation to study periods, from period 1 (P1) to period 5 (P5).

- **Period 1 (P1)**: from 01-01-2019 to 31-12-2019. The baseline period. The CML operated 7 days a week from 8 am to 8 pm.
- **Period 2 (P2)**: from 01-05-2020 to 31-10-2020. Introduction of night shifts in CML organization and Urg-BC. The CML operated 24/7. All BCbs were processed during the daytime (8 a.m. to 8 p.m.), while Urg-BCbs were processed also during the night shift (8 p.m. to 8 a.m.).
- **Period 3 (P3)**: from 01-11-2020 to 30-04-2021. BCb processing activity was implemented during the night shift. The CML operated 24/7. All BCbs were processed during the daytime; Urg-BCbs were also processed during the night shift. Non Urg-BCb processing was gradually introduced during the night shift as personnel training was completed.
- **Period 4 (P4)**: from 01-05-2021 to 30-04-2022. Integration of complete activity into the overall process of sepsis diagnosis improvement. All personnel completed training, enabling all BCbs (Urg-BCbs and Non Urg-BCbs) to be handled equally during both day and night shifts.
- **Period 5 (P5)**: from 01-05-2022 to 31-12-2024. Validation period. In this period, we evaluated whether the results of the previous periods were maintained over time.

Laboratory workflow

Once collected, BCbs were incubated within 2 hours of collection in the BACTEC FX (Becton Dickinson and Company, Franklin Lakes, New Jersey, United States) located in the CML or in the BACTEC FX 40 located in the E.R., Infectious Diseases ward, and ICUs. When a BCb flagged positive, Gram staining was immediately performed (14), the result was recorded in the LIS and promptly communicated by phone to the ward physician. Concurrently, the BCb was plated on appropriate agar media, as indicated by internal procedures.

Rapid identification was guided by the Gram stain result, in accordance with laboratory protocols. For Gram-negative bacilli (GNB), Gram-positive bacilli (GPB), and Gram-positive cocci in chains (GPCC), identification was achieved by MALDI-ToF mass spectrometry (Bruker Daltonik, Bremen, Germany). For Gram-positive cocci (GPC), the Coagulase test was used to identify *Staphylococcus aureus*. When polymicrobial flora or yeasts were observed, a multiplex PCR syndromic panel (FilmArray Blood Culture-2

Panel, BioFire Diagnostics LLC, Salt Lake City, UT, USA) was used to identify the microorganisms involved and their resistance markers.

Following rapid identification, resistance markers in *Enterobacterales* were assessed using either the molecular Xpert® CARBA-R assay (Cepheid, USA) or an immunochromatographic test for carbapenemase production (NG-Test CARBA5, NG Biotech Laboratoire, Guipry, France). For the detection of CTX-M–type extended-spectrum β-lactamases (ESBLs), the NG-Test® CTX-M Multi (NG Biotech Laboratoire, Guipry, France) was used. For *Enterococci*, Molecular Xpert® vanA/vanB (Cepheid, USA) was used to detect vancomycin resistance. For *Staphylococci*, methicillin resistance was evaluated using the molecular MRSA/SA Blood Culture assay (Cepheid, USA).

Data processing and statistical analysis

All data used in this study were extracted from the LIS, which records all BCbs submitted to the CML. Both positive and negative BCbs were included in the analysis. Only the first-encounter BCb was considered in the analysis, and BCbs identified as contamination according to CLSI guidelines were excluded (2,13).

Continuous variables (Gram-TAT, ID-TAT, Marker-TAT) were reported as medians with interquartile ranges (Q1–Q3). Comparisons of variable distributions across different study periods were performed using the Kruskal-Wallis test, followed by Dunn’s post hoc test with Bonferroni correction. To assess whether there was a monotonic decreasing trend in Gram-TAT and ID-TAT across the five consecutive periods, we applied Kendall’s rank correlation test (alternative hypothesis: "less") with periods treated as ordered time points. Within each period, the comparison of Non Urg-BC and Urg-BC regarding continuous variables was performed using the Wilcoxon rank-sum test.

Categorical variables were presented as absolute counts and percentages. Proportions were compared using the chi-squared test, or the binomial test when only two categories were present. To assess whether the proportions of each Gram stain result varied across periods, we performed a proportion test. The significance level was set at 0.05.

Results

From January 1, 2019, to December 31, 2024 (excluding the COVID-19 pandemic period from January 1 to April 30, 2020), a total of 194,171 BCbs were processed, collected from 23,508 hospitalized patients. Of these, 177,224 (91.27%) were classified as Non Urg-BCbs, and 16,947 (8.73%) as Urg-BCbs. Regarding collection time, 115,131 (59.29%) BCbs were collected during the daytime, while the remaining 79,040 (40.70%) were collected during night shifts.

BCbs were mainly collected in the ICUs and the E.R., accounting for 15.36% and 15.29% of the total, respectively, followed by the Hematology Unit (12.63%) and the Infectious diseases ward (9.66%) (Figure 2). Urg-BCbs were primarily collected from patients admitted to ICUs (34.44%) and the ER (32.22%), while the remaining 33.44% originated from other wards.

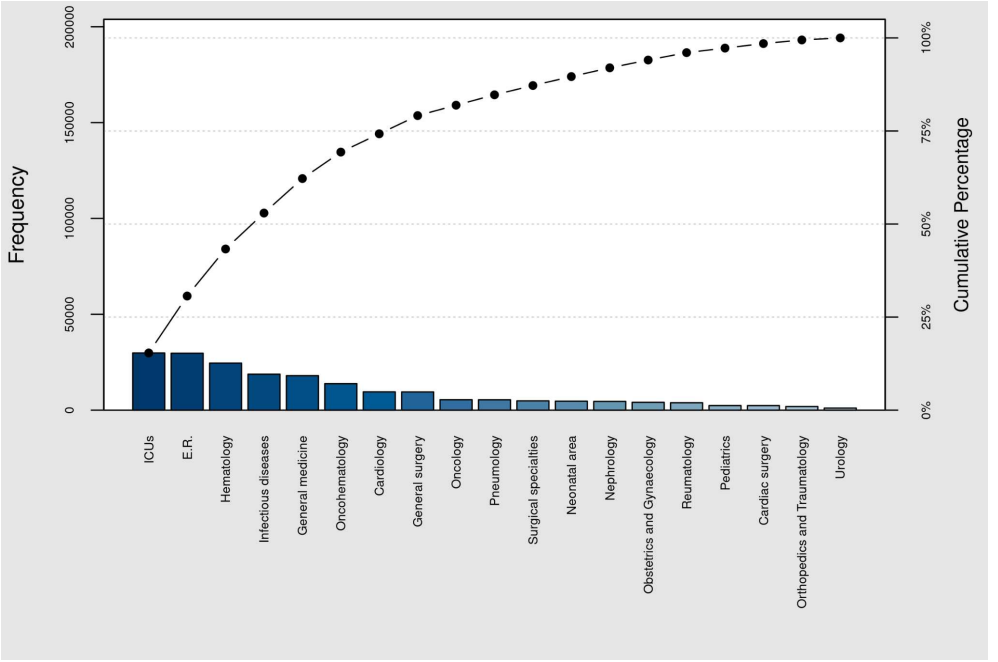


Figure 2 - **Requested BCbs.** Overall pareto chart of requested BCbs from 1st January 2019 to 31st December 2024 in different hospital wards. COVID-19 period from 1st January 2020 to 30th April 2020 was excluded.

During the study period, 25,858 (13.31%) BCbs yielded positive results. Specifically, positivity was observed in 23,109 Non Urg-BCbs and 2,749 Urg-BCbs, corresponding to positivity rates of 13.04% and 16.22%,

respectively (p<0.001). The median TTD for positive bottles was 14.09 hours [IQR: 10.29–19.38] while the overall contamination rate was 5.53%.

Analysis of "first encounters" included 10,340 positive episodes (9,086 Non Urg-BCbs and 1,254 Urg-BCbs). The proportion of BCbs flagging positive during day versus night shifts remained comparable across the five study periods, with no significant differences observed. The median proportion of positive signals detected during the night was 54.56% [53.36–56.58], compared to 45.35% [44.50–45.71] during the day.

Gram stain turnaround time (Gram-TAT)

In 2019 (P1), when the CML operated 12 hours per day, the median time from BCb positivity to Gram stain result (Gram-TAT) was 4.46 hours [1.79-8.27]. However, following the introduction of the night shift and the transition to a 24/7 operational model (P2), the overall median Gram-TAT decreased significantly. Specifically, it declined to 2.43 hours [1.40-4.18] in P3, and further to 1.40 hours [0.94-2.38] in P5 (Figure 3A and Table 1). This decreasing trend in Gram-TAT was statistically significant over the study period (p<0.001). Notably, for BCbs flagging positive during night shifts, the median Gram-TAT was 7.51 hours [5.06-10.06] in P1 and decreased to 1.74 hours [1.06-3.06] in P5 (p< 0.001).
Regarding Non Urg-BCbs, the median Gram-TAT decreased from 4.46 hours in P1 to 2.50 hours in P3 (p<0.001). As shown in Figure 3B, the Gram-TAT continued to decrease significantly in P4 and P5, reaching 1.43 hours in the validation period (P5). The Gram-TAT for Non Urg-BCbs flagging positive during the night decreased from 7.51 hours [5.06-10.06] in P1 to 1.81 hours [1.08-3.17] in P5, showing a statistically significant downward trend (p<0.001), as illustrated in Figure 3D.
For Urg-BCbs (analyzed from P2 onwards), the overall median Gram-TAT was 1.44 hours [0.97-2.39], which was significantly lower than the 1.73 hours [1.05-3.18] observed for Non Urg-BCbs (p<0.001). In P2, the Gram-TAT was 2.04 hours [1.24-3.08], decreasing to 1.27 hours [0.89-1.98] in the validation period P5 (p<0.001; Figure 3C). Specifically, the Gram-TAT for Urg-BCbs flagging positive during the night decreased from 2.37 hours [1.72-3.45] in P2 to 1.48 hours [1.00-2.29] in P5, with a statistically significant decreasing trend (p<0.001; Figure 3E).

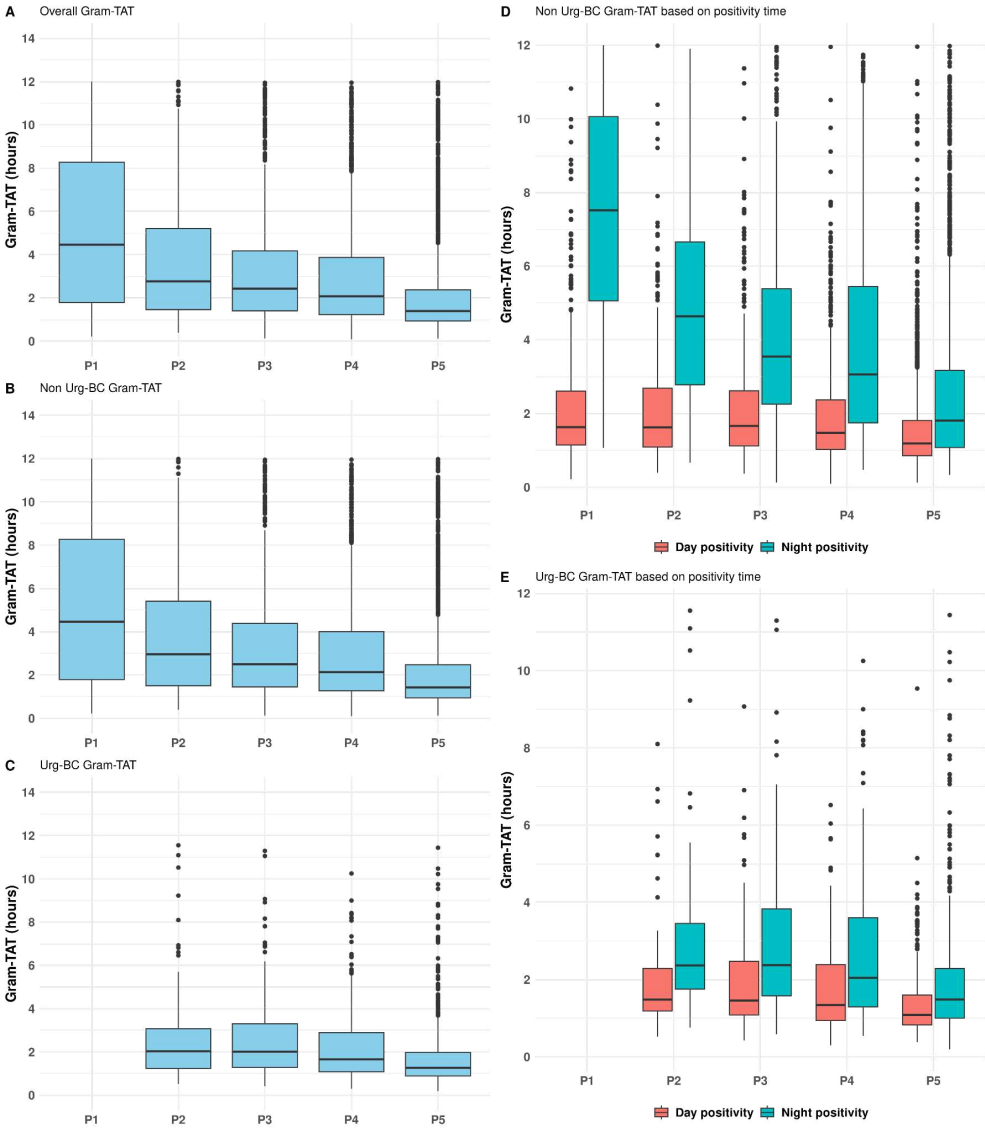


Figure 3 - Distribution of Gram-stain TAT in the five periods. Gram-stain TAT were calculated for (A) overall BCbs, (B) Non Urg-BCbs, (C) Urg-BCbs, (D) Non Urg-BCbs based on positivity time, and (E) Urg-BCbs based on positivity time.

	Gram-TAT			ID-TAT			MARKER-TAT		
Period	Non Urg-BC	Urg-BC	Overall	Non Urg-BC	Urg-BC	Overall	Non Urg-BC	Urg-BC	Overall
P1	4.46 [1.79-8.27]		4.46 [1.79-8.27]	5.75 [3.04-9.59]		5.75 [3.04-9.59]	6.97 [3.63-11.05]		6.97 [3.63-11.05]
P2	2.96 [1.51-5.41]	2.04 [1.24-3.08]	2.77 [1.46-5.21]	4.05 [2.54-6.77]	2.72 [2.15-4.42]	3.91 [2.47-6.61]	5.71 [3.39-7.73]	4.91 [2.30-7.26]	5.59 [3.35-7.73]
P3	2.5 [1.45-4.38]	2.01 [1.28-3.30]	2.43 [1.40-4.18]	3.49 [2.43-5.50]	3.06 [2.57-5.08]	3.41 [2.44-5.50]	3.69 [2.56-6.23]	2.96 [2.33-5.39]	3.61 [2.45-5.84]
P4	2.13 [1.27-4.00]	1.66 [1.09-2.89]	2.08 [1.24-3.87]	2.86 [2.12-4.89]	2.76 [2.07-4.41]	2.85 [2.11-4.83]	4.03 [2.48-6.49]	3.05 [2.43-3.94]	3.70 [2.44-6.15]
P5	1.43 [0.94-2.48]	1.27 [0.89-1.98]	1.40 [0.94-2.38]	2.37 [1.77-3.53]	2.21 [1.60-3.18]	2.35 [1.76-3.42]	2.71 [1.98-4.07]	2.51 [1.79-3.61]	2.67 [1.94-3.97]
Significant trend	↓ (p<0.001)	↓ (p<0.001)	↓ (p<0.001)	↓ (p<0.001)	↓ (p<0.001)	↓ (p<0.001)	↓ (p<0.001)	↓ (p=0.03)	↓ (p<0.001)

Table 1 - TATs summary. Summary of time for Gram-TAT, ID-TAT, and Marker-TAT for overall and type of blood culture (Non Urg-BC and Urg-BC). Median time [Q1-Q3] is indicated in hours.

Among all positive BCbs, Gram-positive cocci (GPC) were the most prevalent, accounting for 39.90% (including contaminants), followed by Gram-negative bacilli (GNB; 36.32%), Gram-positive cocci in chains (GPCC; 11.88%), fungi (7.08%), and polymicrobial flora (2.15%). Across the five study periods, no statistically significant differences were observed in the proportions of GPC, fungi, or polymicrobial flora.

When comparing proportions, a statistically significant difference was found in the prevalence of GPC, which accounted for 40.37% in Non Urg-BCbs compared to 37.08% of Urg-BCbs (p=0.03). No significant differences were observed for the other categories:

- GNB: 35.99% in Non Urg-BCbs vs. 38.68% in Urg-BCbs (p>0.05).
- GPCC: 11.89% in Non Urg-BCbs vs. 11.80% in Urg-BCbs (p>0.05).
- Fungi: 6.99% in Non Urg-BCbs vs. 7.74% in Urg-BCbs (p>0.05).
- Polymicrobial flora: 2.18% in Non Urg-BCbs vs. 1.91% in Urg-BCbs (p>0.05).

Microorganism identification turnaround time (ID-TAT)

For samples undergoing rapid identification, the overall median ID-TAT significantly decreased from 5.75 hours [3.04-9.60] in P1 to 2.35 hours [1.76–3.42] in P5 (p<0.05; Table 1). This improvement was particularly evident for BCbs flagging positive during night shifts: the median ID-TAT dropped from 8.71 hours [6.28-11.13] in P1 to 2.35 hours [1.68-4.00] in P5 (p<0.001).

Among Non Urg-BCbs, the median ID-TAT declined from 5.75 hours [3.04-9.60] in P1 to 2.37 hours [1.77-3.53] in P5, representing a statistically significant decrease (p<0.001). This reduction was driven primarily by improvements during night shifts, where ID-TAT decreased from 8.71 hours [6.28-11.13] in P1 to 2.41 hours [1.70-4.21] in P5 (p<0.001). Conversely, for samples flagging positive during the daytime, the median time remained stable, ranging from 2.73 hours [2.05-3.55] in P1 to 2.37 hours [1.87-3.05] in P5.

Regarding Urg-BCbs, the ID-TAT decreased from 2.72 hours [2.15-4.42] in P2 to 2.21 hours [1.60-3.18] in P5 (p<0.001 for the trend). A significant reduction was observed specifically for Urg-BCbs flagging positive during the night in P5 compared to all previous periods (p<0.001). In contrast, ID-TAT remained unchanged for Urg-BCbs flagging positive during the day.

The microorganisms isolated from positive blood cultures are detailed in Table 2. From this point onward, BCbs identified as contaminants were excluded from the analyses.

	Overall		Non Urg-BC		Urg-BC	
	N	**% / tot Pos BCs (9176)**	**N**	**% /tot pos BCs (8078)**	**N**	**% /tot pos BCs (1098)**
Escherichia coli	1492	16.26%	1284	15,90%	208	18,94%
Klebsiella pneumoniae	609	6,64%	534	6,61%	75	6,83%
Pseudomonas aeruginosa	331	3,61%	288	3,57%	43	3,92%
Enterobacter spp.	168	1,83%	144	1,78%	24	2,19%
Klebsiella spp.	144	1,57%	129	1,6%	15	1,37%
Salmonella spp.	24	0,26%	20	0,25%	4	0,36%
H. influenzae / N. meningitidis	14	0,15%	14	0,17%	0	-
Anaerobic Gram neg.	102	1,11%	93	1,15%	9	0,82%
Other Gram neg.	646	7,04%	565	6,99%	81	7,38%
Staphylococcus spp.	1727	18,82%	1552	19,21%	175	15,94%
Staphylococcus aureus	831	9,06%	744	9,21%	87	7,92%
Enterococcus spp.	624	6,8%	546	6,76%	78	7,1%
Viridans group streptococci	290	3,16%	261	3,23%	29	2,64%
Streptococcus pneumoniae	86	0,94%	72	0,89%	14	1,28%
Streptococcus agalactiae	31	0,34%	24	0,3%	7	0,64%
Streptococcus pyogenes	22	0,24%	16	0,2%	6	0,55%
Listeria spp.	13	0,14%	10	0,12%	3	0,27%
Anaerobic Gram pos.	74	0,81%	64	0,79%	10	0,91%
Other Gram pos.	260	2,83%	220	2,72%	40	3,64%
Candida spp.	704	7,67%	609	7,54%	95	8,65%
Polymicrobial BCbs	984	10,72%	889	11,01%	95	8,65%

Table 2 - Count of microorganisms isolated from Non Urg-BCbs and Urg-BCbs.

Gram-positive cocci (GPC) were the most frequently isolated bacteria. *Staphylococcus spp.* accounted for 18.82% of isolates, including *Staphylococcus aureus* (9.06%), while *Enterococcus spp.* represented 6.80%. Gram-negative bacteria constituted 38.47% of all isolates, with *Escherichia coli* being the most prevalent species (16.26%). Fungi were isolated in 7.67% of cases, and polymicrobial flora in 10.72%.

When analyzing the five main clinical areas of the hospital (E.R., General Medicine, General Surgery, ICUs, and Infection diseases wards), distinctive pathogen distributions emerged (Figure 4). *E. coli* (39.60%) was the most frequently isolated microorganism in the E.R., whereas *Candida spp.* were predominant in the ICUs (22.09%). *S. aureus* was the leading isolate in the Infectious disease ward (23.41%), while polymicrobial infections were most common in General Surgery (20.19%). Notably, *Streptococcus pneumoniae* was isolated more frequently in Urg-BCbs (1.28%) compared to Non Urg-BCbs (0.89%). Overall, blood cultures positive for pathogens associated with meningitis (i.e., *S. pneumoniae*, *Neisseria meningitidis*, and *Haemophilus influenzae*) were detected with equal frequency during day and night shifts.

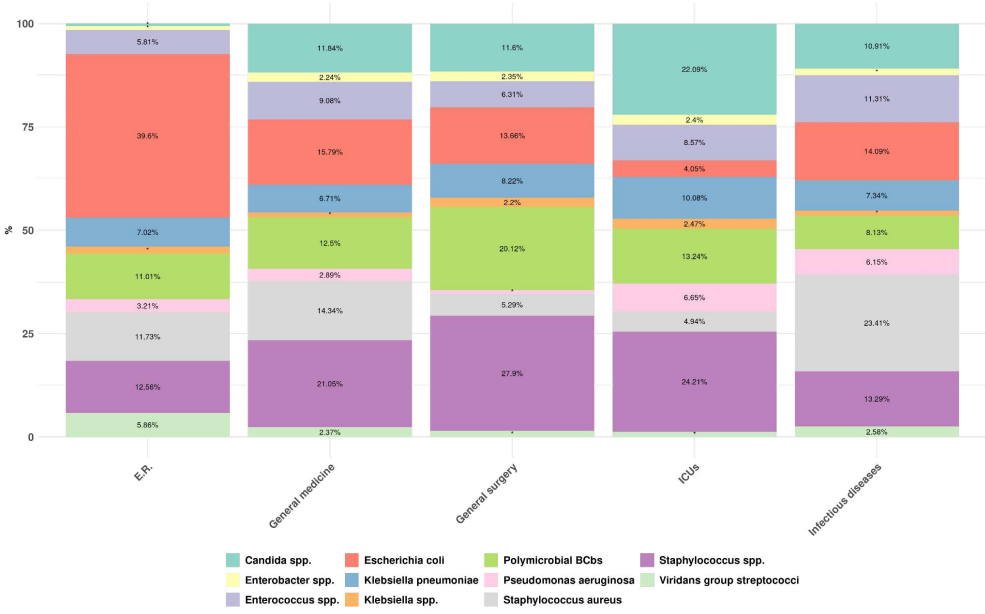


Figure 4 - Distribution of the most frequently isolated microorganisms from BCbs in five hospital areas. Values below 2% are shown as *.

Resistance markers turnaround time (Marker-TAT)

The overall median Marker-TAT significantly decreased from 6.97 hours [3.63–11.05] in P1 to 2.68 hours [1.94–3.97] in P5 ($p<0.05$; Table 1).
Among Non Urg-BCbs, the median Marker-TAT showed a significant decrease from 6.97 hours [3.63-11.05] in P1 to 2.71 hours [1.98-4.07] in P5 ($p<0.05$). Furthermore, for Non Urg-BCbs flagging positive during the night, the Marker-TAT significantly declined from 7.18 hours [5.54-8.63] in P2 to 2.85 hours [1.86-5.53] in P5. For Urg-BCbs, the Marker-TAT decreased from 4.90 hours [2.30–7.26] in P2 to 2.51 hours [1.79–3.61] in P5, although this trend did not reach statistical significance.
Regarding GNB resistance markers across the five periods, the prevalence of ESBL-producing *Enterobacterales* remained stable over time (range: 18.40%–27.59%). In contrast, the prevalence of CRE increased following the first COVID-19 pandemic wave. Specifically, CRE rates rose from 3.23% in P2 to 18.97% in P3, before stabilizing in P4 and P5 (7.98% and 10.90%, respectively).

Discussion

Sepsis remains a critical public health challenge and a leading cause of in-hospital mortality, with case fatality rates ranging from 20% to 45% (15). Since mortality associated with septic shock increases by approximately 7% for every hour of delay in appropriate antimicrobial therapy (16), sepsis is universally acknowledged as a time-sensitive emergency requiring rapid diagnosis and coordinated intervention (17). Consequently, the efficiency of the microbiological workflow, spanning pre-analytical, analytical, and post-analytical phases, is pivotal for optimizing patient outcomes (1,5,18,19).
This study demonstrates that reorganizing the CML to a 24/7 operational model, combined with a fast-track pathway for critical patients, effectively reduced TATs for all analytical phases.

The necessity of a 24/7 workflow is underscored by our finding that over half (54%) of positive BCbs flagged during nighttime hours. These nocturnal detections frequently involved time-critical pathogens, including agents of meningitis (e.g., *Neisseria meningitidis*, *Streptococcus pneumoniae*, *Haemophilus influenzae*) and fungi such as *Candida spp.* (20,21). In the absence of night shifts, the diagnosis of these life-threatening infections would be significantly delayed, potentially compromising patient survival. Historically, sepsis onset during night shifts has been associated with longer therapeutic delays and higher mortality compared to day shifts (25,26).

Implementation of the 24/7 model led to drastic and sustained reductions in TATs. At baseline (P1), the median Gram-TAT for samples flagging positive at night was approximately 8 hours. By the validation period (P5), this decreased to 1.27 hours for Urg-BCbs and 1.48 hours for Non Urg-BCbs. Crucially, these results demonstrate that the performance improvements achieved in earlier phases were maintained over time, despite the high staff turnover and increased diagnostic workload characterizing the post-COVID-19 period. This consistency highlights the robustness, standardization, and reproducibility of the reorganized workflow (22).

The Gram-staining results continue to play a crucial role in the management of septic patients, and can be associated with a decreased patient's mortality (23,27,28). Cattoir and colleagues evidenced that about one third of positive blood cultures incited an antibiotic initiation or switch, based on the Gram stain result, mostly due to the presence of Gram positives or yeasts (29). In our study, Gram-positive accounted for 43.13%, followed by Gram-negative (38.47%), and *Candida spp*. (7.67%).
Furthermore, by integrating continuous processing with rapid molecular and mass spectrometry methods, we achieved a median time from bottle positivity to full identification and resistance profiling of under 3 hours. When combined with a median time-to-positivity of 14 hours, definitive microbiological data were available to clinicians within a median of 17 hours from sample collection. This speed is essential for optimizing empirical therapy, particularly given the rising prevalence of multidrug-resistant organisms (e.g., CRE-producing *Enterobacterales* or *P. aeruginosa*), a trend that has intensified in the aftermath of the COVID-19 pandemic (33,34). Due to rising resistance rates, especially among Gram negative bacteria, the empirical therapy for septic patients may often be inappropriate, leading to increased mortality (35).

A key innovation of this model is the establishment of a preferential diagnostic pathway for critically ill patients ("Urg-BC"). This system functions as a form of tailored diagnostic medicine: by flagging samples from high-risk patients (who accounted for ~10% of the total volume but showed a higher positivity rate of 16%), the laboratory can prioritize processing and ensure immediate communication of results (36). Notably, *S. pneumoniae* was more frequently isolated from Urg-BCbs, further validating the clinical relevance of this prioritization.

This study is limited by its single-center design, which may affect the generalizability of the findings to other healthcare settings. However, the extensive five-year observation period supports the robustness of the results. A primary challenge during the reorganization involved standardizing procedures across a multidisciplinary team, including staff from virology and parasitology, within a large, second-level laboratory.

Conclusion

The reorganization of the CML into a fully integrated 24/7 service significantly reduced diagnostic TATs, particularly for samples flagging positive overnight. By combining a prioritized workflow for critically ill patients with advanced diagnostic technologies, this model enhances the hospital's capacity to manage sepsis effectively. This approach demonstrates the value of precision microbiology in supporting antimicrobial stewardship and improving early targeted treatment for life-threatening infections.

Acknowledgement

All medical residents who supported the diagnostic process through their dedicated commitment. San Matteo Pavia Microbiology and Virology Working Group: Jessica Bagnarino, Daniela Barbarini, Alice Bonetti, Caterina Cavanna, Vincenzina Monzillo, Antonio Piralla, Roberta Maserati, Giulia Campanini, Irene Cassaniti, Milena Furione, Daniele Lilleri, Guglielmo Ferrari, Stefania Paolucci, Francesca Rovida, Giuditta Comolli, Pasquale Piemontese, Chiara Rebuffa, Vincenzo Brunco, Alessia Girello, Marco Ardizzone, Caterina Scotti, Alessandro Russo, Riccardo Cavanna, Ermanna Cabano, Luca Dossena, Cristina Zollino, Alice Fratini, Randa Abu Shweimeh, Margherita Torso, Laura Demurtas, Simona Pezzaia, Giuliana Casali, Viviana Landini, Lucia Municchi.

Ethical statement

The study was approved by the local Research Ethics Committee Foundation (P-20200109218, protocol 20210015825).

References

1. Van Heuverswyn J, Valik JK, Desirée van der Werff S, Hedberg P, Giske C, Nauclér P. Association Between Time to Appropriate Antimicrobial Treatment and 30-day Mortality in Patients With Bloodstream Infections: A Retrospective Cohort Study. Clin Infect Dis. 2023 Feb 8;76(3):469–78.
2. Clinical and Laboratory Standards Institute. CLSI. Principles and Procedures for Blood Cultures. 2nd ed. CLSI guideline M47. 2022.

3. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 2021 Nov;47(11):1181–247.
4. King J, Chenoweth CE, England PC, Heiler A, Kenes MT, Raghavendran K, et al. Early recognition and initial management of sepsis in adult patients. Ann Arbor (MI): Michigan Medicine University of Michigan; 2023.
5. Lockwood AM, Perez KK, Musick WL, Ikwuagwu JO, Attia E, Fasoranti OO, et al. Integrating rapid diagnostics and antimicrobial stewardship in two community hospitals improved process measures and antibiotic adjustment time. Infect Control Hosp Epidemiol. 2016 Apr;37(4):425–32.
6. Lamy B, Dargère S, Arendrup MC, Parienti J-J, Tattevin P. How to Optimize the Use of Blood Cultures for the Diagnosis of Bloodstream Infections? A State-of-the Art. Front Microbiol. 2016 May 12;7:697.
7. Tabak YP, Vankeepuram L, Ye G, Jeffers K, Gupta V, Murray PR. Blood culture turnaround time in U.S. acute care hospitals and implications for laboratory process optimization. J Clin Microbiol. 2018 Dec;56(12).
8. Seminari E, Colaneri M, Corbella M, De Silvestri A, Muzzi A, Perlini S, et al. Reduction of BSI associated mortality after a sepsis project implementation in the ER of a tertiary referral hospital. Sci Rep. 2023 Mar 29;13(1):5142.
9. Mariani B, Corbella M, Seminari E, Sacco L, Cambieri P, Capra Marzani F, et al. Evaluation of a model to improve collection of blood cultures in patients with sepsis in the emergency room. Eur J Clin Microbiol Infect Dis. 2018 Feb;37(2):241–6.
10. Liu W, Liao K, Wu J, Liu S, Zheng X, Wen W, et al. Blood culture quality and turnaround time of clinical microbiology laboratories in Chinese Teaching Hospitals: A multicenter study. J Clin Lab Anal. 2024 Jan 18;38(1–2):e25008.
11. Royal College of Physicians of London. National Early Warning Score (NEWS) : Standardising the Assessment of Acute-Illness Severity in the NHS. Royal College of Physicians; 2012.
12. Royal College of Physicians of London. National Early Warning Score (NEWS) 2 [Internet]. 2017 [cited 2025 Jul 24]. Available from: https://www.rcp.ac.uk/improving-care/resources/national-early-warning-score-news-2/
13. Bekeris LG, Tworek JA, Walsh MK, Valenstein PN. Trends in blood culture contamination: a College of American Pathologists Q-Tracks study of 356 institutions. Arch Pathol Lab Med. 2005 Oct;129(10):1222–5.
14. Smith AC, Hussey MA. Gram Stain Protocols [Internet]. 2005 [cited 2025 Jul 24]. Available from: https://asm.org/protocols/gram-stain-protocols

15. Stoller J, Halpin L, Weis M, Aplin B, Qu W, Georgescu C, et al. Epidemiology of severe sepsis: 2008-2012. J Crit Care. 2016 Feb;31(1):58–62.
16. Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med. 2006 Jun;34(6):1589–96.
17. Kahn JM, Davis BS, Yabes JG, Chang C-CH, Chong DH, Hershey TB, et al. Association Between State-Mandated Protocolized Sepsis Care and In-hospital Mortality Among Adults With Sepsis. JAMA. 2019 Jul 16;322(3):240–50.
18. Morquin D, Lejeune J, Agostini C, Godreuil S, Reynes J, Le Moing V, et al. Time is of the essence: achieving prompt and effective antimicrobial therapy of bloodstream infection with advanced hospital information systems. Clin Infect Dis. 2024 Jun 14;78(6):1434–42.
19. Venturelli C, Righi E, Borsari L, Aggazzotti G, Busani S, Mussini C, et al. Impact of Pre-Analytical Time on the Recovery of Pathogens from Blood Cultures: Results from a Large Retrospective Survey. PLoS ONE. 2017 Jan 3;12(1):e0169466.
20. Deghmane A-E, Taha S, Taha M-K. Global epidemiology and changing clinical presentations of invasive meningococcal disease: a narrative review. Infect Dis (Lond). 2022 Jan;54(1):1–7.
21. Pappas PG, Kauffman CA, Andes DR, Clancy CJ, Marr KA, Ostrosky-Zeichner L, et al. Clinical practice guideline for the management of candidiasis: 2016 update by the infectious diseases society of america. Clin Infect Dis. 2016 Feb 15;62(4):e1-50.
22. Antonios K, Croxatto A, Culbreath K. Current state of laboratory automation in clinical microbiology laboratory. Clin Chem. 2021 Dec 30;68(1):99–114.
23. Barenfanger J, Graham DR, Kolluri L, Sangwan G, Lawhorn J, Drake CA, et al. Decreased mortality associated with prompt Gram staining of blood cultures. Am J Clin Pathol. 2008 Dec;130(6):870–6.
24. French K, Evans J, Tanner H, Gossain S, Hussain A. The Clinical Impact of Rapid, Direct MALDI-ToF Identification of Bacteria from Positive Blood Cultures. PLoS ONE. 2016 Dec 30;11(12):e0169332.
25. Ginestra JC, Kohn R, Hubbard RA, Auriemma CL, Patel MS, Anesi GL, et al. Association of Time of Day with Delays in Antimicrobial Initiation among Ward Patients with Hospital-Onset Sepsis. Ann Am Thorac Soc. 2023 Sep;20(9):1299–308.
26. Lyons PG, Hough CL. Antimicrobials in sepsis: time to pay attention to when delays happen. Ann Am Thorac Soc. 2023 Sep;20(9):1239–41.
27. Perez KK, Olsen RJ, Musick WL, Cernoch PL, Davis JR, Peterson LE, et al. Integrating rapid diagnostics and antimicrobial stewardship improves outcomes in patients with antibiotic-resistant Gram-negative bacteremia. J Infect. 2014 Sep;69(3):216–25.

28. Messacar K, Parker SK, Todd JK, Dominguez SR. Implementation of rapid molecular infectious disease diagnostics: the role of diagnostic and antimicrobial stewardship. J Clin Microbiol. 2017 Mar;55(3):715–23.
29. Cattoir L, Coorevits L, Leroux-Roels I, Claeys G, Verhasselt B, Boelens J. Improving timelines in reporting results from positive blood cultures: simulation of impact of rapid identification on therapy on a real-life cohort. Eur J Clin Microbiol Infect Dis. 2018 Dec;37(12):2253–60.
30. Banerjee R, Komarow L, Virk A, Rajapakse N, Schuetz AN, Dylla B, et al. Randomized Trial Evaluating Clinical Impact of RAPid IDentification and Susceptibility Testing for Gram-negative Bacteremia: RAPIDS-GN. Clin Infect Dis. 2021 Jul 1;73(1):e39–46.
31. Yuceel-Timur I, Thierry E, Chainier D, Ndao I, Labrousse M, Grélaud C, et al. Retrospective evaluation of rapid genotypic ID and phenotypic AST systems on positive blood culture turnaround time and simulated potential impacts on bloodstream infection management. J Antimicrob Chemother. 2024 Sep 19;79(Supplement_1):i26–31.
32. Schneiderhan W, Grundt A, Wörner S, Findeisen P, Neumaier M. Work flow analysis of around-the-clock processing of blood culture samples and integrated MALDI-TOF mass spectrometry analysis for the diagnosis of bloodstream infections. Clin Chem. 2013 Nov;59(11):1649–56.
33. Fasciana T, Antonelli A, Bianco G, Lombardo D, Codda G, Roscetto E, et al. Multicenter study on the prevalence of colonization due to carbapenem-resistant Enterobacterales strains before and during the first year of COVID-19, Italy 2018-2020. Front Public Health. 2023 Dec 19;11:1270924.
34. European Centre for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-Net) -Annual Epidemiological Report 2022. . Stockholm: ECDC; 2023.
35. Kumar A, Ellis P, Arabi Y, Roberts D, Light B, Parrillo JE, et al. Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. Chest. 2009 Nov;136(5):1237–48.
36. Halstead FD, Pinjuh G, Antonacci G, Proudlove N. Reducing laboratory delays in blood culture pathogen identification: a quality improvement project. BMJ Open Qual. 2025 Mar 22;14(1).

CHAPTER 8:

Gut microbiota and B Cell receptor (BCR) inhibitors for the treatment of Chronic Lymphocytic Leukemia:
Is biodiversity correlated with clinical response or immune-related adverse event occurrence? A Cross-Sectional study

Valentina Zuccaro [1], **Greta Petazzoni** [2,3] , Irene Mileto [2], , Marta Corbella [2], Erika Asperges [1], Paolo Sacchi [1], Sara Rattotti [4], Marzia Varettoni [4], Irene Defrancesco [3,4], Patrizia Cambieri [2], Fausto Baldanti [2,3], Luca Arcaini [4,5] and Raffaele Bruno [1,3]

1. Division of Infectious Diseases I, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
2. Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
3. Department of Medical, Surgical, Diagnostic and Pediatric Science, University of Pavia, 27100 Pavia, Italy
4. Division of Hematology, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
5. Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy

This manuscript has been **published on Microorganisms**, May 2023, Volume 11, Issue 1305, https://doi.org/10.3390/microorganisms11051305

In this work I conducted all the bioinformatic analyses and I contributed in drafting and revising the manuscript.

Abstract

The relationship between the gut microbiota (GM) and the response to immunotherapy in cancer patients has been substantiated by several studies, underscoring the potential utility of GM as a biomarker of response. B-cell receptor (BCR) inhibitors (BCRi), which constitute targeted therapies, represent the most recent approach for the treatment of chronic lymphocytic leukemia (CLL). However, a satisfactory response is not achieved by all patients, and immune-related adverse events (irAEs) can also diminish efficacy.

The objective of this study was to compare GM biodiversity in patients with CLL who had been treated with BCRi for a minimum of 12 months. Twelve patients were enrolled in the study: 10 patients in the responder group (R) and 2 patients in the non-responder group (NR). Seven patients (58.3%) experienced adverse reactions (AE).

Although a significant difference was not observed across the study population in terms of relative abundance, as well as alpha and beta diversity, a distinct distribution of bacterial taxa was identified between the analyzed groups. Specifically, a higher level of the class *Bacteroidia* and the order *Bacteroidales* was noted in the R group, along with an inversion in the *Firmicutes* and *Bacteroidetes* ratio within the AE group. Prior research has not focused on establishing a link between GM and the response to BCRi in this patient population. While these analyses are preliminary, they offer suggestions to guide future research endeavors.

Introduction

Chronic lymphoproliferative disorders (LPD) constitute a heterogeneous spectrum of pathologies defined by the clonal proliferation of lymphocytes, resulting in monoclonal lymphocytosis, lymphadenopathy, and bone marrow involvement (1). Within this category, Chronic Lymphocytic Leukemia (CLL) represents the most prevalent leukemia in the adult population. Adhering to the updated consensus of the 2018 International Workshop on Chronic Lymphocytic Leukemia (2), therapeutic intervention is mandated solely for patients exhibiting "active disease" (characterized by marrow failure-induced cytopenias, significant or symptomatic organomegaly, progressive lymphocytosis, or severe disease-related symptoms).

While the therapeutic armamentarium is diverse, the advent of targeted therapies has fundamentally revolutionized CLL management. Notably, agents targeting Bruton tyrosine kinase (BTK) have exhibited remarkable efficacy in disease control (3). Both Ibrutinib, a first-in-class BTK inhibitor, and Idelalisib, a first-in-class phosphatidylinositol 3-kinase δ (PI3K-inhibitor), have secured approval for managing indolent B-cell malignancies.

During the recruitment phase of the present study, these B-cell receptor inhibitors (BCRi) were primarily deployed to address unmet clinical needs, such as in patients with relapsed/refractory disease, high-risk genetic profiles, or significant comorbidities (4).
Despite the high potency of BCRi regimens, a subset of patients manifests primary resistance or develops immune-related adverse events (irAEs) (2). Given that the majority of these toxicities affect the gastrointestinal tract (manifesting as colitis or diarrhea), the Gut Microbiota (GM) has been hypothesized as a potential modulator of both drug tolerance and efficacy.

Substantial evidence supports the existence of a crosstalk between the GM and host immunity, particularly in solid tumors. For instance, in melanoma, the GM composition appears to influence the success of anti-CTLA4 and anti-PD1 checkpoint inhibitor therapies. Preclinical models demonstrated that the efficacy of anti-CTLA4 antibodies in controlling sarcoma growth correlated with the abundance of *Bacteroides fragilis* and *Burkholderia cepacia* (5). Furthermore, metagenomic analyses in melanoma patients revealed an enrichment of *Clostridiales* (specifically *Ruminococcaceae* and *Faecalibacterium*) in responders to anti-PD1 therapy compared to non-responders (6).

In the hematological setting, the bidirectional relationship between the microbiome and therapeutic response is increasingly recognized: while the GM can modulate drug efficacy, chemotherapeutic agents can reciprocally induce dysbiosis (7). Although a definitive consensus is lacking, emerging data suggest a link between specific microbial signatures and adverse

outcomes; for example, Hakim et al. reported associations between *Enterococcaceae* or *Streptococcaceae* abundance and infection risk, as well as between *Proteobacteria* and febrile neutropenia in acute lymphocytic leukemia (8). Consequently, the GM profile is emerging as a putative biomarker for stratification (9) and a potential therapeutic target, as evidenced by improved progression-free survival (PFS) in refractory melanoma patients following fecal microbiota transplantation from complete responders (10,11).

To date, the specific alterations of the GM in CLL patients treated with BCRi, and their potential association with clinical outcomes or irAEs, remain unexplored. We hypothesized that the distinct immunomodulatory and antitumor properties of BCRi (specifically Ibrutinib and Idelalisib) might be influenced by the host microbiome configuration. Therefore, the primary aim of this study was to investigate correlations between GM characteristics, therapeutic response, and the incidence of irAEs in a cohort of CLL patients undergoing BCRi treatment.

Materials and methods

Study design and patient cohort

A cross-sectional observational design was employed for this study. The cohort comprised adult patients diagnosed with CLL attending the Hematology Unit of the Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy), who had initiated BCRi therapy (specifically BTK or PI3K inhibitors) at least 12 months prior to enrollment.
Fecal specimen collection was scheduled to coincide with the 12-month follow-up visit post-treatment initiation. In cases of premature therapy discontinuation, sample collection was performed at the time point closest to the 12th month. Written informed consent was obtained from all participants prior to inclusion.
Based on the response criteria, the population was stratified into two primary groups:

- Responder Group (R): patients exhibiting maintained clinical response (absence of progression criteria) while remaining on active BCRi therapy.
- Non-Responder Group (NR): patients manifesting criteria for clinical progression or those requiring BCRi discontinuation within the first 12 months of treatment.

Comprehensive metadata were curated, including demographics (age, sex), clinical and hematological history (CLL diagnosis date, cytogenetic/molecular profiles, lines of therapy, adverse events), antibiotic usage within the preceding three months, and eventual outcomes (cause of death). Dietary habits, specifically restrictive regimens (e.g., gluten-free, lactose-free, casein-free), were also logged. Therapeutic response assessment was conducted starting two months post-initiation and monitored according to standard clinical practice. Additionally, a secondary stratification was applied based on the occurrence of immune-related adverse events (irAEs): the AE group (patients experiencing irAEs at any stage) and the NAE group (patients free from irAEs).

Objectives

The primary endpoint was to characterize the GM landscape in CLL patients to identify distinct compositional signatures differentiating responders (R) from non-responders (NR). The secondary endpoint focused on evaluating GM biodiversity variations associated with the development of irAEs.

Sample collection and sequencing

Patients were instructed to provide stool samples at the first scheduled visit falling within the study window (at least 12 months from BRCi initiation). Specimens were stored at -80 °C at the Microbiology and Virology Unit of the HSM.
Genomic DNA extraction was performed using the DNeasy® PowerSoil Pro Kit (Qiagen, Hilden, Germany), following the manufacturer's protocol. The V3–V6 hypervariable regions of the 16S rRNA gene were amplified utilizing the Microbiota Solution B Kit (Arrow Diagnostics s.r.l., Genova, Italy). Subsequently, the amplicon libraries underwent paired-end sequencing (2x250) on the Illumina MiSeq platform (Illumina, CA, USA).

Bioinformatic and statistical analysis

Raw sequencing data underwent quality control and processing via the MicrobAT (Microbiota Analysis Tool) system (SmartSeq S.r.l., Novara, Italy). Following filtration for read length and quality, sequences were aligned against the RDP database (v11.4). Taxonomic assignment was performed based on stringent criteria: query coverage ≥80% and sequence similarity ≥97% enabling classification from phylum down to species level.
Subsequent statistical analyses were conducted using a custom pipeline in R. Feature tables were filtered to remove noise, excluding taxa with low variance or low counts (defined as <4 counts in <20% of samples). Relative abundances were then calculated for the retained features. Comparative analysis of microbial profiles was restricted to taxa with a prevalence >1% in at least one sample. Differences between groups (R vs. NR; AE vs. NAE) were assessed using the Mann–Whitney U-test, with statistical significance defined as $p<0.05$.
Alpha-diversity (within-sample diversity) was estimated using the phyloseq R package (13), comparing Observed Richness and Shannon indices between groups via the Mann–Whitney U-test ($p<0.05$). Beta-diversity (between-sample diversity) was computed at all taxonomic levels using the vegan R package (14). Dissimilarity matrices were constructed using Bray-Curtis distances and visualized via Principal Coordinate Analysis (PCoA). Statistical separation between groups was evaluated using Permutational Multivariate Analysis of Variance (PERMANOVA), with a significance threshold of $p < 0.05$.

Results

Between January 2019 and October 2019, a total of 13 patients were initially recruited at the Hematology Unit of the HSM. Following the exclusion of one subject, the final analytical cohort consisted of 12 patients. The mean age of the study population was 63.6 years, with a perfectly balanced gender distribution (6/12 males; 50%). All participants had a confirmed diagnosis of CLL and reported no adherence to specific dietary regimens.

Regarding pharmacological management, seven patients (58.3%) received Ibrutinib, while the remaining five (41.7%) were treated with Idelalisib. BCRi therapy was administered as a first-line treatment in a single case; for the vast majority of the cohort, it represented a subsequent line of therapy. Stratification identified ten patients satisfying the definition of Responders (R Group), whereas two patients met the criteria for clinical progression and were classified as Non-Responders (NR Group). In terms of safety, adverse events (AEs) were documented in seven patients (58.3%), who constituted the AE Group. The most frequently reported toxicity was diarrhea (affecting three of the seven patients), followed by pyrexia and hepatotoxicity. Notably, Idelalisib was the therapeutic agent in five of these seven cases. Conversely, both NR patients were undergoing treatment with Ibrutinib. It is important to note that, although both NR patients experienced immune-related AEs (irAEs), these toxicities were not the primary determinants for treatment discontinuation. Detailed demographic and hematological profiles of the cohort are summarized in Table 1.

ID Patient	Age (YO)	Sex	BCRi	Genomic Features	Treatment Lines	Antibiotic Intake	irAEs	Outcome
01	40–50	M	Ibrutinib	TP53 mut., IGVH unmut.	3°	No	Yes	NR
03	60–70	M	Idelalisib	TP53 mut., IGVH unmut.	2°	Yes	Yes	R
04	60–70	F	Ibrutinib	TP53 mut., IGVH unmut., del(13q), del(17p)	2°	Yes	No	R
05	60–70	F	Ibrutinib	TP53 mut., IGVH unmut.	2°		No	R
06	60–70	M	Idelalisib	TP53 mut., IGVH unmut., del(17p), Trisomy 13	2°	Yes	Yes	R
07	60–70	F	Idelalisib	TP53 mut., IGVH unmut., del(13q)	4°	Yes	Yes	R
08	60–70	F	Idelalisib	TP53 mut., IGVH unmut., del(13q)	2°	Yes	Yes	R
09	60–70	M	Ibrutinib	TP53 mut., IGVH unmut., del(17p)	5°	Yes	Yes	NR
10	60–70	M	Ibrutinib	IGVH unmut., del(13q)	4°	Yes	No	R
11	70–80	F	Ibrutinib	TP53 mut., IGVH unmut., del(17p)	1°	No	No	R
12	70–80	F	Idelalisib	IGVH unmut., del(13q)	2°	No	Yes	R
13	60–70	M	Ibrutinib	IGVH unmut., Trisomy 13	3°	Yes	No	R

Table 1 - Patients' characteristics. B-cell receptor inhibitors used to treat patients are shown in the BCRi. Mutated (mut.) genes, unmutated (unmut.) genes, deletions of

chromosomes and the trisomy recorded are listed in genomic features. irAEs are the immune-related adverse events experienced by each patient. Non-responder (NR) and responder (R) patients are shown in the outcome.

Taxonomic profiling and diversity in clinical response groups

The gut microbial architecture was characterized and compared between patients achieving a clinical response (R) and those who did not (NR). Statistical analysis revealed no significant disparities in the mean relative abundance of taxa at the phylum, class, or order levels between the two cohorts (Figure 1).
At the phylum level, *Firmicutes* constituted 38.3% of the microbiome in the R group compared to 47.3% in the NR group. Conversely, *Bacteroidetes* abundance was recorded at 35.7% in responders versus 28.2% in non-responders.
While lacking statistical significance, observable trends at the class level indicated that the R group harbored lower proportions of *Clostridia* (32.3%) and elevated levels of *Bacteroidia* (34.9%). In contrast, the NR group displayed an inverted pattern, with *Clostridia* and *Bacteroidia* accounting for 40.9% and 27.9%, respectively (Figure 1b).

A similar divergence in relative abundance was noted at the order level, though it did not reach statistical significance. Specifically, *Bacteroidales* represented the predominant order in the R group (34.9%), whereas *Clostridiales* prevailed in the NR group (40.1%) (Figure 1c).

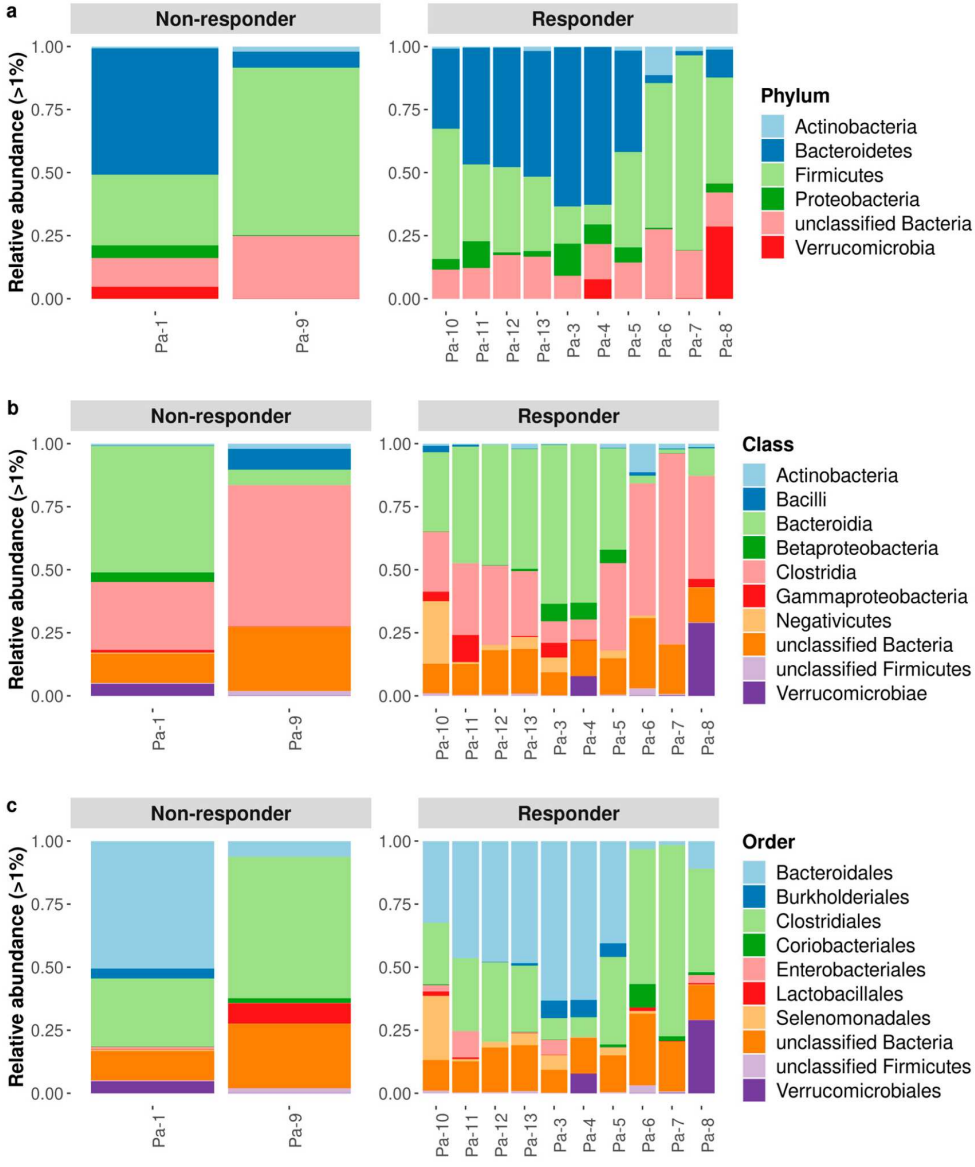


Figure 1 - Taxonomic composition and diversity of fecal bacterial communities for each patient stratified for responder and non-responder group. The most represented phyla (a), classes (b), and orders (c) identified in the study groups are shown with relative abundance. Only taxa whose relative abundance was >1% in at least one group were included.

At the species level, a significant differential abundance was observed: *Bacteroides* sp. S-18 was significantly enriched in the NR group compared to the R group ($p = 0.034$; Mann–Whitney U-test).
To ensure sample homogeneity and minimize confounding factors, a sub-analysis was restricted exclusively to Ibrutinib-treated patients, given that the NR cohort consisted solely of individuals receiving this specific agent. Consequently, a refined responder group (designated as RI) was established. The taxonomic architecture of the RI cohort mirrored the profile observed in the broader R group.
At the phylum rank, *Firmicutes* accounted for 32.2% of the microbiome in the RI group, while *Bacteroidetes* reached 46.3%. Consistent with the findings in the general R group, class-level analysis of the RI cohort confirmed reduced levels of fecal *Clostridia* (32.3%) and elevated levels of fecal *Bacteroidia* (34.9%) when compared to the NR group (Supplementary Figure S1). This trend persisted at the order level, where *Bacteroidales* emerged as the dominant taxon in the RI group, exhibiting a relative abundance of 44.9%. Statistical evaluation of biodiversity indices revealed no significant disparities between the RI and NR cohorts in terms of either alpha-diversity or beta-diversity (Supplementary Figures S2 and S3).

The α-diversity indexes (observed richness and Shannon) were computed at all taxonomic levels to analyze the within-sample diversity, and there were no significant differences among the Responder vs. Non-Responder Groups (Figure 2).

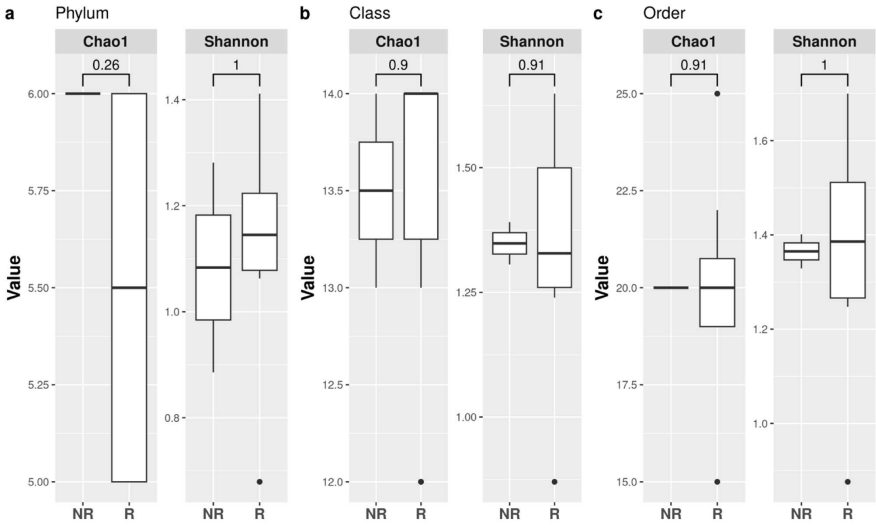


Figure 2 - α-diversity. Observed richness and Shannon indices are presented at the taxonomic level of phylum (a), class (b), and order (c). Significant ($p < 0.05$) comparisons between responder (R) and non-responder (NR) patients are indicated in the boxplot.

The β-diversity (diversity in composition among samples) was computed at all taxonomic levels, and we did not observe any significant differences, suggesting the lack of a different gut microbial community between the analyzed groups (Figure 3).

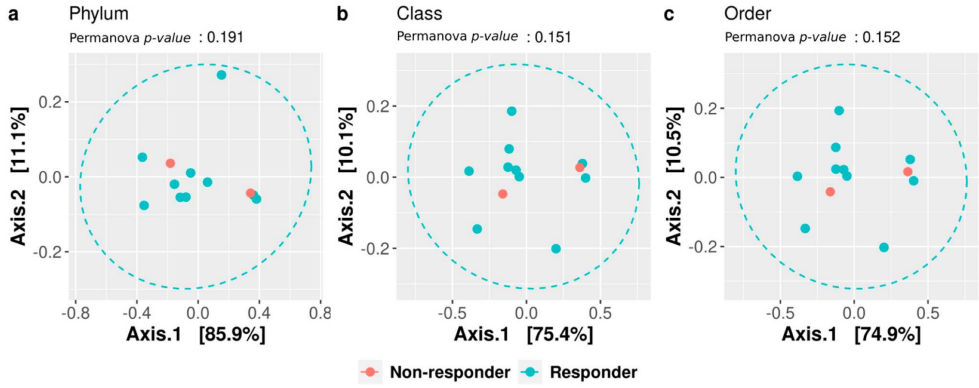


Figure 3 - β-diversity. The microbiota distances were evaluated through the Bray-Curtis dissimilarity matrix at the taxonomic level of phylum (a), class (b), and order (c) and visualized through principal coordinate analysis (PCoA). Each point represents the microbiota composition of one sample stratified for outcome (responder and non-responder).

Taxonomic profiling and diversity in relation to immune-related adverse events

The cohort was subsequently stratified based on safety profiles into two distinct categories: the AE group (comprising patients who manifested irAEs) and the Non-Adverse Events (NAE) group (patients free from such toxicities). Although the observed differences did not reach statistical significance, discernible shifts in mean relative abundance were identified across taxonomic ranks.

At the phylum level, the AE group was characterized by an enrichment of *Firmicutes* (45.7%) and a concomitant depletion of *Bacteroidetes* (26.1%). In contrast, the NAE group exhibited an inverted distribution pattern, with values of 31.5% and 46.1%, respectively (Figure 4a).

This compositional divergence persisted at the class level: *Clostridia* emerged as the predominant taxon in the AE group (41%), whereas *Bacteroidia* represented the most abundant class in the NAE cohort (44.7%) (Figure 4b). Similarly, analysis at the order level revealed that *Clostridiales* were elevated in patients experiencing adverse events (40.9%), while *Bacteroidales* constituted the dominant order in the toxicity-free group (44.7%) (Figure 4c).

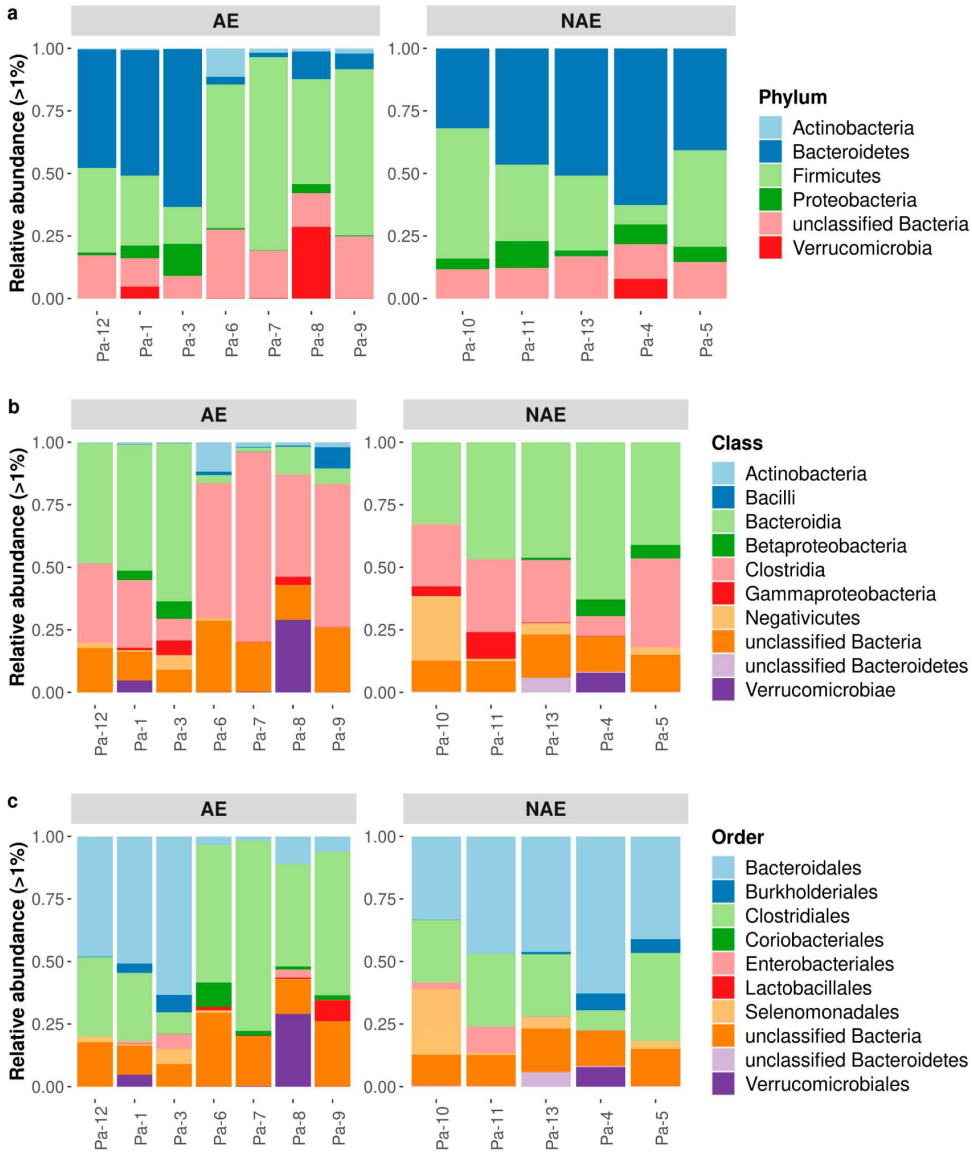


Figure 4 - Taxonomic composition and diversity of fecal bacterial communities for each patient who experienced an irAE (AE) or did not (NAE). The most represented phyla (a), classes (b), and orders (c) identified in the study groups are shown with relative abundance. Only taxa whose relative abundance was >1% in at least one group were included.

Interestingly, some bacterial taxa were only found in one of the two groups, as shown in Table 2.

	AE Group	NAE Group
Phylum	Actinobacteria	-
Class	Actinobacteria	unclassified Bacteroidetes
	Bacilli	
Order	*Coriobacteriales*	unclassified Bacteroidetes
	Lactobacillales	
Family	*Coriobacteriaceae*	*Veillonellaceae*
		Prevotellaceae
Species	*Akkermansia muciniphila* ATCC BAA-835	*Bacteroides eggerthii*
	Blautia luti	*Veillonella spp* ICM51a
	Ruminococcus obeum	*Faecalibacterium prausnitzii*
	Barnesiella spp EBA4-14	*Bacteroides ovatus*
		Butyrivibrio fibrisolvens
		Akkermansia muciniphila
		Bacteroidaceae bacterium Smarlab 3301643
		Parabacteroides distasonis ATCC 8503

Table 2 - Bacterial taxa found only in the AE or NAE group.

α-diversity indexes were also calculated according to irAE categorization, and we did not observe any significant differences between the two groups for all of the taxonomic ranks (Figure 5).

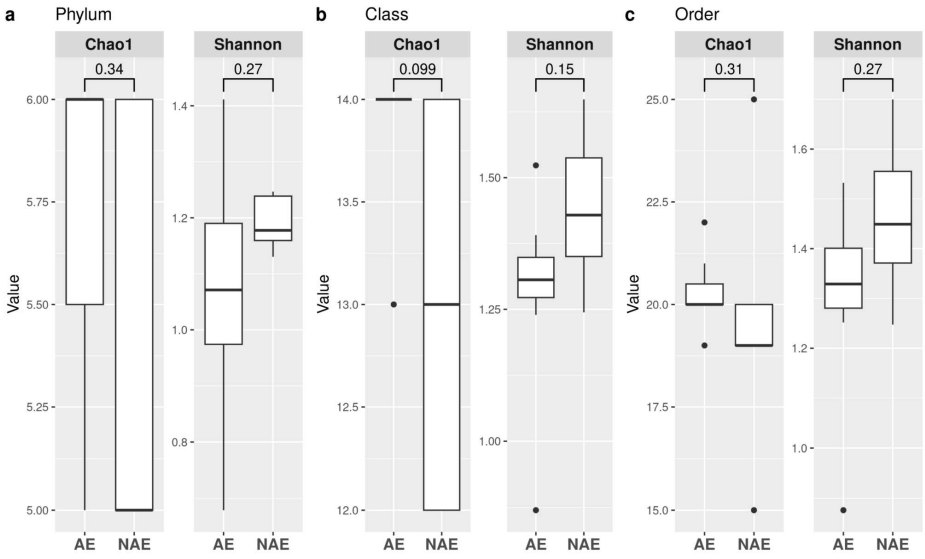


Figure 5 - α-diversity according to the experience of an irAE (AE) or not (NAE). Observed richness and Shannon indices are presented at the taxonomic level of phylum (a), class (b), order (c). Significant ($p < 0.05$) comparisons between AE and NAE are indicated in the boxplot.

Similarly, we performed the between-sample diversity (β-diversity) evaluation, and the data suggested no significant differences in all bacterial consortia when comparing the samples according to irAE categorization (Figure 6).

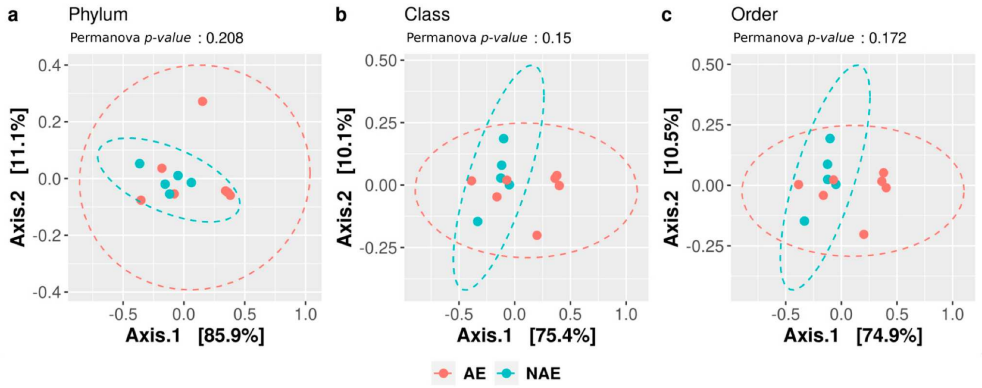


Figure 6 - Beta-diversity according to the experience of an irAE (AE) or not (NAE). The microbiota distances were evaluated through the Bray-Curtis dissimilarity matrix at the taxonomic level of phylum (a), class (b), and order (c) and visualized through principal coordinate analysis (PCoA). Each point represents the microbiota composition of one sample.

Discussion

In this cross-sectional exploratory study, we reported an overall analysis of the GM of patients with CLL receiving BCRi treatment, and we speculated that GM signature was associated with clinical response and the occurrence of irAEs. Although the analysis did not yield statistically significant deviations in relative abundance or diversity metrics across the cohort, likely due to sample size constraints, discernible trends in taxonomic distribution were identified. Specifically, the Responder (R) cohort exhibited an enrichment of bacteria belonging to the class *Bacteroidia* and the order *Bacteroidales*. In contrast, the Non-Responder (NR) group was characterized by a preponderance of the class *Clostridia* and order *Clostridiales*. Notably, this latter profile was mirrored in the AE group, which also displayed elevated levels of *Clostridia* and *Clostridiales*. These observations align with previous reports by Karpinets et al. in cervical cancer models, where Clostridia-rich communities were linked to stress response pathways (15). Not surprisingly, the emerging view based on numerous studies is that collective interactions within the bacterial community drive disease through the dysregulation of mucosal immunity and the disruption of the mucosal barrier.

Furthermore, the AE cohort displayed a notable inversion of the Firmicutes to Bacteroidetes ratio. It is well established that GM metabolic activity generates Short Chain Fatty Acids (SCFAs) - such as acetate, propionate, and butyrate - which are pivotal for modulating inflammation. Butyrate, predominantly synthesized by Firmicutes, maintains colonocyte homeostasis and exerts anti-inflammatory effects (16). With this in mind, the lower abundance of Firmicutes, which means lower butyrate production, is not an unexpected finding in the AE group. Regarding specific taxa within Firmicutes, the butyrate-producing Faecalibacterium prausnitzii had been widely proved to be associated with a good response to immune checkpoint inhibitors (ICIs) and a higher risk of developing irAEs (in particular, immune-related colitis) in metastatic melanoma patients (17). Conversely, within Bacteroidetes, Akkermansia muciniphila is recognized for its protective role against colitis (16). In agreement with literature data, in our cohort, we did not find A. muciniphila in the AE group and we failed to demonstrate the association between F. prausnitzii and irAEs; conversely, we found this species only in the NAE group. An explanation of these contradictory data could be the nonhomogeneity of irAEs other than colitis that we observed. Moreover, we speculated that the metabolites secreted by *F. prausnitzii,* blocking NF-κB activation and IL-8 secretion (18), might play a role in attenuating the manifestation of irAEs, as demonstrated by Liu et al. in patients treated with anti-PD-1 inhibitors: a higher abundance of *Faecalibacterium* was associated with mild irAEs (19).

The interpretation of these microbial signatures is complicated by inherent confounding factors, particularly the retrospective design and limited sample size. Moreover, the majority of patients had a history of chemotherapy and antibiotic prophylaxis. The interplay between antibiotics, GM modulation, and therapeutic efficacy remains a critical challenge. Chemotherapeutic agents like cyclophosphamide are known to alter GM composition, inducing the translocation of Gram-positive species to secondary lymphoid organs and promoting pathogenic Th17 tumor invasion (7,20). With this in mind, in a large German cohort of patients affected by CLL and treated with cyclophosphamide, according to Viad's findings, Pflug et al. showed that treatment with anti-Gram-positive reduced anticancer efficacy (21).

As discussed above, our study has several limitations that might impact the interpretation of our findings. Furthermore, the small sample size prevented us from reaching statistical significance. Further validation of our findings in prospective cohorts or analyses conducted in association with randomized controlled trials should be promoted. Despite these limitations, the strength of our study comes from a strict definition of patient characteristics, and, to the best of our knowledge, no previous research on this topic is available.

Conclusions

In conclusion, despite the absence of statistical significance, this study identified distinct distributional patterns of bacterial taxa across the clinical subgroups. These observations parallel the well-established evidence concerning the interplay between the GM and host immune responses in melanoma settings. Within the framework of precision medicine, further extensive research is warranted to validate the prognostic utility of GM signatures and to implement targeted therapeutic strategies, including microbiome modulation via fecal microbiota transplantation.

Acknowledgments

We are grateful for the help and time of all the staff involved in the sampling process. This work was supported by funding from Ricerca Corrente 2018, Fondazione IRCCS Policlinico San Matteo, Pavia (Grant No. 08020418).

Ethical statement

The study adhered to the STROBE guidelines for observational research (12) and received approval from the local Ethics Committee (Comitato Etico Area Pavia) and the Institutional Review Board (Protocol P-20180107349, approved 7 December 2018).

Data availability

The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/microorganisms11051305/s1, Figure S1: Taxonomic composition and diversity of Fecal Bacterial Communities for each patient treated with Ibrutinib stratified for Responder and Non-responder group; Figure S2: α-diversity. Figure S3: Beta-diversity.

References

1. Justiz Vaillant, A.A.; Stang, C.M. Lymphoproliferative Disorders; NCBI Bookshelf; StatPearls: Treasure Island, FL, USA, 2022.
2. Hallek, M.; Cheson, B.D.; Catovsky, D.; Caligaris-Cappio, F.; Dighiero, G.; Döhner, H.; Hillmen, P.; Keating, M.; Montserrat, E.; Chiorazzi, N.; et al. IwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL. Blood 2018, 131, 2745–2760.
3. Gianfelici, V.; Levato, L.; Molica, S. The Evolution of Targeted Therapies in Chronic Lymphocytic Leukaemia. Curr. Hematol. Malig. Rep. 2020, 15, 343–349.
4. de Weerdt, I.; Koopmans, S.M.; Kater, A.P.; van Gelder, M. Incidence and Management of Toxicity Associated with Ibrutinib and Idelalisib: A Practical Approach. Haematologica 2017, 102, 1629–1639.
5. Vétizou, M.; Pitt, J.M.; Daillère, R.; Lepage, P.; Waldschmitt, N.; Flament, C.; Rusakiewicz, S.; Routy, B.; Roberti, M.P.; Duong, C.P.M.; et al. Anticancer Immunotherapy by CTLA-4 Blockade Relies on the Gut Microbiota. Science 2015, 350, 1079–1084.
6. Gopalakrishnan, V.; Spencer, C.N.; Nezi, L.; Reuben, A.; Andrews, M.C.; Karpinets, T.V.; Prieto, P.A.; Vicente, D.; Hoffman, K.; Wei, S.C.;

et al. Gut Microbiome Modulates Response to Anti–PD-1 Immunotherapy in Melanoma Patients. Science 2018, 359, 97–103.
7. Uribe-Herranz, M.; Klein-González, N.; Rodríguez-Lobato, L.G.; Juan, M.; Fernández de Larrea, C. Gut Microbiota Influence in Hematological Malignancies: From Genesis to Cure. Int. J. Mol. Sci. 2021, 22, 1026.
8. Hakim, H.; Dallas, R.; Wolf, J.; Tang, L.; Schultz-Cherry, S.; Darling, V.; Johnson, C.; Karlsson, E.A.; Chang, T.-C.; Jeha, S.; et al. Gut Microbiome Composition Predicts Infection Risk During Chemotherapy in Children With Acute Lymphoblastic Leukemia. Clin. Infect. Dis. 2018, 67, 541–548.
9. Huang, J.; Jiang, Z.; Wang, Y.; Fan, X.; Cai, J.; Yao, X.; Liu, L.; Huang, J.; He, J.; Xie, C.; et al. Modulation of Gut Microbiota to Overcome Resistance to Immune Checkpoint Blockade in Cancer Immunotherapy. Curr. Opin. Pharmacol. 2020, 54, 1–10.
10. Lee, S.M.; Donaldson, G.P.; Mikulski, Z.; Boyajian, S.; Ley, K.; Mazmanian, S.K. Bacterial Colonization Factors Control Specificity and Stability of the Gut Microbiota. Nature 2013, 501, 426–429.
11. Baruch, E.N.; Youngster, I.; Ortenberg, R.; Ben-Betzalel, G.; Katz, L.H.; Lahat, A.; Barshack, I.; Dick-Necula, D.; Mamtani, R.; Bloch, N.; et al. Fecal Microbiota Transplant Promotes Response in Immunotherapy-Refractory Melanoma Patients. Science 2021, 371, 602–609.
12. Cuschieri, S. The STROBE Guidelines. Saudi J. Anaesth. 2019, 13, S31–S34.
13. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 2013, 8, e61217.
14. Oksanen, J.; Simpson, G.L.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.R.; O'Hara, R.; Solymos, P.; Stevens, M.H.; Szoecs, E.; et al. Vegan: Community Ecology Package. 2022. Available online: https://cran.r-project.org/web/packages/vegan/index.html (accessed on 20 March 2023).
15. Karpinets, T.V.; Wu, X.; Solley, T.; El Alam, M.B.; Sims, T.T.; Yoshida-Court, K.; Lynn, E.; Ahmed-Kaddar, M.; Biegert, G.; Yue, J.; et al. Metagenomes of Rectal Swabs in Larger, Advanced Stage Cervical Cancers Have Enhanced Mucus Degrading Functionalities and Distinct Taxonomic Structure. BMC Cancer 2022, 22, 945.
16. Portincasa, P.; Bonfrate, L.; Vacca, M.; De Angelis, M.; Farella, I.; Lanza, E.; Khalil, M.; Wang, D.Q.-H.; Sperandio, M.; Di Ciaula, A. Gut Microbiota and Short Chain Fatty Acids: Implications in Glucose Homeostasis. Int. J. Mol. Sci. 2022, 23, 1105.
17. Effendi, R.M.R.A.; Anshory, M.; Kalim, H.; Dwiyana, R.F.; Suwarsa, O.; Pardo, L.M.; Nijsten, T.E.C.; Thio, H.B. Akkermansia Muciniphila

and Faecalibacterium Prausnitzii in Immune-Related Diseases. Microorganisms 2022, 10, 2382.
18. Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermúdez-Humarán, L.G.; Gratadoux, J.-J.; Blugeon, S.; Bridonneau, C.; Furet, J.-P.; Corthier, G.; et al. Faecalibacterium Prausnitzii Is an Anti-Inflammatory Commensal Bacterium Identified by Gut Microbiota Analysis of Crohn Disease Patients. Proc. Natl. Acad. Sci. USA 2008, 105, 16731–16736.
19. Liu, W.; Ma, F.; Sun, B.; Liu, Y.; Tang, H.; Luo, J.; Chen, H.; Luo, Z. Intestinal Microbiome Associated With Immune-Related Adverse Events for Patients Treated With Anti-PD-1 Inhibitors, a Real-World Study. Front. Immunol. 2021, 12, 756872.
20. Viaud, S.; Saccheri, F.; Mignot, G.; Yamazaki, T.; Daillere, R.; Hannani, D.; Enot, D.P.; Pfirschke, C.; Engblom, C.; Pittet, M.J.; et al. The Intestinal Microbiota Modulates the Anticancer Immune Effects of Cyclophosphamide. Science 2013, 342, 971–976.
21. Pflug, N.; Kluth, S.; Vehreschild, J.J.; Bahlo, J.; Tacke, D.; Biehl, L.; Eichhorst, B.; Fischer, K.; Cramer, P.; Fink, A.-M.; et al. Efficacy of Antineoplastic Treatment Is Associated with the Use of Antibiotics That Modulate Intestinal Microbiota. OncoImmunology 2016, 5, e1150399.

CHAPTER 9:

General discussion

The field of bioinformatics within hospital settings is increasingly expanding and has revolutionised the field of infectious diseases in the post COVID-19 era. Many hospitals and institutions worldwide have now introduced genomic-based systems for surveillance, often acting in real-time. These systems have proven to be valid tools for high-resolution tracking of strains and deep characterization of antimicrobial resistance and virulence. They represent an advancement in clinical microbiology, guiding antibiotic stewardship at the patient level and informing the Hospital Infection Control Committee about local epidemiology to optimize patient management.

Accordingly, Chapter 3 addresses the long-term epidemiology of carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp), a leading cause of HAIs and colonisation events among hospitalized patients at IRCCS Fondazione Policlinico San Matteo (HSM, Pavia). In high-risk patients, HAIs or colonisation caused by CR-Kp can rapidly progress to sepsis; therefore, a deep understanding of local epidemiology is crucial for guiding clinical decision-making. Genomic analysis on over 500 genomes revealed a dynamic population structure characterized by the progressive replacement of historical lineages (CG258) with the emerging ST307 and more recently ST6668 clones. This finding carries a practical implication: phenotypic surveillance alone is insufficient to detect fine epidemiological shifts or, in some cases, to identify reservoirs that facilitate the re-introduction of persistent resistant strains.
Moreover, the investigation of the plasmid landscape highlighted the genomic plasticity of ST307, which harbored a diverse array of resistance determinants compared to other clones. While carbapenem resistance in this clone is primarily mediated by KPC enzymes, ST307 occasionally presents a combination of multiple determinants. This plasticity likely drives its evolutionary success and adaptability over other clones within HSM.

Chapter 4 expands the scope of genomic surveillance to encompass all carbapenem-resistant *Enterobacterales* (CRE) isolated in HSM over a one-year period. While CR-Kp confirmed itself as a primary driver of hospital transmission, characterized by the sustained intra-hospital dissemination of both major and low-prevalence clones, other CRE such as *Escherichia coli* and *Enterobacter spp.* exhibited a sporadic distribution, mostly attributable to independent introduction events. However, limiting the clonal expansion of dominant clones is insufficient to halt the spread of resistance

mechanisms. In fact, genomic analysis detected the horizontal transfer of the *bla*$_{NDM-7}$ gene via an IncX3 plasmid across distinct species (*K. pneumoniae*, *K. aerogenes*, and *E. coli*) within the same patient, as well as across different patients in the same ward.
These findings demonstrate that "minor" or sporadic CRE can act as silent reservoirs. In addition they support the implementation of a dual strategy that integrates the extensive surveillance of the CRE population to monitor and identify emerging clones before they trigger outbreaks, while concurrently tracking mobile genetic elements, thereby mitigating their spread into the bacterial population.

Moving from broad surveillance to targeted intervention, Chapters 5 and 6 demonstrate the value of high-resolution of genomics in resolving hospital outbreaks. While routine surveillance monitors trends, SNP-based investigation provides the granularity required to reconstruct transmission chains.

Chapter 5 details the emergence of a novel *K. pneumoniae* clone harboring *bla*$_{NDM-1}$ that was first identified in February 2023 within HSM and the surrounding Pavia area. Genomic analysis enabled the identification of this isolate as ST6668, alongside a comprehensive profiling of its antimicrobial resistance and virulence traits. This outbreak signaled a pivotal shift in local epidemiology, marking a transition from the historical endemicity of KPC-producers to the rising circulation of NDM-producers. While this shift is still ongoing, the specific contribution of ST6668 remains unclear. Nevertheless, current evidence confirms that this clone has spread beyond Pavia and is now affecting healthcare settings also in the Milan area [80]. This evolving scenario underscores the need to act as a network, establishing local and regional surveillance systems.

Similarly, Chapter 6 investigates a carbapenem-resistant *A. baumannii* outbreak that occurred during the COVID-19 pandemic in HSM. By employing a hybrid sequencing approach (combining short and long reads), the study generated high-quality genome assemblies that enabled in-depth plasmid characterization. This study identified four circulating ST2 subclones, two of which were responsible for the outbreak; however, no specific virulence determinants were detected to explain its persistence and diffusion. Therefore, this evidence suggests that transmission was driven by extrinsic pressures, specifically the breakdown of standard infection control practices due to the health emergency, rather than intrinsic bacterial hyper-virulence. Consequently, these findings reinforce the pivotal role of strict adherence to standard hygiene measures as the primary defense against the dissemination of high-risk clones.

Ultimately, to fully exploit the potential of clinical genomics, we must shift from descriptive epidemiology to causal functional inference. To explain the biological drivers behind the success of specific high-risk lineages, future

analytical frameworks should integrate comparative phylogenetics to quantify evolutionary rates, including plasmid transfer dynamics and accessory genome flux. Furthermore, understanding pathogen adaptation entails probing beyond standard coding sequences by evaluating per-gene selection pressures (dN/dS ratios), annotating non-coding RNAs (ncRNAs), and characterizing regulatory elements that modulate phenotypic expression. Finally, Genome-Wide Association Studies applied to bacteria can maximize the explanatory power of these datasets by statistically linking complex traits to specific genetic signatures (e.g., epistasis or cryptic mutations).

Shifting the focus from epidemiological tracking to patient management, Chapter 7 addresses the operational dimension of clinical microbiology: the timeliness of diagnosis. While genomic surveillance provides high-resolution data for epidemiological control, the immediate management of individual patients, particularly those with Bloodstream Infections (BSIs), relies both on clinicians and the speed of diagnostic output from the Clinical Microbiology Laboratory (CML).
This study evaluated the impact of transitioning the CML to a 24/7 continuous workflow for Blood culture (BC) processing, coupled with the implementation of a specific 'fast-track' pathway for positive BCs from critically ill patients. By significantly reducing Turnaround Times (TATs), especially for the initial Gram stain result, this operational efficiency bridges the gap between the CML and the bedside. This reduction in TAT directly translates to clinical efficacy. Indeed, the new model may empower clinicians to make rapid, evidence-based therapeutic decisions, which are essential for effective sepsis management. Moreover, reduced TAT may enable earlier de-escalation from empirical to targeted therapy, thereby limiting selective pressure on bacterial populations.

Finally, Chapter 8 represents a different application of NGS in clinical microbiology, shifting the focus from single-colony genomics to the characterization of the entire microbial community of the host. The study applied the targeted metagenomics (using 16S rRNA) to characterize the taxonomic composition of Gut Microbiota (GM) of patients experiencing CLL undergoing treatment with B-cell receptor inhibitors. Although the limited sample size precluded statistical significance, the GM analysis revealed detectable taxonomic trends distinguishing clinical responders from non-responders, as well as patients prone to immune-related adverse events. In the era of precision medicine, these findings highlight the translational potential of the gut microbiota as a non-invasive predictive biomarker to guide personalized interventions and improve therapeutic outcomes.

The overall diversity of the projects that I carried on during my PhD reflected the heterogeneous application of bioinformatics in clinical microbiology. By spanning across pathogen genomics, outbreak reconstruction, and metagenomics, this thesis has exemplified the integration of 'dry' computational analysis with 'wet' laboratory routine. I believe this synergy represents the key to translating data into actionable insights that improve public health.

REFERENCES

1. Otter JA, Yezli S, French GL (2011) The role played by contaminated surfaces in the transmission of nosocomial pathogens. Infect Control Hosp Epidemiol 32:687–699
2. Cruz-López F, Martínez-Meléndez A, Garza-González E (2023) How Does Hospital Microbiota Contribute to Healthcare-Associated Infections? Microorganisms. https://doi.org/10.3390/microorganisms11010192
3. European Centre for Disease Prevention and Control. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals. Stockholm: ECDC; 2024.
4. World Health Organization. Report on the burden of endemic health care-associated infection worldwide. Geneva: World Health Organization; 2011.
5. Healthcare-associated infections in European hospitals (PPS survey) 2022-2023. . https://www.ecdc.europa.eu/en/publications-data/healthcare-associated-infections-european-hospitals. Accessed 14 Jan 2026
6. Khan HA, Baig FK, Mehboob R (2017) Nosocomial infections: Epidemiology, prevention, control and surveillance. Asian Pac J Trop Biomed 7:478–482
7. Organization WH (2024) WHO bacterial priority pathogens list, 2024 bacterial pathogens of public health importance, to guide research, development, and strategies to prevent and control antimicrobial resistance. 72
8. Singer M, Deutschman CS, Seymour CW, et al (2016) The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 315:801–810
9. King J, Chenoweth CE, England PC, et al (2023) Early recognition and initial management of sepsis in adult patients. Michigan Medicine University of Michigan, Ann Arbor (MI)
10. Van Heuverswyn J, Valik JK, Desirée van der Werff S, Hedberg P,

Giske C, Nauclér P (2023) Association Between Time to Appropriate Antimicrobial Treatment and 30-day Mortality in Patients With Bloodstream Infections: A Retrospective Cohort Study. Clin Infect Dis 76:469–478

11. Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D (2020) Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019- results from a systematic review and meta-analysis. Crit Care 24:239
12. CLSI. (2022) M47 PRINCIPLES AND PROCEDURES FOR BLOOD CULTURES: 2ND EDITION CLSI. 98
13. Evans L, Rhodes A, Alhazzani W, et al (2021) Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med 47:1181–1247
14. Ho CS, Wong CTH, Aung TT, et al (2025) Antimicrobial resistance: a concise update. Lancet Microbe 6:100947
15. Davies J, Davies D (2010) Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev 74:417–433
16. Organization WH (2025) Global antibiotic resistance surveillance report 2025 WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS). 138
17. Tahmasebi H, Arjmand N, Monemi M, Babaeizad A, Alibabaei F, Alibabaei N, Bahar A, Oksenych V, Eslami M (2025) From Cure to Crisis: Understanding the Evolution of Antibiotic-Resistant Bacteria in Human Microbiota. Biomolecules. https://doi.org/10.3390/biom15010093
18. Mendelson M, Sharland M, Mpundu M (2022) Antibiotic resistance: calling time on the “silent pandemic”. JAC Antimicrob Resist 4:dlac016
19. Antimicrobial Resistance Collaborators (2022) Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399:629–655
20. GBD 2021 Antimicrobial Resistance Collaborators (2024) Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. Lancet 404:1199–1226
21. Council of the European Union Council recommendation on stepping up EU actions to combat antimicrobial resistance in a one health approach (2023/C 220/01). Council of the European Union, Brussels,

2023. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOC_2023_220_R_0001. Accessed 14 Jan 2026

22. WHO Regional Office for Europe and European Centre for Disease Prevention and Control. Surveillance of antimicrobial resistance in Europe, 2024 data: executive summary. Copenhagen: WHO Regional Office for Europe; 2025.
23. European Centre for Disease Prevention and Control. ECDC; Stockholm: 2025. Antimicrobial consumption in the EU/EEA (ESAC-Net) - annual epidemiological report 2024.
24. European Centre for Disease Prevention and Control. ECDC; Stockholm: 2025. Antimicrobial resistance in the EU/EEA (EARS-Net) - annual epidemiological report 2024.
25. Pearce N (1996) Traditional epidemiology, modern epidemiology, and public health. Am J Public Health 86:678–683
26. Fernando SA, Gray TJ, Gottlieb T (2017) Healthcare-acquired infections: prevention strategies. Intern Med J 47:1341–1351
27. Rodríguez-Acelas AL, de Abreu Almeida M, Engelman B, Cañon-Montañez W (2017) Risk factors for health care-associated infection in hospitalized adults: Systematic review and meta-analysis. Am J Infect Control 45:e149–e156
28. Dayhoff MO National Biomedical Research Foundation. Atlas of Protein Sequence and Structure, Vol. 1. Silver Spring, 1965.
29. Dayhoff MO, Schwartz RM, Orcutt BC (1978) Chapter 22: a model of evolutionary change in proteins. Atlas of Protein Sequence and Structure
30. Gauthier J, Vincent AT, Charette SJ, Derome N (2019) A brief history of bioinformatics. Brief Bioinformatics 20:1981–1996
31. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
32. Lander ES, Linton LM, Birren B, et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
33. Venter JC, Adams MD, Myers EW, et al (2001) The sequence of the human genome. Science 291:1304–1351
34. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age:

ten years of next-generation sequencing technologies. Nat Rev Genet 17:333–351

35. Rodriguez R, Krishnan Y (2023) The chemistry of next-generation sequencing. Nat Biotechnol 41:1709–1715

36. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF (2021) Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol 39:1348–1365

37. Peng K, Li C, Wang Q, Xin X, Wang Z, Li R (2025) The applications and advantages of nanopore sequencing in bacterial antimicrobial resistance surveillance and research. npj Antimicrob Resist 3:87

38. Roberts MC, Holt KE, Del Fiol G, Baccarelli AA, Allen CG (2024) Precision public health in the era of genomics and big data. Nat Med 30:1865–1873

39. Waddington C, Carey ME, Boinett CJ, Higginson E, Veeraraghavan B, Baker S (2022) Exploiting genomics to mitigate the public health impact of antimicrobial resistance. Genome Med 14:15

40. Sherry NL, Gorrie CL, Kwong JC, et al (2022) Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis. Lancet Reg Health West Pac 23:100446

41. Roberts LW, Forde BM, Hurst T, et al (2021) Genomic surveillance, characterization and intervention of a polymicrobial multidrug-resistant outbreak in critical care. Microb Genom. https://doi.org/10.1099/mgen.0.000530

42. European Centre for Disease Prevention and Control. Survey protocol for the ECDC genomic-based survey of carbapenem-resistant Acinetobacter baumannii in Europe – Version 1.3. Stockholm: ECDC; 2024.

43. European Centre for Disease Prevention and Control. ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations –2019–2021. Stockholm: ECDC; 2019.

44. Popovich KJ, Snitkin ES (2017) Whole Genome Sequencing-Implications for Infection Prevention and Outbreak Investigations. Curr Infect Dis Rep 19:15

45. Wang X, Stelzer-Braid S, Scotch M, Rawlinson WD (2022) Detection

of respiratory viruses directly from clinical samples using next-generation sequencing: A literature review of recent advances and potential for routine clinical use. Rev Med Virol 32:e2375

46. Clarridge JE (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin Microbiol Rev 17:840–62, table of contents

47. Matchado MS, Rühlemann M, Reitmeier S, Kacprowski T, Frost F, Haller D, Baumbach J, List M (2024) On the limits of 16S rRNA gene-based metagenome prediction and functional profiling. Microb Genom. https://doi.org/10.1099/mgen.0.001203

48. Naccache SN, Federman S, Veeraraghavan N, et al (2014) A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. Genome Res 24:1180–1192

49. Bağcı C, Patz S, Huson DH (2021) DIAMOND+MEGAN: fast and easy taxonomic and functional analysis of short and long microbiome sequences. Curr Protoc 1:e59

50. Alawi M, Burkhardt L, Indenbirken D, Reumann K, Christopeit M, Kröger N, Lütgehetmann M, Aepfelbacher M, Fischer N, Grundhoff A (2019) DAMIAN: an open source bioinformatics tool for fast, systematic and cohort based analysis of microorganisms in diagnostic samples. Sci Rep 9:16841

51. Miller S, Naccache SN, Samayoa E, et al (2019) Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. Genome Res 29:831–842

52. Blauwkamp TA, Thair S, Rosen MJ, et al (2019) Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nat Microbiol 4:663–674

53. Torres Montaguth OE, Buddle S, Morfopoulou S, Breuer J (2026) Clinical metagenomics for diagnosis and surveillance of viral pathogens. Nat Rev Microbiol 24:61–75

54. Chiu CY, Miller SA (2019) Clinical metagenomics. Nat Rev Genet 20:341–355

55. Mitchell SL, Simner PJ (2019) Next-Generation Sequencing in Clinical Microbiology: Are We There Yet? Clin Lab Med 39:405–418

56. Ladner JT, Grubaugh ND, Pybus OG, Andersen KG (2019) Precision

epidemiology for infectious disease control. Nat Med 25:206–211

57. Neville SA, Lecordier A, Ziochos H, Chater MJ, Gosbell IB, Maley MW, van Hal SJ (2011) Utility of matrix-assisted laser desorption ionization-time of flight mass spectrometry following introduction for routine laboratory bacterial identification. J Clin Microbiol 49:2980–2984

58. Li Y, Shan M, Zhu Z, Mao X, Yan M, Chen Y, Zhu Q, Li H, Gu B (2019) Application of MALDI-TOF MS to rapid identification of anaerobic bacteria. BMC Infect Dis 19:941

59. Shropshire WC, Hanson BM, Shelburne SA (2025) Genome-wide approaches to bacterial strain typing: a history and review of recent methodological advances. Curr Opin Infect Dis 38:329–338

60. Vinatzer BA, Tian L, Heath LS (2017) A proposal for a portal to make earth's microbial diversity easily accessible and searchable. Antonie Van Leeuwenhoek 110:1271–1279

61. Budia-Silva M, Kostyanev T, Ayala-Montaño S, Bravo-Ferrer Acosta J, Garcia-Castillo M, Cantón R, Goossens H, Rodriguez-Baño J, Grundmann H, Reuter S (2024) International and regional spread of carbapenem-resistant Klebsiella pneumoniae in Europe. Nat Commun 15:5092

62. Wyres KL, Lam MMC, Holt KE (2020) Population genomics of Klebsiella pneumoniae. Nat Rev Microbiol 18:344–359

63. Li S, Jiang G, Wang S, et al (2025) Emergence and global spread of a dominant multidrug-resistant clade within Acinetobacter baumannii. Nat Commun 16:2787

64. Shakya M, Ahmed SA, Davenport KW, Flynn MC, Lo C-C, Chain PSG (2020) Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life. Sci Rep 10:1723

65. Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL (2017) Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev 30:1015–1063

66. Bouckaert R, Vaughan TG, Barido-Sottani J, et al (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15:e1006650

67. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM

(2022) ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom. https://doi.org/10.1099/mgen.0.000748

68. Alcock BP, Huynh W, Chalil R, et al (2023) CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res 51:D690–D699

69. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q (2005) VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 33:D325-8

70. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE (2021) A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun 12:4188

71. Feldgarden M, Brover V, Gonzalez-Escalona N, et al (2021) AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci Rep 11:12728

72. Turton JF, Ward ME, Woodford N, Kaufmann ME, Pike R, Livermore DM, Pitt TL (2006) The role of ISAba1 in expression of OXA carbapenemase genes in Acinetobacter baumannii. FEMS Microbiol Lett 258:72–77

73. Martínez-Martínez L, Pascual A, Hernández-Allés S, Alvarez-Díaz D, Suárez AI, Tran J, Benedí VJ, Jacoby GA (1999) Roles of beta-lactamases and porins in activities of carbapenems and cephalosporins against Klebsiella pneumoniae. Antimicrob Agents Chemother 43:1669–1673

74. Valavarasu S, Sangu Y, Mahapatra T (2025) Prediction of antibiotic resistance from antibiotic susceptibility testing results from surveillance data using machine learning. Sci Rep 15:30509

75. Arnold BJ, Huang I-T, Hanage WP (2022) Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol 20:206–218

76. Rodríguez-Beltrán J, DelaFuente J, León-Sampedro R, MacLean RC, San Millán Á (2021) Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev Microbiol 19:347–359

77. Orlek A, Stoesser N, Anjum MF, et al (2017) Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. Front Microbiol 8:182

78. Navon-Venezia S, Kondratyeva K, Carattoli A (2017) Klebsiella pneumoniae: a major worldwide source and shuttle for antibiotic resistance. FEMS Microbiol Rev 41:252–275

79. Di Pilato V, Pollini S, Miriagou V, Rossolini GM, D'Andrea MM (2024) Carbapenem-resistant Klebsiella pneumoniae: the role of plasmids in emergence, dissemination, and evolution of a major clinical challenge. Expert Rev Anti Infect Ther 22:25–43

80. Rimoldi SG, Comandatore F, Cogliati C, Odelli S, Morelli L, Santus P, Antinori S, Dolci A (2025) New NDM-producing and blaKPC-2 harbouring Klebsiella pneumoniae ST6668 (CC147) clone in Milan, Northern Italy. J Glob Antimicrob Resist 46:126–127

ACKNOWLEDGMENT

Journal of Infection and Public Health 19 (2026) 103187

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: www.elsevier.com/locate/jiph

Original article

Empowering active surveillance of Carbapenem-resistant *Enterobacterales*: Insights from one year of genomic surveillance

Greta Petazzoni [a,b], Stefano Gaiarsa [a,*], Irene Mileto [a], Angela Kuka [a], Marta Corbella [a,*], Cristina Merla [a], Antonio Piralla [a], Marina Ramus [a], Aurora Piazza [b,c], Vittoria Mattioni Marchetti [b], Roberta Migliavacca [b], Alba Muzzi [c], Patrizia Cambieri [a], Fausto Baldanti [a,b]

[a] *Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy*
[b] *Department of Clinical, Surgical, Diagnostic, and Paediatric Sciences, University of Pavia, Pavia, Italy*
[c] *Fondazione IRCCS Policlinico San Matteo, Pavia, Italy*

ARTICLE INFO

Article history:
Received 29 October 2025
Received in revised form 20 February 2026
Accepted 23 February 2026

Keywords:
Surveillance
Genomics
CRE
Clonal diffusion
Klebsiella pneumoniae

ABSTRACT

Background: Healthcare-associated infections (HAIs) in Europe affect millions of patients annually, driving antimicrobial resistance and imposing significant economic burdens. Carbapenem-resistant Enterobacterales (CRE) are particularly concerning due to their ability to spread, making genomic surveillance essential for tracking their transmission.

Methods: From October 2022 to September 2023, we sequenced and analysed the genome of 340 CRE obtained from patients admitted to Fondazione IRCCS Policlinico San Matteo (Italy). Sequencing was performed using a short-read approach and genomes were in-silico characterized to identify Sequence Types (STs), antimicrobial resistance and virulence determinants, and plasmid content. Moreover, a core SNP-based phylogeny was performed to identify potential epidemic clusters within the same clone.

Results: Among the 340 CRE isolated from surveillance (N = 234) and other clinical specimens (N = 106), 87.7% belonged to the Klebsiella genus, primarily K. pneumoniae (CR-Kp, N = 284). Escherichia coli (CR-Ec) accounted for 8.5% (N = 29); while the remaining species were Enterobacter spp. (N = 13) and Citrobacter spp. (N = 9). bla_{KPC} and bla_{NDM} gene variants were detected in 55.9% and 25.6% of CRE, respectively. The most prevalent CR-Kp clones were ST307, ST6668, ST258, and ST512, whereas ST131 was dominant among CR-Ec. Large epidemic clusters (ECs; ≥ 5 strains in each clade) were mainly found in CR-Kp associated with ST307 and ST6668, with 80.63% of CR-Kp grouped into nine ECs. Conversely, other species were observed as sporadic cases or within transmission clusters with limited spread.

Conclusion: Our comprehensive analysis of CRE circulation at HSM revealed that their dissemination was largely driven by highly successful clones, shedding light on several of them.

© 2026 Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Healthcare-associated infections (HAIs) are among the most common complications in healthcare, affecting one in ten inpatients, with a higher prevalence in low-/middle-income countries and high-risk patients [1]. Europe records 4.5 million HAIs annually [2], leading to over 90,000 deaths and 2.5 million disability-adjusted life years. HAIs are strongly linked to antimicrobial resistance, with 71% of resistant infections occurring in healthcare settings [3]. Globally, 136 million antibiotic-resistant HAIs occur every year [4], imposing significant economic burdens of billion euros.

Carbapenem-resistant *Enterobacterales* (CRE) pose a major global threat due to their ability to spread undetected as colonisers of the

* Corresponding authors.
E-mail addresses: g.petazzoni@smatteo.pv.it (G. Petazzoni), s.gaiarsa@smatteo.pv.it (S. Gaiarsa), i.mileto@smatteo.pv.it (I. Mileto), a.kuka@smatteo.pv.it (A. Kuka), m.corbella@smatteo.pv.it (M. Corbella), c.merla@smatteo.pv.it (C. Merla), a.piralla@smatteo.pv.it (A. Piralla), marina.ramus01@universitadipavia.it (M. Ramus), aurora.piazza@unipv.it (A. Piazza), vittoria.mattionimarchetti@unipv.it (V. Mattioni Marchetti), roberta.migliavacca@unipv.it (R. Migliavacca), a.muzzi@smatteo.pv.it (A. Muzzi), p.cambieri@smatteo.pv.it (P. Cambieri), f.baldanti@smatteo.pv.it (F. Baldanti).

https://doi.org/10.1016/j.jiph.2026.103187
1876-0341/© 2026 Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

human gut. Italy is among the countries in south Europe with the highest percentage of invasive isolates resistant to carbapenem; indeed a recent report of European Centre for Disease Prevention and Control (ECDC) pointed out that 24% of invasive *Klebsiella pneumoniae* isolates in Italy were resistant to carbapenems. In this context, Italy has witnessed a rapid spread of resistance primarily driven by *Klebsiella pneumoniae* producing KPC carbapenemases [5].

Surveillance plays a crucial role in limiting the transmission of HAI-causing bacteria and thereby identifying epidemic events. Many healthcare facilities worldwide have incorporated genomic characterisation into surveillance programs, improving strain typing and outbreak tracing. Genomics has been estimated to reduce HAIs by 18–37% and to prevent thousands of related deaths [6,7].

At Fondazione IRCCS Policlinico San Matteo (HSM), a 900-bed teaching hospital in Pavia (Italy), active surveillance for CRE is conducted on patients admitted to Intensive Care Units (ICUs) or those in other wards presenting risk factors (e.g., prior hospitalization, prolonged antibiotic therapies, surgical intervention). Surveillance specimens (rectal swabs or faeces) are collected on admission and weekly, identifying patients with CRE colonisation. As a part of infection prevention and control measures for multi-drug resistant (MDR) diffusion, CRE isolated from other specimens are also monitored. From October 2022 to September 2023, whole-genome sequencing of CRE was integrated into traditional surveillance to provide a comprehensive overview of bacterial circulation. Here, the results of one year of genomic surveillance of CRE at HSM are presented

Materials and methods

Bacterial isolate selection

The study was conducted from 1st October 2022–30 th September 2023 at Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy). Surveillance specimens (rectal swabs or faeces) were plated onto selective chromogenic agar CHROMID® Carba medium (BioMérieux, Marcy-l'Etoile, France). Plates were incubated at 37°C for 16 h into WASPLab® (Copan Italia). Other clinical specimens (blood, respiratory, urine samples and non-rectal swabs) were processed following internal procedures [8]. Isolates from all specimens were identified with Matrix Assisted Laser Desorption/Ionisation Time-Of-Flight Mass Spectrometry (MALDI-TOF MS; Bruker Daltonics, Bremen, Germany) equipped with Bruker Biotyper 3.1 software. Antibiotic susceptibility was tested using the BD Phoenix NMIC-474 panel loaded on a BD Phoenix M50 automated instrument (Becton Dickinson, Franklin Lakes, NJ, USA) and Sensititre EUMDRXXF (Thermo Scientific Diagnostics B.V., The Netherlands). Minimum inhibitory concentration (MIC) values were interpreted according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints (version 12.0 [9] and 13.0 [10]) and carbapenemase production confirmed using the NG-Test® CARBA-5 lateral flow immunoassay (NG Biotech laboratories) or Xpert Carba-R (Cepheid, Italy). For this study, we considered all *Enterobacterales* isolates with MIC of meropenem > 0.125 mg/L obtained from specimens processed during diagnostic routine. The meropenem MIC cut-off was chosen since above this value, carbapenemase production should be suspected [9,10]. For each patient (aged ≥ 18 years), only the first isolate for each bacterial species and the antibiotic susceptibility profile were included, obtaining 234 CRE from surveillance and 106 CRE from other clinical specimens.

DNA extraction and sequencing

Once obtained bacterial isolates, the genomic DNA was extracted using QIAsymphony DSP Virus/Pathogen Midi Kit (QIAGEN, Germany) and sequenced using Illumina MiSeq (San Diego, California, U.S.) after Nextera XT library preparation. Short reads were assembled using Shovill (https://github.com/tseemann/shovill). Forty-four genomes were retrieved from previous work (BioProject: PRJNA1019273). Kleborate [11,12] was used to assign Sequence Types (ST) and predict resistance and virulence genes for *Klebsiella* spp. Whether a new carbapenem gene variant was predicted with Kleborate, we manually checked using BLAST. Regarding other CRE, the Pasteur multilocus sequence typing (MLST) scheme and Achtman scheme for *Escherichia coli* were used to type strains whenever feasible using mlst (https://github.com/tseemann/mlst). For all CRE with no assigned MLST, sequences were submitted to the PubMLST database (https://pubmlst.org/). ResFinder [13] and VirulenceFinder [14] were used to search for resistance and virulence genes. Finally, SerotypeFinder was used to assess the *E. coli* serotype [15] PlasmidFinder [16] was performed on the dataset to identify replicon sequences.

The P-DOR pipeline [17] was used to assess the phylogenetic relationships among strains of the same ST (n ≥ 2) based on core-SNP analysis. This approach allowed the contextualization of local isolates within the global diversity, by incorporating closely related genomes from the BV-BRC database (www.bv-brc.org). We then identified transmission clusters which were defined as monophyletic clades with bootstrap support > 80/100.

To assess the association between species and wards, as well as ST and wards, the *chi-square test* was performed, followed by post hoc analysis with Bonferroni correction to adjust for multiple comparisons in the R environment (vr 4.4.0). The significance threshold was set at 0.05.

Results

From October 2022 to September 2023, 3806 *Enterobacterales* were isolated from samples processed during routine diagnostic (patients ≥ 18 years old) in Fondazione IRCCS Policlinico San Matteo (HSM; Pavia, Italy). Of these, 51.9% (N = 1975) were beta-lactamases susceptible, 23.1% (N = 881) produced extended-spectrum beta-lactamases, and 25.0% (N = 950) were carbapenem-resistant (CRE). CRE from surveillance specimens were 737 while the remaining 213 from other specimens. We selected the first CRE per patient, obtaining 340 CRE from 294 patients. One hundred ninety-five patients were male (66.3%) with a median age of 67 years (Q1-Q3: 55–76), further details are present in Table S1 and Supplementary results 1. The majority of CRE were obtained from patients in Intensive Care Units (ICUs; N = 130; 38.2%), Infectious diseases (N = 28; 8.2%), and Pneumology (N = 22; 6.5%). Two hundred thirty-four (68.5%) CRE were isolated from surveillance specimens, 57 from urine, 20 from bloodstream, 17 from respiratory tract, while the remaining from other specimens. The 340 CRE isolates comprised 11 species: the majority were within the *Klebsiella* genus (N = 298; 87.7%), including *K. pneumoniae* (N = 284), *K. quasipneumoniae* (N = 3), *K. variicola* (N = 2), *K. aerogenes* (N = 7), and *K. oxytoca* complex (N = 2). Other less frequently isolated CRE were *Escherichia coli* (N = 29), *Enterobacter* spp. (eight *E. hormaechei*, one *E. asburiae*), and *Citrobacter* spp. (two *C. freundii*, one *C. amalonaticus*, and one *C. telavivensis*). *K. pneumoniae* was significantly associated with ICUs ($p < 0.05$), while other species were evenly distributed (Fig. 1a and b and Table S1).

High-quality genomes of the 340 CRE were obtained (Table S2). Genomic characterization revealed that 190 (55.9%) CRE carried bla_{KPC} variants, while 87 (25.6%) bla_{NDM} variants, and 22 (6.5%) bla_{VIM-1}. Six were found to harbour bla_{OXA-48}-*like* genes. The co-occurrence of multiple carbapenem resistance genes was observed in six cases of *K. pneumoniae* and one case of *E. coli* (Fig. 1c). In 32 CRE, no acquired gene associated with carbapenem resistance was found (Table S3).

For species with at least two genomes assigned to the same Sequence Type (ST), we performed core SNP-based phylogenetic

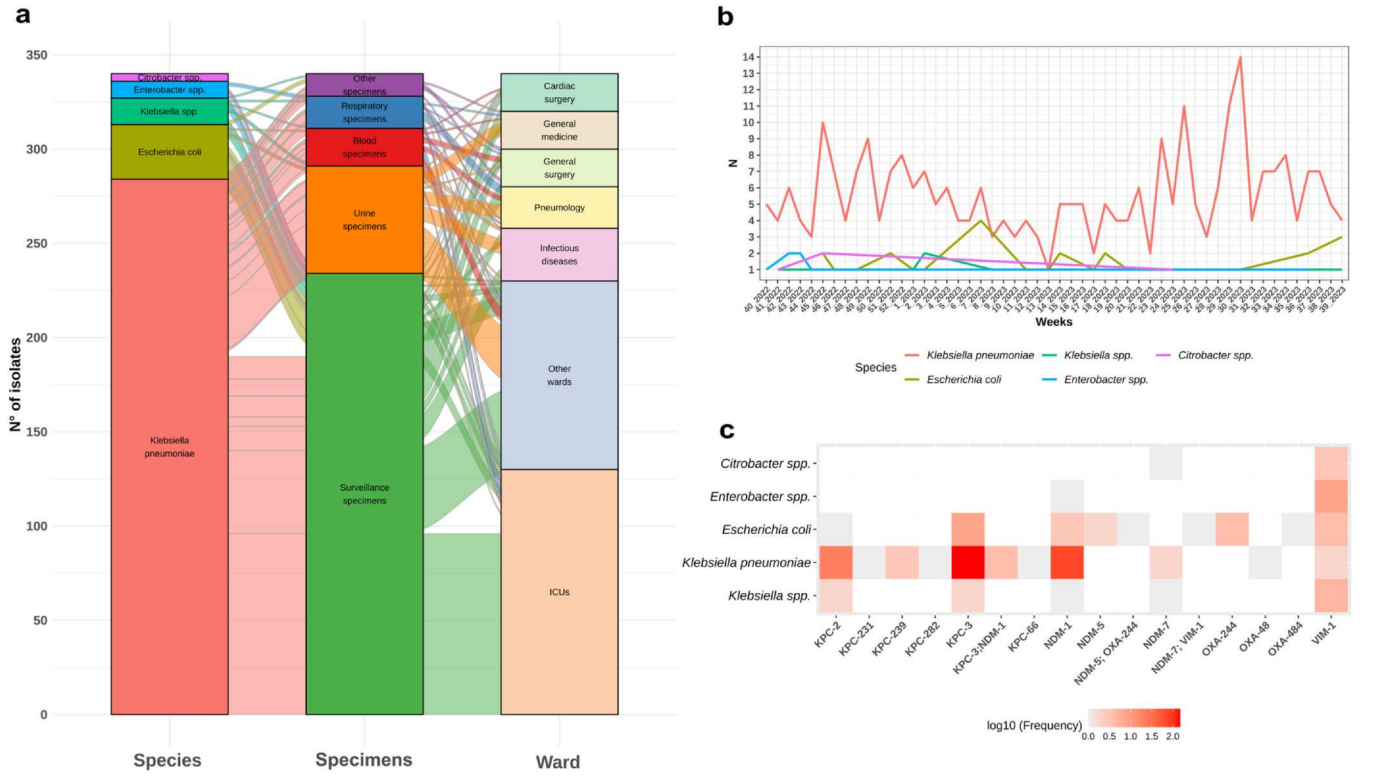


Fig. 1. Epidemiological and genomic landscape of the 340 carbapenem-resistant *Enterobacterales* (CRE) isolates. (**a**) Alluvial diagram shows the dataset characterization, illustrating the relationship between species, specimens, and wards (see Table S1 for further details). (**b**) Weekly incidence of CRE from October 1, 2022 (Week 40 of 2022), to September 31, 2023 (Week 39 of 2023), stratified by species. (**c**) Heatmap illustrates the distribution of resistance mechanisms across CRE species; color intensity reflects the log_{10}-transformed frequency.

analysis, integrating BV-BRC database genomes for evolutionary contextualisation. We identified four types of transmission: Epidemic Clusters (ECs; monophyletic clades with at least five study CRE); moderate transmission clusters (3–4 study CRE in the monophyletic clade), one-to-one transmission clusters (2 study CRE), and singletons, representing independent hospital entries with no further spread.

Klebsiella pneumoniae

Regarding carbapenem-resistant *K. pneumoniae* (CR-Kp), clones detected were ST307 (N = 130), ST6668 (N = 56), ST512 (N = 26) with four ST258, ST101 (N = 16), ST11 (N = 15), and also other less prevalent clones (Table S4). In general, the diffusion of CR-Kp was particularly prominent: 19 transmission clusters were identified. Of these, nine consisted of ECs, five were moderate transmission clusters while the other five were one-to-one transmission clusters. Additionally, 28 CR-Kp were singletons. The nine CR-Kp ECs identified comprised most of the CR-Kp strains (N = 229, 80.63%). See Fig. 2.

K. pneumoniae ST307

The ST307 core SNP phylogeny revealed three epidemic clusters (ECs): ST307_A, ST307_B, and ST307_C, along with one moderate transmission cluster (ST307_D, N= 3) and two singletons (Fig. 3).

In ST307_D, two CR-Kp strains (8447_U390 and 8404_U406) harbored bla_{NDM-7} within large genomic islands (36.309 bp and 38.246 bp), identical to a region of a complete IncX3 plasmid from *E. coli* 8501_U437 (hybrid assembly available at [18]). This *E. coli* was isolated later in time from the same patient harbouring CR-Kp 8404_U406; therefore the genomic island with bla_{NDM-7} appears to have undergone intra-patient horizontal transfer between CR-Kp and *E. coli*, likely through plasmid mediation. Conversely, the high genomic similarity (13 SNPs) between the NDM-7 CR-Kp isolated from two different patients hospitalized in the same ICU within the same month suggests an inter-patient transmission event (Fig. 3). EC ST307_A (SNPs range: 20–61) comprised nine genomes. Eight harboured bla_{KPC-3}, four of which also carried bla_{NDM-1}. One strain (8813_U449) had only bla_{NDM-1}, suggesting KPC loss. In this strain, bla_{NDM-1} was located in a ~14Kbp genomic island (14,402 bp) that was identical to a NDM-1-harbouring genomic island present in the three CR-Kp with KPC-3 +NDM-1 (8603_U455, 8494_U410, 8377_U307), while the remaining CR-Kp (8504_U391) had bla_{NDM-1} on a 5969-bp contig. These CR-Kp were isolated from patients in Pneumology (N = 3), Infectious Diseases, and Nephrology across different months (Fig. 3).

EC ST307_B (N = 30, SNPs range: 4–45) was mainly bla_{KPC-3}-positive (N = 21), whereas nine CR-Kp lacked carbapenem resistance genes. One strain exhibited an OmpK35 truncation (88% amino acid length) with $bla_{CTX-M-15}$, suggesting that phenotypic carbapenem resistance was due to a combination of factors. The absence of genomic determinants correlated with the lack of the IncFIB(pQil) plasmid. ST307_A and ST307_B formed a highly supported, Italy-exclusive clade, comprising a singleton (strain 7815_U174) and three Italian online genomes (BD_573.18714, BD_573.49061, and BD_573.49056). ST307_A isolates were mainly from June–October 2023, whereas ST307_B spanned October 2022–May 2023 (Fig. 3).

ST307_C was the largest EC (N = 84, SNPs range: 6–277), and it was predominantly bla_{KPC-3}-positive (N = 76), with three bla_{KPC-2}. Five CR-Kp strains lacked carbapenem resistance genes, correlating with IncFIB(pQil) plasmid absence. However, one of these (8626_U520) had an OmpK36 truncated. ST307_C was widely distributed across wards, particularly ICUs ($p < 0.05$), suggesting extensive ICU dissemination and potential reservoir. None of the three ECs carried hypervirulence genes (Fig. 3).

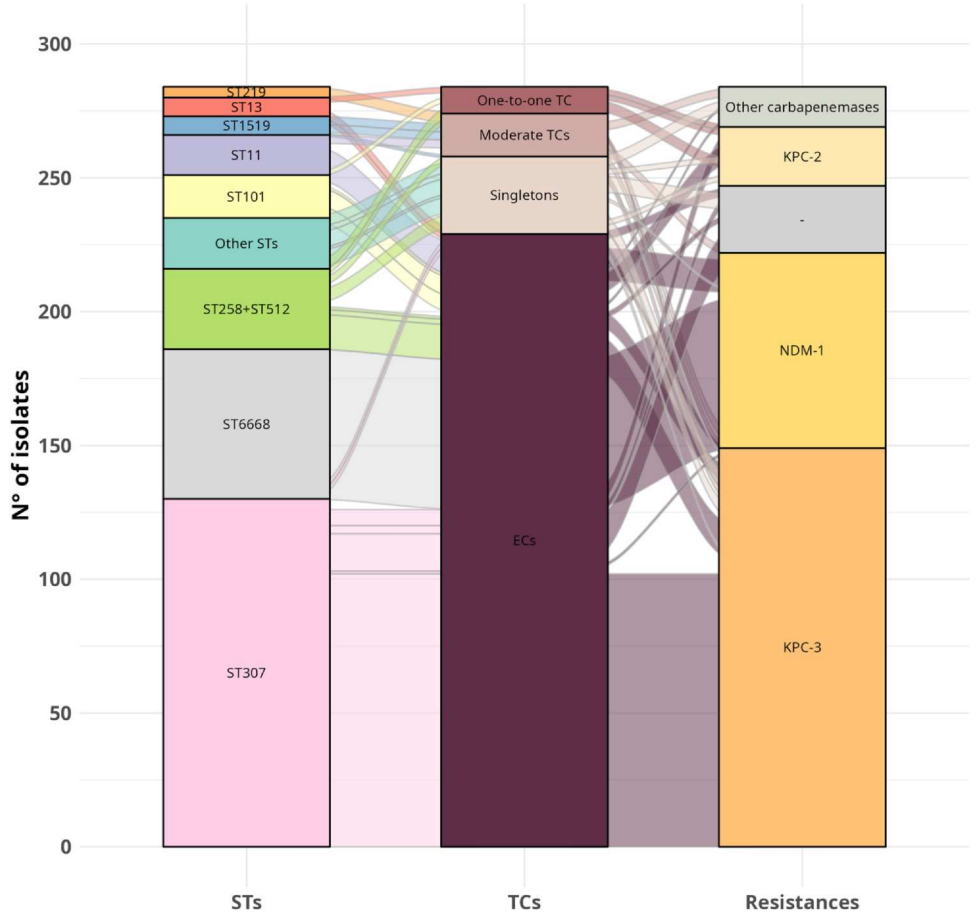


Fig. 2. Distribution of carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) strains. The alluvional plot depicts majoritarian STs (n ≥ 2 within the same ST); the category "Other STs" encompass STs detected only once. The second stratum indicates the distribution of STs across transmission clusters (TC), identified by inspecting the core-SNP phylogeny: Epidemic clusters (ECs; n ≥ 5 CR-Kp within the monophyletic clade), Moderate TCs (n = 3–4 CR-Kp within the monophyletic clade), and One-to-one TC (n = 2 in the monophyletic clade). Resistance profiles correspond to the associated carbapenemase enzymes. The category "Other carbapenemases" includes less prevalent enzymes (see Tables S3 and S4 for details).

For ST307 CR-Kp exhibiting phenotypic carbapenem resistance but no genomic resistance determinants, we hypothesised an *in vitro* loss during sample processing before sequencing (Supplementary results 2).

Other highly prevalent CR-Kp clones

K. pneumoniae ST6668 (N = 56) is a single-locus variant of ST147 (phoE: C414T) and a member of the Clonal Group (CG) 147. ST6668 was responsible for a large NDM-1 outbreak in the HSM and other nearby hospitals [19]. All ST6668 strains were monophyletic (bootstrap=100; SNPs range: 3–202; Figure S1) and carbapenem- and ceftazidime-avibactam-resistant, but retained resistance to gentamicin.

The *K. pneumoniae* part of the CG258, in particular ST258 and ST512, were divided into one EC (ST258_ST512_A; N = 16; SNP range: 8–103), one moderate transmission cluster, three one-to-one transmission clusters, and five singletons (Figure S2). Most (14/16) CR-Kp strains carried bla_{KPC-3}, whereas two Nephrology isolates lacked carbapenem resistance genes. Three CR-Kp (7642_U98, 33663_U302 and the singleton 8975_U560) were part of a highly supported larger cluster of only Italian strains. Within this cluster, there were strains involved in a previous HSM outbreak during 2015–2016 [8] (i.e. BD_573.27635, BD_573.27631, and BD_573.27630), and one CR-Kp isolated in Genoa (Italy) in 2021.

We also detected 15 *K. pneumoniae* of ST11, harbouring bla_{NDM-1} and having the gene *mgrB* mutated (conferring resistance to colistin). In HSM, most were part of a large EC (N = 12; ST11_A; SNPs range: 8–41) and carried *ybt15* for Yersiniabactin, while three formed a moderate transmission cluster with *ybt10* (Figure S3).

Finally, we also detected 7 CR-Kp ST1519, predominantly isolated in ICUs and 4/7 were bla_{KPC-3}-positive forming a moderate transmission cluster and one singleton.The remaining three harboured the $bla_{KPC-239}$ in a moderate transmission cluster and were resistant to ceftazidime-avibactam, while they remained susceptible at increased exposure to meropenem (Figure S4).

K. pneumoniae ST101 (N = 16) was divided into two ECs (ST101_A and ST101_B) and one-to-one transmission cluster (Figure S5). Notably, ST101_A CR-Kp (N = 8; SNPs range: 5–27), were mainly collected in ICUs (December 2022-August 2023). Most of them (7/8) were bla_{KPC-3}-positive except for 8817_U532 that presented a novel variant. This variant was identical at 99.7% to $bla_{KPC-162}$ (806_808delGAC) and it was found for the first time in Argentina from a ceftazidime-avibactam resistant strain; also in our case, the strain was resistant to this molecule. All CR-Kp within the EC harboured the *ybt9* for Yersiniabactin. The one-to-one transmission cluster was separated from ST101_A by a genome from the BV-BRC database (BD_573.36604 isolated in the Czech Republic) and it was composed of two CR-Kp isolated in General surgery and General medicine Units in September and December 2022. CR-Kp in EC ST101_B (N = 6; SNPs range: 15–255) carried bla_{KPC-2} and were isolated from multiple wards. These CR-Kp were associated with multidrug and hypervirulence traits, such as *iuc1* for Aerobactin (except for 8453_U401 with a partial iuc1), *ybt9* for Yersiniabactin and

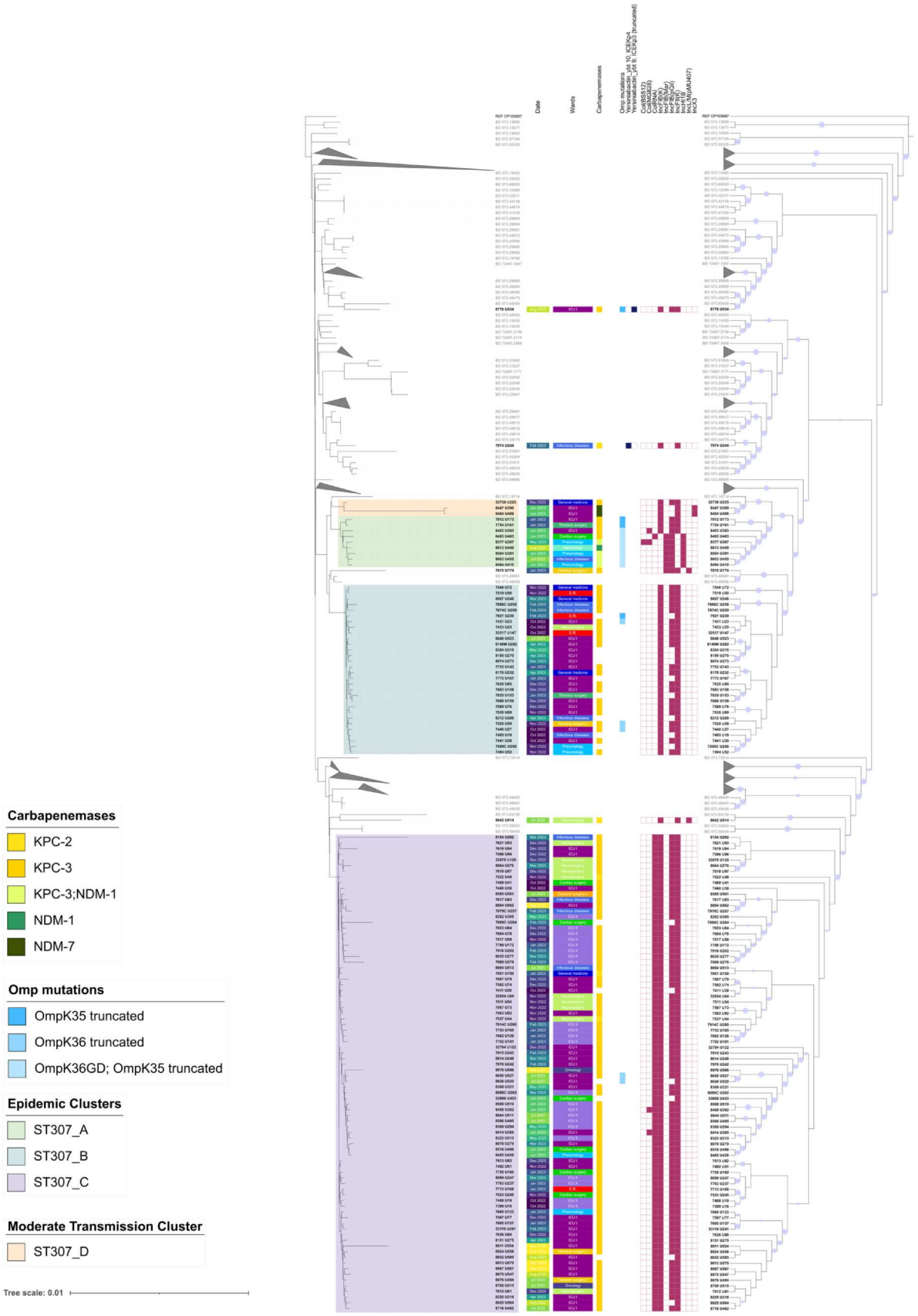


Fig. 3. Core SNP phylogeny of 310 CR-Kp ST307: 180 from the HSM (black labels) and 258 from the BV-BRC database (grey labels). Left: phylogenetic tree with branch length with date of collection, wards, carbapenemases and Omp mutations mapped. Heatmaps display the presence/absence of virulence factors (blue) and plasmid incompatibility groups (magenta). Right: tree topology on which bootstrap values (ranging from 80 to 100) are represented by light purple circles proportional to their support value. This figure was obtained using iTol (https://itol.embl.de/).

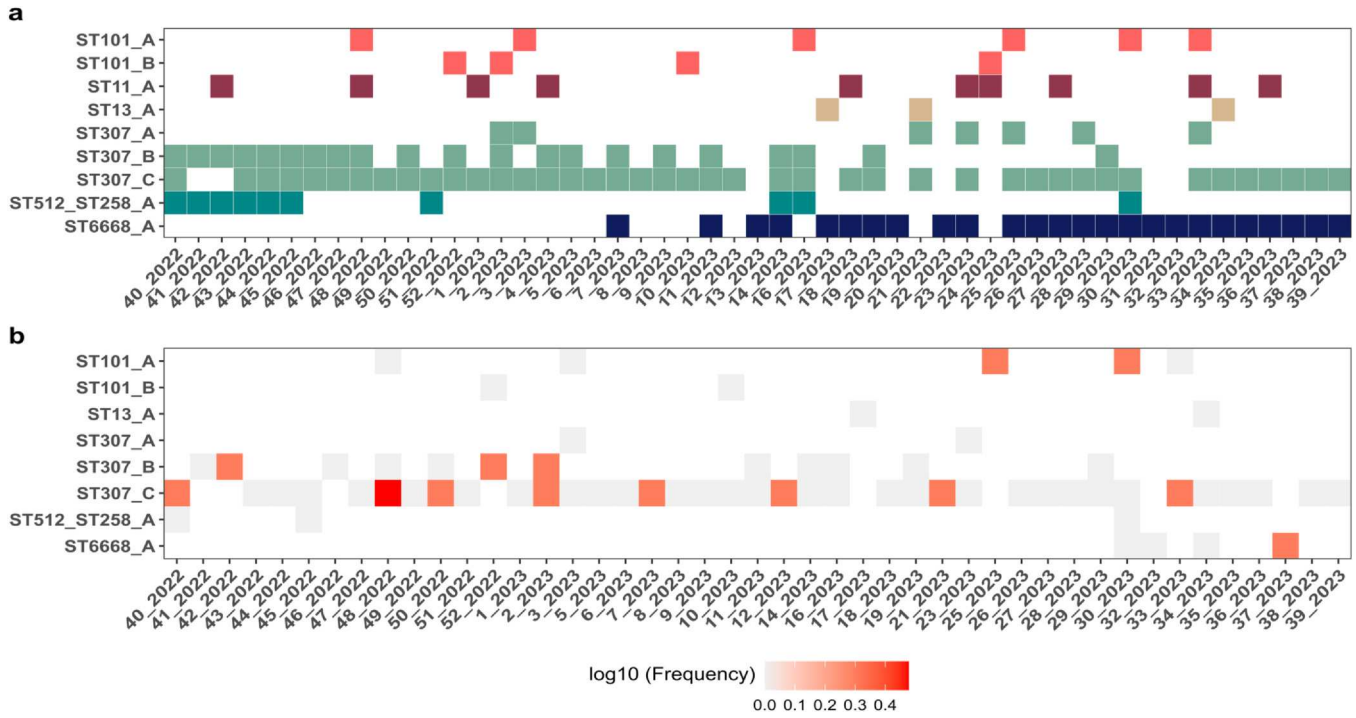


Fig. 4. Temporal distribution of Epidemic Clusters (ECs) in carbapenem-resistant *Klebsiella pneumoniae* (CR-Kp) clones. (a) Weekly presence/absence of CR-Kp ECs from October 1, 2022 (Week: 40_2022), to September 31, 2023 (Week: 39_2023), coloured by ST. (b) Heatmap illustrates the temporal distribution of each CR-Kp ECs in ICUs; color intensity reflects the log10-transformed CR-Kp frequency observed in each bin.

rmpA2 gene in two CR-Kp. The sister clade of ST101_B (N = 6) encompassed only Italian strains from the BV-BRC database, carrying bla_{KPC}–variants, *iuc1* for Aerobactin, and *ybt9* for Yersiniabactin. The aforementioned CR-Kp, together with ST101_B, were part of a broader Italian cluster of 2013–2021 CR-Kp (N = 25), supporting the Italian origin of ST101_B.

Other less prevalent clones and singleton STs were discussed in Supplementary results 3 and Figure S6 and S7.

In general, ST307 was mainly associated with the ICUs, whereas ST101 and ST11 were associated with General surgery and General medicine wards, respectively ($p < 0.05$). For the most prevalent clones, we identified nine ECs that persisted over the year and spread across HSM, particularly in ICUs (Fig. 4**a** and **b**).

Virulence factors and MGEs in K. pneumoniae

Siderophores were detected at varying prevalence: Yersiniabactin in 43.7% (124/284), Aerobactin in 3.2% (9/284), and colibactin in 2.11% (6/284) of CR-Kp, mainly in ECs ST101_A and B, ST11_A, ST13_A, ST512_ST258_A, and ST6668_A (Table S5). A truncated *rmpA2* for the hypermucoid phenotype was also found in two ST101. All KPC-Kp strains carried at least one plasmid incompatibility (Inc) group, with a median of 4 (Q1-Q3: 3–5). ST219 and ST6668 had the highest Inc counts (median: 6 [Q1-Q3: 6–6] and 5.5 [Q1-Q3: 4–6]), while ST307 and ST258/ST512 had median of 4 (IQR: 3–4) and 4 (IQR: 4–5), respectively.

Klebsiella spp.

We identified 14 additional CRE within the *Klebsiella* genus. Seven were *K. aerogenes* of diverse STs, including two novel ones (ST731, ST732; Table S6). These clones, unrelated to highly prevalent human-associated lineages (e.g., ST4, ST93), were sporadic in HSM. Two carried bla_{KPC-2}, one bla_{KPC-3}, one bla_{NDM-1}, and one bla_{NDM-7}. Notably, strain 8457_U388 harboured bla_{NDM-7} on a ~45 kb contig identical to the IncX3 plasmid sequence found in *E. coli* 8501_U437. Additionally, we detected three *K. quasipneumoniae* subsp. *quasipneumoniae* (with bla_{VIM-1}), two *K. variicola* subsp. *variicola* (with $bla_{VIM-1}$and bla_{KPC-3}), and two *Klebsiella oxytoca* complex, one carrying bla_{VIM-1} and one lacking carbapenem resistance determinants.

Escherichia coli

We identified 29 carbapenem-resistant *E. coli* (CR-Ec) strains spanning multiple STs, with ST131 (N = 10) as the most prevalent. Other detected STs included ST69, ST410, ST401, and ST58, with additional sporadic clones (Table S7). Carbapenemase genes were found in 25/29 CR-Ec (83.3%), with bla_{KPC-3} (8/9) being the most frequent among KPC variants, followed by bla_{NDM-1} (in 3 out of 7 CR-Ec), $bla_{OXA-244}$ (5/6), and bla_{VIM-t1} (N = 5). Co-occurring resistance was observed in two CR-Ec: ST10 (strain 8234M_U285) with bla_{NDM-5} and $bla_{OXA-244}$ and ST3564 (strain 8501_U437) carrying bla_{NDM-7} plus bla_{VIM-1}, recently described by Mattioni Marchetti and colleagues [18]. No association was found between clones and HSM wards.

The 10 ST131 CR-Ec strains were isolated from December 2022 to July 2023 across multiple wards. All belonged to phylogroup B2, with 9/10 showing serotype O25b:H4. However, the first CR-Ec belonged to the less common O16:H5. The most common carbapenemase was the chromosomally encoded $bla_{OXA-244}$ (N = 4), with the detection of two bla_{KPC-3} and one bla_{NDM-1}. Three strains lacked carbapenemase genes and were susceptible to Imipenem and meropenem, with MIC= 0.25 mg/L. In the core SNP phylogeny, 7/10 isolates were singletons, while three consisted of a moderate transmission cluster (Figure S8). The most common plasmid Inc were of Col-type (i.e. MG828 and 156) and IncFIB, with IncHI1B_pNDM-MAR and IncI1 in bla_{NDM-1}-positive, and IncN and IncX3 in the bla_{KPC-3} strain (further details in [20]). Other *E. coli* clones were discussed in Supplementary results 3 and Figure S9-S12.

Other species

Among other CRE, nine *Enterobacter* spp. were carbapenem-resistant: eight *E. hormaechei* and one *E. asburiae*. All but one carried bla_{VIM-1}; the remaining carried bla_{NDM-1}. Most *E. hormaechei* clones were sporadic (Table S8), except two ST190 consisting of a one-to-one transmission cluster (41 SNPs, see Figure S13). Similarly, two ST66 formed a one-to-one transmission cluster, but they were 257 SNPs apart (Figure S14).

Additionally, four carbapenem-resistant *Citrobacter* spp. were detected: two *C. freundii* (with bla_{VIM-1}), one *C. telavivensis* of newly assigned ST1266, carrying bla_{VIM-1} and one *C. amalonaticus* with bla_{NDM-7} (Table S9). The latter was isolated in June 2023 from the same patient with *K. aerogenes* 8457_U388, CR-Kp 8404_U406, and CR-Ec 8501_U437, all bla_{NDM-7}-positive. As in other NDM-7-producing CRE, in this case, four contigs covered 99% of the entire sequence (100% identity) of the IncX3 plasmid found in *E. coli* 8501_U437.

Discussion

Real-time genomics is increasingly integrated into traditional surveillance practices to reduce the spread of CRE in healthcare settings [21]. In fact, robust screening campaigns coupled with reliable sequencing technologies were essential for identifying Epidemic clusters (ECs) and singletons that could potentially evolve into larger outbreaks. During one year of surveillance at the HSM, the prevalence of CRE was 25% of the total number of *Enterobacterales*. Genomic characterisation was performed on the first CRE for each patient considering the bacterial species and antimicrobial susceptibility profile, obtaining 68.5% CRE from surveillance specimens. Few patients experienced invasive infections and nearly half were previously colonised, confirming colonisation as a key driver of transmission and infection risk [22,23].

In this study *K. pneumoniae* accounted for more than 80% of CRE and it has been confirmed as the dominant healthcare-associated pathogen in Italy. Twenty-seven STswere identified, with the majority of strains grouped into nine ECs, indicating significant intra-hospital spread. In this study ST307 was the prevalent clone, consistent with its wide diffusion across other Italian hospitals in recent years [24–26]. This emerging high-risk clone has been implicated in healthcare associated outbreaks throughout several European countries, including Portugal, Spain, France, the United Kingdom, the Netherlands, Slovenia, and Romania. Moreover ST307 has increasingly been regarded as a replacement for other high-risk clones, such as ST258, in countries such as Colombia, South Africa, and Italy [24]. At HSM ST258/ST512 has circulated since the early 2010s alongside ST307 [8,27]. Although this clone was not the most prevalent in the present study, ST258/512 remains the dominant carbapenem-resistant lineage in Europe [28].

Among the KPC-producing ST, ST101 has been frequently isolated particularly during a short period in ICUs. All the ST101 isolates of this study brought the KPC gene, differently from the emerging OXA-48 ST101 strain in Northern Africa [29,30], Middle East [31], Serbia and Romania [28].

In contrast, NDM-producing ST6668 emerged in February 2023 and spread widely in Northern Italy in a short timeframe [19,32], further narrowing patients' treatment options [33]. The other NDM-producing CR-Kp was colistin-resistant ST11 bearing bla_{NDM-1}. These strains were not linked to the ST11 clone circulating in China, which also harbors colistin resistance determinants associated with KPC-2, NDM-1, or OXA-48 carbapenemases [34].

ECs ST101_A, ST307_B, and ST307_C were ICU-associated, confirming that the ICU wards are reservoirs for resistant pathogens. Although the sampling bias of surveillance specimens may enlarge the percentage of isolations, these results align with evidence that ICUs having high antibiotic usage and fragile patients, create ideal conditions for the survival and spread of MDR pathogens [35]. In contrast, ST6668_A spread across multiple wards, suggesting diverse transmission routes within the HSM.

Notably, 3.5% CR-Kp exhibited hypervirulent traits, indicating a convergence between these two characteristics. The Yersiniabactin siderophore was found in 43.5% of the patients, reflecting the overall prevalence of the CR-Kp population. Other siderophores were detected at lower prevalence in specific ECs, such as *iro1* (3.2%) for Aerobactin in ST101_B and *clb3* (2.1%) for Colibactin in ST13_A. This evidence supports the general trend of largely non-overlapping resistant and hypervirulent populations [36].

Finally, less prevalent species included *E. coli*, *Enterobacter* spp., and *Citrobacter* spp., predominantly detected as singletons (58.9%) without evidence of HSM dissemination. Notably, the IncX3 plasmid carrying NDM-7 was identified across multiple species, despite not being associated with any major worldwide spread clones in which the plasmid has been detected [33]. Specifically, the IncX3 plasmid was observed to have likely transferred from *K. pneumoniae* to *K. aerogenes*, *C. amalonaticus*, and *E. coli* within a single patient in the ICU during June 2023. On the contrary, the other NDM-7 CR-Kp was isolated from a second patient admitted to the same ICU during the same month, suggesting inter-patient plasmid transmission.

This study has some limitations. First, only the first CRE isolate per patient was sequenced, prioritizing distinct species and resistance profiles. However, identical susceptibility profiles may arise from different resistance mechanisms, potentially introducing selection bias. Second, some genomes lacked identifiable carbapenem resistance genes, possibly due to plasmid loss during sub-culturing, as supported by *in vitro* observations of plasmid instability after a single plating. These findings underscore the need to minimize handling and sub-culturing prior to DNA extraction.

Conclusion

Although sequencing was not performed prospectively, our findings support the concept that integrating genomic data with epidemiological information is essential to fully exploit the potential of WGS for infection prevention and control, as epidemiological data provide critical temporal and spatial context. The effective management of CRE demands a dual approach: controlling the proliferation of dominant clones while curtailing the dissemination of sporadic, emerging strains. To achieve this in practice, future implementation would require full integration of the WGS pipeline within the hospital laboratory information system, allowing timely generation of clinically meaningful reports to share and discuss with clinicians and the Hospital Infection Committee. Despite the technical, infrastructural, and communication challenges associated with implementing WGS surveillance in healthcare settings, the substantial associated costs, the use of WGS in MDR surveillance is justified by the potential system-wide savings and the enhanced prevention of healthcare-associated infections, particularly among vulnerable patient populations.

Abbreviations

BV-BRC: Bacterial and Viral Bioinformatics Resource Center
CG: Clonal Group
CR-Ec: Carbapenem-resistant *Escherichia coli*
CR-Kp: Carbapenem-resistant *Klebsiella pneumoniae*
CRE: Carbapenem-resistant *Enterobacterales*
ECs: Epidemic Clusters
ECDC: European Centre for Disease Prevention and Control
EUCAST: European Committee on Antimicrobial Susceptibility Testing
HAIs: Healthcare-associated infections

HSM: Fondazione IRCCS Policlinico San Matteo (Pavia, Italy)
ICUs: Intensive Care Units
Inc: Incompatibility group
MIC: Minimum inhibitory concentration
SNP: Single Nucleotide Polymorphism
ST: Sequence Type

Authors' contributions

APir, FB, PC, MC, and SG conceptualised the study. PC, RM, and FB administered the project. IM, AK, CM, MR, VMM, and APia performed laboratory investigations. GP and SG performed computational investigations. GP performed statistical analysis. GP, VMM, APia, and SG performed bioinformatic analysis. GP, SG, IM, AK, MC, CM, APir, MR, APia, and VMM curated the data. GP, SG, IM, AK, CM, MR, VMM, and APia wrote the manuscript. GP, SG, IM, AK, MC, CM, APir, MR, APia, VMM, RM, AM, CP, and FB reviewed and edited the manuscript.

Ethics approval

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023–3.11/105). The work described herein is a study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

Funding

This work was supported by NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project PE00000007, INF-ACT), by Ricerca Corrente n° 08075723 (1034-rcr2023–43), and by JPIAMR 2024_ IMPACT 197.

Data Availability

Genome sequence data analyzed in this paper are available under BioProject PRJNA1266098 and PRJNA1019273 in the National Center for Biotechnology Information.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors declare no usage of Artificial Intelligence.

Acknowledgement

Not applicable.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jiph.2026.103187.

References

[1] Key facts and figures on World Hand Hygiene Day [Internet]. [cited 2025 Jan 30];Available from: ⟨https://www.who.int/campaigns/world-hand-hygiene-day/key-facts-and-figures⟩.
[2] European Centre for Disease Prevention and Control. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals. Stockholm: ECDC; 2024.
[3] Healthcare-associated infections [Internet]. [cited 2024 Nov 4];Available from: ⟨https://www.ecdc.europa.eu/en/healthcare-associated-infections⟩.
[4] Balasubramanian R, Van Boeckel TP, Carmeli Y, Cosgrove S, Laxminarayan R. Global incidence in hospital-associated infections resistant to antibiotics: an analysis of point prevalence surveys from 99 countries. PLoS Med 2023;20(6):e1004178.
[5] European Centre for Disease Prevention and Control. ECDC country visit to Italy to discuss antimicrobial resistance issues. Stockholm: ECDC; 2017.
[6] Fox JM, Saunders NJ, Jerwood SH. Economic and health impact modelling of a whole genome sequencing-led intervention strategy for bacterial healthcare-associated infections for England and for the USA. Micro Genom 2023;9(8).
[7] Gordon LG, Elliott TM, Forde B, Mitchell B, Russo PL, Paterson DL, et al. Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant bacterial pathogens in Queensland, Australia. BMJ Open 2021;11(2):e041968.
[8] Ferrari C, Corbella M, Gaiarsa S, Comandatore F, Scaltriti E, Bandi C, et al. Multiple Klebsiella pneumoniae KPC Clones Contribute to an Extended Hospital Outbreak. Front Microbiol 2019;10:2767.
[9] The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 2022;12(0) [Internet]. [cited 2024 Oct 15];Available from⟨http://www.eucast.org⟩.
[10] The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 2023;13(1)⟨http://www.eucast.org⟩ [Internet]. [cited 2024 Oct 15];Available from.
[11] Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun 2021;12(1):4188.
[12] Wyres KL, Wick RR, Gorrie C, Jenney A, Follador R, Thomson NR, et al. Identification of Klebsiella capsule synthesis loci from whole genome data. Micro Genom 2016;2(12):e000102.
[13] Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 2020;75(12):3491–500.
[14] Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 2014;52(5):1501–10.
[15] Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy in silico serotyping of Escherichia coli isolates by use of whole-genome sequencing data. J Clin Microbiol 2015;53(8):2410–26.
[16] Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 2014;58(7):3895–903.
[17] Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics 2023;39(9).
[18] Mattioni Marchetti V, Corbella M, Piazza A, Gaiarsa S, Mileto I, Merla C, et al. Polyclonal carbapenemase-producing *Escherichia coli* in Northern Italy: the emergence of NDM-7. Front Cell Infect Microbiol 2025;15.
[19] Mileto I, Petazzoni G, Corbella M, Gaiarsa S, Merla C, Kuka A, et al. Rapid spread of a novel NDM-producing clone of Klebsiella pneumoniae CC147, Northern Italy, February to August 2023. Eur Surveill 2023;28(42).
[20] Piazza A, Corbella M, Mattioni Marchetti V, Merla C, Mileto I, Kuka A, et al. Clinical isolates of ST131 blaOXA-244-positive Escherichia coli, Italy, December 2022 to July 2023. Eur Surveill 2024;29(8).
[21] Baker KS, Jauneikaite E, Hopkins KL, Lo SW, Sánchez-Busó L, Getino M, et al. Genomics for public health and international surveillance of antimicrobial resistance. Lancet Microbe 2023;4(12):e1047–55.
[22] Xiao Y, Duan J, Tan C, Zou J, Chen S, Liu T, et al. Correlation between intestinal CRE colonization and consequent systemic infection in hospitalized patients in China. Sci Rep 2024;14(1):26017.
[23] David S, Reuter S, Harris SR, Glasner C, Feltwell T, Argimon S, et al. Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. Nat Microbiol 2019;4(11):1919–29.
[24] Loconsole D, Accogli M, De Robertis AL, Capozzi L, Bianco A, Morea A, et al. Emerging high-risk ST101 and ST307 carbapenem-resistant Klebsiella pneumoniae clones from bloodstream infections in Southern Italy. Ann Clin Microbiol Antimicrob 2020;19(1):24.
[25] Magi G, Tontarelli F, Caucci S, Sante LD, Brenciani A, Morroni G, et al. High prevalence of carbapenem-resistant Klebsiella pneumoniae ST307 recovered from fecal samples in an Italian hospital. Future Microbiol 2021;16:703–11.
[26] Fox V, Mangioni D, Renica S, Comelli A, Teri A, Zatelli M, et al. Genomic characterization of Klebsiella pneumoniae carbapenemase-producing Klebsiella pneumoniae (KPC-Kp) strains circulating in three university hospitals in Northern Italy over three years. Antimicrob Resist Infect Control 2024;13(1):70.
[27] Rossi M, Chatenoud L, Gona F, Sala I, Nattino G, D'Antonio A, et al. Characteristics and Clinical Implications of Carbapenemase-Producing Klebsiella pneumoniae Colonization and Infection, Italy. Emerg Infect Dis 2021;27(5):1416–26.
[28] Budia-Silva M, Kostyanev T, Ayala-Montaño S, Bravo-Ferrer Acosta J, Garcia-Castillo M, Cantón R, et al. International and regional spread of carbapenem-resistant Klebsiella pneumoniae in Europe. Nat Commun 2024;15(1):5092.
[29] Kedišaletše M, Phumuzile D, Angela D, Andrew W, Mae N-F. Epidemiology, risk factors, and clinical outcomes of carbapenem-resistant Enterobacterales in Africa: a systematic review. J Glob Antimicrob Resist 2023;35:297–306.

[30] Gamaleldin P, Alseqely M, Evans BA, Omar H, Abouelfetouh A. Comparison of genotypic features between two groups of antibiotic resistant Klebsiella pneumoniae clinical isolates obtained before and after the COVID-19 pandemic from Egypt. BMC Genom 2024;25(1):983.
[31] Idrees EK, Alkhulaifi MM, Aldriwesh MG, Alqurainy N, Okdah L, Alswaji AA, et al. Genomic characterization and resistance mechanisms of carbapenem-resistant Klebsiella pneumoniae ST101 isolates from Saudi Arabia. Int J Mol Sci 2025;26(23).
[32] Rimoldi SG, Comandatore F, Cogliati C, Odelli S, Morelli L, Santus P, et al. New NDM-producing and blaKPC-2 harbouring Klebsiella pneumoniae ST6668 (CC147) clone in Milan, Northern Italy. J Glob Antimicrob Resist 2025;46:126–7.
[33] Wu W, Feng Y, Tang G, Qiao F, McNally A, Zong Z. NDM metallo-β-lactamases and their bacterial producers in health care settings. Clin Microbiol Rev 2019;32(2).
[34] Liao W, Liu Y, Zhang W. Virulence evolution, molecular mechanisms of resistance and prevalence of ST11 carbapenem-resistant Klebsiella pneumoniae in China: a review over the last 10 years. J Glob Antimicrob Resist 2020;23:174–80.
[35] Lee C-R, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global dissemination of carbapenemase-producing Klebsiella pneumoniae: epidemiology, genetic context, treatment options, and detection methods. Front Microbiol 2016;7:895.
[36] Wyres KL, Lam MMC, Holt KE. Population genomics of Klebsiella pneumoniae. Nat Rev Microbiol 2020;18(6):344–59.

Check for updates

Received: 20 September 2025 | Accepted: 9 February 2026

DOI: 10.1111/eci.70188

ORIGINAL ARTICLE

WILEY

Infectious seeds of valve calcification: Exploring the bacterial hypothesis in the pathogenesis of calcific aortic valve disease

Antonella Forlino[1,2] | **Paola Giordani**[3] | **Cristina Merla**[4] | **Silvia Roda**[5] | **Roberta Besio**[1] | **Abeer Ahmed Qaed Ahmed**[1] | **Angela Kuka**[4] | **Irene Mileto**[4] | **Greta Petazzoni**[4] | **Marta Corbella**[4] | **Patrizia Cambieri**[4] | **Baldanti Fausto**[4] | **Annalisa De Silvestri**[2] | **Totaro Pasquale**[6] | **Filippo Amoroso**[6] | **Pelenghi Stefano**[6] | **Maraschi Federica**[7] | **Eloisa Arbustini**[8] | **Viviana Vilardo**[8] | **Alexandra Smirnova**[8] | **Raffaele Bruno**[3,9] | **Seminari Elena**[3]

[1]Department of Molecular Medicine, Biochemistry Unit, University of Pavia, Pavia, Italy

[2]Biostatistics and Clinical Epidemiology, Fondazione IRCCS Policlinico san Matteo, Pavia, Italy

[3]Malattie Infettive, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

[4]Microbiologia e Virologia, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

[5]Infectious Diseases Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[6]Cardiac Surgery Unit, Fondazione IRCCS Policlinico san Matteo, Pavia, Italy

[7]Department of Chemistry, University of Pavia, Pavia, Italy

[8]Department of Research, Centre for Inherited Cardiovascular Diseases, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

[9]Department of Medicine, Surgical, Diagnostic and Pediatric Science, University of Pavia, Pavia, Italy

Correspondence
Seminari Elena, Malattie Infettive, Fondazione IRCCS Policlinico San

Abstract

Background: Calcific aortic valve disease (CAVD) is a multifactorial condition characterized by progressive leaflet calcification with a potential role for bacterial colonization in its pathogenesis. This study investigates clinical, microbiological and molecular features of calcified versus non-calcified aortic regurgitation (AR) valves.

Methods: This is a prospective, observational study, whose primary objective was to compare the occurrence of bacterial detection between CAVD and AR. The secondary objectives included the evaluation of bone-related calcification markers in valves from CAVD and AR patients.

Results: We analysed 31 CAVD and 8 AR valves, yielding 111 leaflets (84 calcified, 27 non-calcified). Light microscopy of CAVD leaflets revealed near-complete disruption of the three-layered valve architecture, with calcified masses extending through the leaflets, sparse cellularity and focal micro-angiogenesis; no bacteria were detected by GRAM, PAS or TEM. Enrichment culture detected low-virulence bacteria in 5.95% of CAVD and 4.16% of AR leaflets; 16S rRNA PCR was positive in 22.5% of CAVD and 12.5% of AR cases, with *Staphylococcus* and *Streptococcus* spp. predominating. Calcium content was significantly higher in CAVD leaflets ($p = .001$) and correlated with dyslipidemia ($p = .02$). Osterix expression was higher in valves with positive microbiological findings ($p < .0001$), while ALP was increased in CAVD and bicuspid valves regardless of microbial status. Valve interstitial cells from CAVD exhibited spontaneous in vitro calcification, unlike controls.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *European Journal of Clinical Investigation* published by John Wiley & Sons Ltd on behalf of Stichting European Society for Clinical Investigation Journal Foundation.

Matteo, Pavia, Italy.
Email: e.seminari@smatteo.pv.it

Funding information
Ministero della Salute, Grant/Award Number: 08072322, 2022 23683266 and E03 2022 23683266 DIAGNOSTICA AVANZATA - (HLS-DA)

Conclusion: The early osteogenic marker osterix was found to be upregulated in patients whose valves tested positive for microbial DNA, suggesting a potential role for bacteria in driving cellular differentiation towards an osteoblastic phenotype in CAVD.

KEYWORDS
calcific aortic valve disease, osteoblast, osterix, staphylococcus spp., streptococcus spp., VIC

1 | INTRODUCTION

Calcific aortic valve disease (CAVD) is highly prevalent worldwide, affecting nearly 50% of individuals over the age of 85. CAVD is characterized by slowly progressive fibro-calcific remodelling of the valve leaflets causing aortic stenosis.[1] Risk factors for CAVD partly overlap those for atherosclerosis, but also include age-related tissue changes and the effects of comorbidities (i.e. renal failure).[2–4] The spectrum of the disease progression starts with leaflet degeneration and progresses from early lesions to valve stenosis/obstruction.[1,5] Calcific aortic stenosis is characterized by osteogenic metaplasia with osteoblast-like cells and chondrocytes associated with dense inflammatory infiltrates.[6] Bacteria have been detected in calcific aortic valves in the absence of diagnosis of acute infective endocarditis, but their role in the pathogenesis remains unclear.[7,8] Indeed, it has been recently demonstrated that bacterial infections can directly affect osteoblast differentiation/activation.[9]

This study aimed to assess whether calcified aortic valves harbour bacterial infiltration and to explore its potential correlation with ectopic calcification and the upregulation of osteogenic markers.

2 | METHODS

2.1 | Design of the study

This is a prospective, observational study conducted at Fondazione IRCCS Policlinico San Matteo between 15 May 2023 and 31 July 2024.

To investigate the infective, biochemical and structural features of CAVD, we established a multidisciplinary team comprehensively addressing the clinical, histological, cellular, genetic and molecular features.

The case series was constituted of native aortic valves removed from patients with CAVD (stage III and IV) who underwent aortic valve replacement in accordance with International Guidelines.[10] The control series is constituted of native non-calcified aortic valves removed from patients who underwent surgery for aortic regurgitation (AR). None of the patients had a diagnosis of infective endocarditis according to the modified Duke criteria.[11]

The primary objective of the study was to compare the occurrence of bacterial detection between CAVD and AR patients. The secondary objectives included the evaluation of bone-related calcification markers in calcified and non-calcified cusps, and the comparison of these markers in valve tissues from CAVD and AR patients.

Patient's risk factors evaluated at baseline included gender, age, the diagnosis of bicuspid aortic valve, diabetes, kidney disease, cardiovascular concomitant diseases, cancer, chronic liver disease, chronic kidney disease, smoking habit, anticoagulant or antiplatelet therapy. Markers of lipid metabolism (LDL, HDL, triglycerides), cell blood count and differential and fasting glucose level were evaluated at baseline and correlated with the expression of valve osteogenic markers and total calcium valve content.

The study has been approved by the local Ethic Committee (study approbation number 0025690/23) and has been published on clinicaltrial.gov (NCT05806411).

2.2 | Valve collection

At the time of surgery, patients' valves were explanted and immediately processed under a sterile hood. Macroscopic evaluation was performed and documented with photographs. Each leaflet was then dissected into four sections within one hour of removal using a scalpel for subsequent microbiological, histological, biochemical and calcium quantification analyses. The three cusps of the valve (or two in case of bicuspid aorta) were grouped according to their anatomy (left coronary LC, right coronary RC, non-coronary NC); each leaflet was macroscopically described and assigned a consecutive identification code. The description included an evaluation of calcification in the hinge or commissural regions, or both.[12]

Samples for direct and enrichment cultures were immediately processed. Samples for calcium quantification were stored at −80°C, samples for RNA extraction were

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

collected in Qiazol (Qiagen) and stored at −80°C. Samples for bacterial DNA assessment were immediately stored at −80°C before DNA extraction.

Details on sample processing (microbiology, pathology, expression analyses of osteogenic markers, calcium determination, cell culture and VICs staining) are provided in Appendix S1.

2.3 | Statistical analysis

The sample size was calculated considering a probability of bacterial detection among sampled control of 10%,[13] a sample of 31 patients and 8 controls could be enrolled to achieve 80% power to detect as significant a proportion of 50% in cases versus the alternative of equal proportions using a chi-square test with a .05 significance level. The statistical difference in calcium amount and relative expression between pooled data from CAVD and AR leaflets across the three analysed regions was evaluated using a *t*-test. A *p*-value of less than .05 was considered statistically significant. As 9 out of 31 patients had a bicuspid aorta, this group of patients was considered in the analysis as a subgroup.

Quantitative markers were compared among groups with multilevel models to take into account the clustered nature of the data with linear regression models having patients and leaflets as random factors. Results are expressed as beta coefficient and 95% confidence interval (95% CI). Dependent variables considered in the analysis were calcium content and expression of early (osterix and COL1A1) and late (ALP, BGLAP) osteoblast markers. Covariates included in the analysis included gender and age, group (cases or controls or bicuspid aorta), aortic leaflet (left coronary or right coronary or non-coronary), microbiologic findings (leaflets with a PCR result positive for bacterial RNA 16S gene were considered as positive in the analysis), chronic kidney disease, diabetes and dyslipidemia.

3 | RESULTS

3.1 | Clinical and demographic characteristics

The study included 31 patients with CAVD and eight controls; among the controls, four had aortic bulb dilatation, two had aortic regurgitation following subaortic ridge resection, one had valvular degeneration, and one had combined mitral–aortic–tricuspid valvular disease. Among the 31 CAVD valves, nine were bicuspid (29%), resulting in 108 valve leaflets, of which 84 were from CAVD and 24 from AR patients. Calcification was mainly localized at the commissural level in 12 leaflets, at the hinge in 21 leaflets and in both sites in 51 leaflets. Localized foci of calcification were randomly observed also in AR patients. No visible sign of infection was observed during surgical procedures either in cases or in controls. Clinical and demographic data are reported in Table 1. Cases and controls were comparable for body mass index (BMI) and clinical characteristics, while dyslipidemia was more common in CAVD than in AR patients (54.8% vs. 12.5%, $p = .03$) and fast glucose level (108 vs. 99.5 mg/dL, $p = .05$) and total WBC (8505 vs. 6540 cells/mmc, $p = .02$) were higher in CAVD than in AR patients.

3.2 | Histopathology

In CAVD samples, light microscopy showed near-complete loss of the structural integrity of the three layers of the semilunar cusps in both tricuspid and bicuspid valves. In all cases, calcified masses extended into the thickness of the leaflets (Figure 1) or protruded from the leaflet profile. No superficial thrombotic stratifications were observed. GRAM and PAS staining did not reveal the presence of bacteria in any of the leaflets examined. The non-calcified areas showed structural disruption due to fibrosclerosis or focal, non-specific myxoid patterns. Cellularity in these areas was sparse and predominantly composed of fibroblasts and smooth muscle cells, with scattered chronic inflammatory cells, mainly macrophages and rare lymphocytes (indicative of very low-grade inflammation). Focal micro-angiogenesis was also observed in the non-calcified areas. Despite the presence of substantial calcium deposits, no areas of cartilage or bone metaplasia were detected.

At transmission electron microscopy (Figure 2) a spectrum of dense, compact, amorphous calcifications and calcified collagen fibrils was variably present in all calcific deposits; calcified bacterial bodies were not observed. Osteoblasts and chondrocytes were also not detected.

In AR samples light microscopy (Figure 3) showed remodelling of the valve structure, with dense fibrosis/fibrosclerosis alternating with areas of loose fibrosis, occasional adipocytes and rare lipid droplets. Cellularity was low, consisting mainly of fibroblasts and smooth muscle cells. The sparse presence of inflammatory cells (macrophages and small lymphocytes) was common in all samples. Von Kossa-positive microcalcifications were variably scattered throughout the cusps. GRAM and PAS staining tested negative. Transmission electron microscopy

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Epidemiological characteristics of cases and controls.

	Calcific aortic valve disease patients (cases)	Aortic regurgitation (controls)	*p*-Value
Males, *n* (%)	20 (64.5%)	5 (62.5%)	.9
Age (mean, range)	72 (67–75)	66 (61–79.5)	.7
BMI	27 (24–30)	25 (22–26.5)	.2
Hypertension	23 (74.2%)	6 (75.0%)	1
Diabetes	11 (35.5%)	1 (12.5%)	.2
Chronic kidney disease	2 (6.5%)	1 (12.5%)	.6
Cancer	6 (19.4%)	1 (12.5%)	.7
Chronic liver disease	2 (6.5%)	0	.5
Smoking	10 (32.3%)	1 (12.5%)	.3
Dyslipidemia on treatment	17 (54.8%)	1 (12.5%)	.03
Anticoagulant therapy	6 (19.4%)	2 (25.0%)	.7
Antiplatelet therapy	16 (51.6%)	3 (37.5%)	.5
Metformin	3 (9.7%)	1 (12.5%)	.8
Other hypoglycemic agents	7 (22.5%)	1 (12.5%)	.5
Aortic stenosis	22 (70.9%)	0	.000
Aortic insufficiency	1 (3.2%)	7 (87.5%)	.000
Stenoinsufficiency	8 (25.8%)	1 (12.5%)	.000
Coronary Artery Disease (CAD)	30 (96.8%)	7 (87.5%)	.29
WBC	8505 (6480–9910)	6540 (5325–7440)	.02
Neutrophils	5290 (4180–7130)	3890 (3500–5290)	.12
Lymphocytes	1730 (1450–2510)	1490 (1290–1685)	.06
Blood glucose level	108 (96–135)	99.5 (80–105)	.05
Triglycerides	106 (85–143)	95 (65–134)	.21
Cholesterol	154 (114–215)	148 (128.5–198)	.9

revealed microcalcifications involving collagen and elastic fibres, amorphous matrix, lipids and cellular debris, with variable density and distribution. No calcified bacteria were identified in the control cusps.

3.3 | Microbiology

3.3.1 | Direct and enrichment culture

Direct culture of all 108 valve leaflets yielded negative results, whereas enrichment culture was positive for five leaflets from 31 CAVD patients (5.95%) and for three leaflets from one single non-calcific valve out of eight (4.16%). Coagulase-negative *Staphylococci* were isolated from the enrichment culture of four CAVD leaflets (in two cases *S. warneri* and in other two cases *S. pettenkoferi*), and *Streptococcus oralis* was isolated from the enrichment culture of one CAVD leaflet. Furthermore, *S. oralis* was isolated from the three leaflets of a single AR valve.

3.3.2 | 16S rRNA sequencing detected low-virulent bacteria

PCR amplification of the V3-V5 region of 16S rRNA was positive in 7 of the 31 CAVD valves (22.5%) and in 1 valve (3 leaflets) in the control valves (12.5%) ($p = .5$). In detail, eight leaflets from seven calcific valves (8 out of 84 leaflets, 9.52%) and three leaflets from a single patient in the control group (3 out of 24 leaflets, 12.5%) tested positive. PCR amplification of the V3-V5 region was negative for all the negative controls, warranting the sterility of the amplification procedure.

Therefore, the V3-V4 region of 16S rRNA gene of eight CAVD leaflets and three AR leaflets was sequenced. A total of 3,996,077 ASVs were retrieved from sequencing, and after the removal of non-bacterial ASVs, 3,822,263 ASVs (95.65%) were available for further analysis. The relative abundance of bacterial phyla, classes, orders and families for both cases and controls is shown respectively in Figures S1–S4, respectively. *Firmicutes* was the predominant phylum

13652262, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

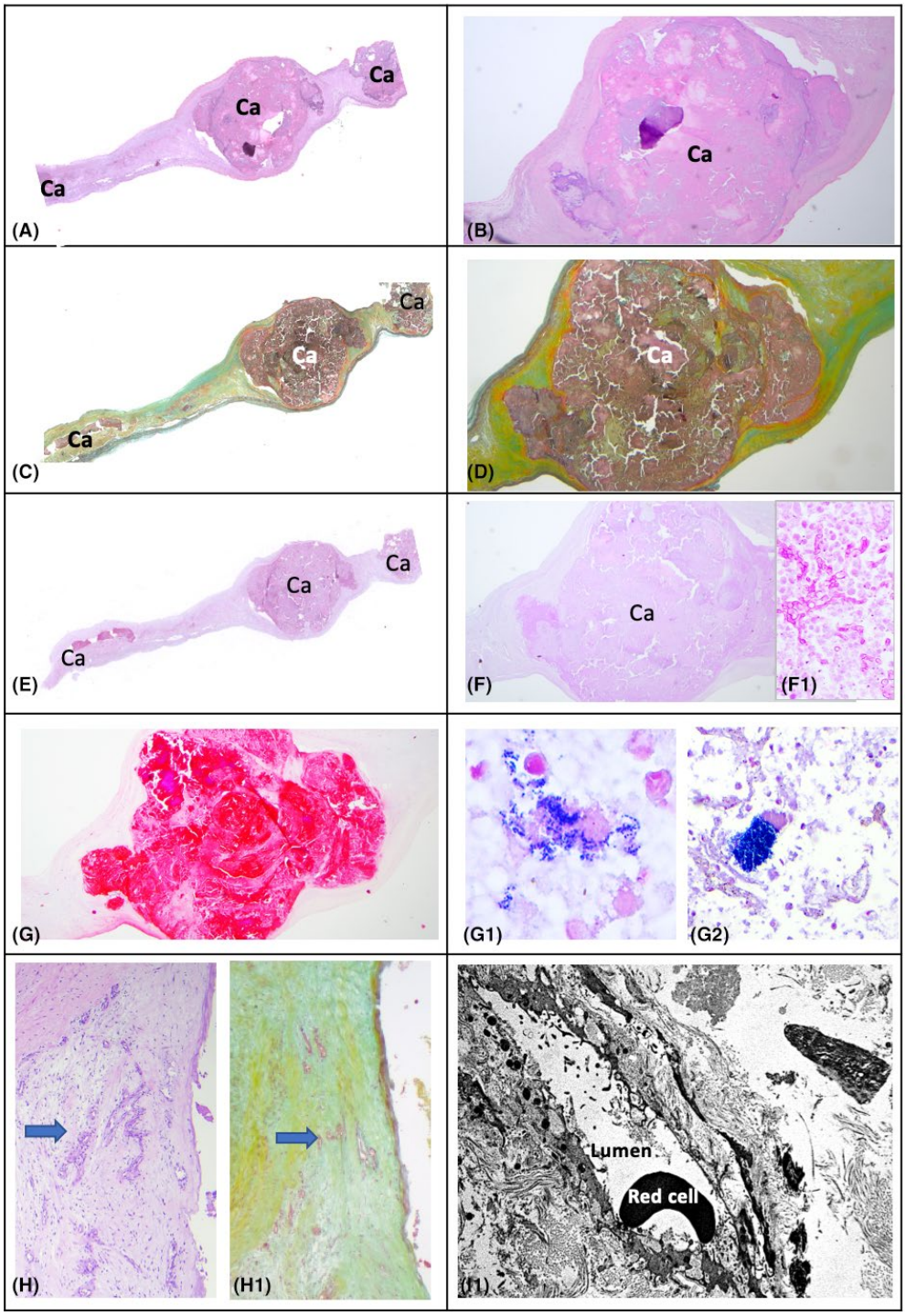


FIGURE 1 The figure shows the spectrum of major morphological changes occurring in CaVD. Panel (A) H&E stain, low magnification view of a valve sample showing a large nodular calcification (nearly acellular) in the context of the thickness of the leaflet; enlarged in panel (B) which highlights the fibro-sclerotic context of the calcified nodule. The panel (E) shows the corresponding PAS stain, enlarged in (F), highlighting the absence of PAS-positive pathogens that, vice versa, stain a positive pneumonia control (F1). The panel (G) shows the same sample staining negative with Gram, vs. Positive Gram controls on bacterial pneumonia (panels G1 & G2). Panel (H) and H1 show H&E- and Movat-stained, non-calcified area of a calcified leaflet showing micro-angioneogenesis in the absence of significant inflammation; a feature confirmed also with ultrastructural study (Panel I).

followed by *Proteobacteria*. The most prevalent families were *Staphylococcaceae*, *Streptococcaceae* and *Enterobacteriaceae*.

ASVs assigned to *Staphylococcus* spp. were found in 5 CAVD leaflets, in all with relative abundance greater than 80% (Figure 4). For four of these leaflets,

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

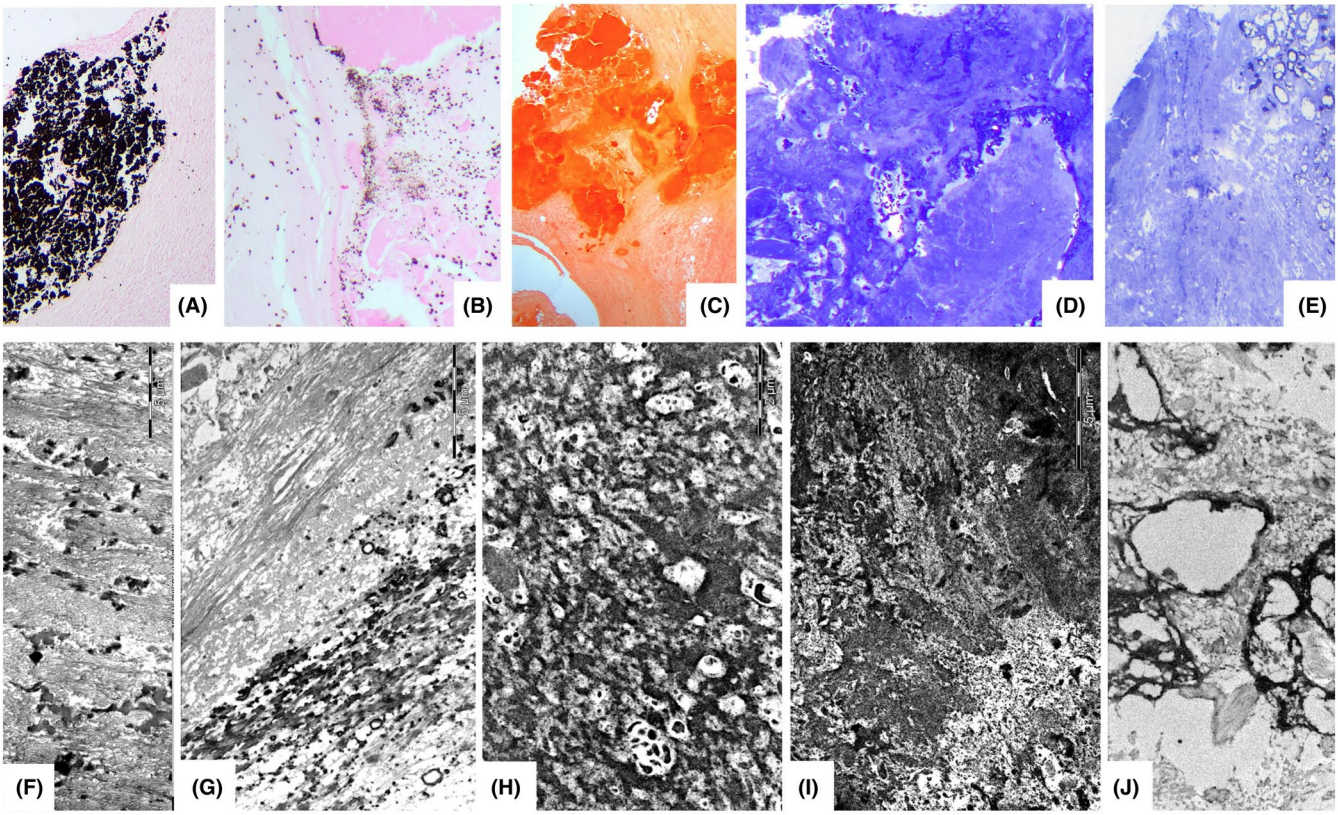


FIGURE 2 Light microscopy. Spectrum of calcifications observed at light and electron microscopy in the leaflets of CaVD. (A, B): Von Kossa Stain; (C) Alizarin red stain; (D, E): Toluidine blue-stained semi-thin section showing the dense calcific masses (dark blue in D) and micro-calcification (E). Ultrastructural studies. Electron micrographs showing the spectrum of calcifications observed in CaVD leaflets, from microcalcifications scattered within the collagen bundles (F), to calcification of entire bundles of collagen (G), dense reticular calcification ground with microdeposits in spared spaces (H), calcification masses (I) and calcified cellular debris (J).

Staphylococcus spp. was also isolated in the enrichment culture.

ASVs assigned to *Streptococcus* spp. were found in one leaflet of a CAVD valve with a relative abundance of 99.8%. This same leaflet tested positive for *S. oralis* in the enrichment culture.

ASVs assigned to the Enterobacteriaceae family were found in two CAVD leaflets from a single CAVD valve, with relative abundances ranging from 53% to 64%. However, no bacteria belonging to the *Enterobacteriaceae* family were isolated in the enrichment culture.

ASVs assigned to other microorganisms with low pathogenic potential, such as *Elizabethkingia* spp., *Chryseobacterium* spp., *Comamonas* spp., *Serratia* spp., *Pseudomonas* spp., *Delftia* spp. and *Morganella* spp., were found among the ASVs of four leaflets from three valves, with relative abundances between 1% and 15%. In all cases, the leaflets were already considered positive as they had ASV predominance of other bacterial genera (Figure 4).

For the AR valves (controls), three leaflets of a single valve showed *Streptococcus* spp. as the most prevalent genus in ASVs. Concordance between the genus grown in culture and the prevalent genus in the ASVs was observed in this case as well.

3.4 | Calcium amount is significantly increased in CaVD samples and was associates with dyslipidemia

At univariate analysis, the ICP-OES analysis revealed a significantly higher calcium amount in CAVD samples compared to AR (coeff. 111, 95% CI 46–176, $p = .001$). Bicuspid aorta leaflets calcium content was increased compared to AR calcium content (coeff. 186, 95% CI 98–273, $p < .0001$). A significant increase was also observed by *t*-test analysis comparing controls versus CAVD tricuspid aorta valves and CAVD bicuspid aorta valves and specific leaflets in the LC and NC, but not in the RC samples (Figure 5). Moreover, at the univariate analysis, dyslipidemia was associated with increased calcium content (coeff. 74, 95% CI 12.6–137, $p = .02$). At multivariable analysis, factors associated with calcium contents were the CAVD group (coeff. 77, 95% CI 4.2–150, $p = .04$) and bicuspid aorta group (coeff. 167, 95% CI 78–257, $p < .0001$) compared to AR.

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

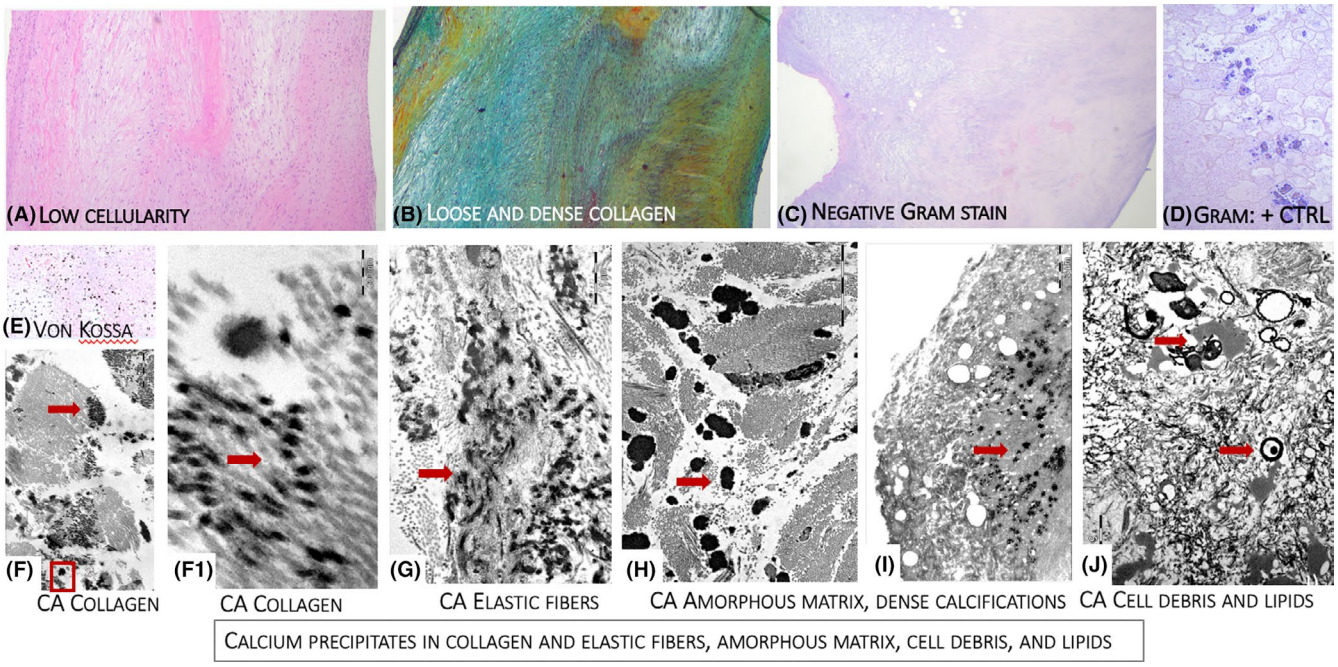


FIGURE 3 The figure shows the common fibrous remodelling in valves with AR, including loose and dense fibrous tissue within the thickness of the leaflets (A: H&E, B: Movat). Negative Gram-stained leaflets (C), paired with its corresponding positive control (blue bacterial colonies) from a bacterial pneumonia (D). A common finding in these valves is the presence of microcalcifications detectable with both light (Von Kossa Stain – E) and electron microscopy in panel (F): The red-squared area Is enlarged in F1), showing early calcium precipitates on single collagen fibrils (red arrow). The panel (G) shows that elastic fibres other than collagen fibrils are common sites of calcium precipitation (all black fibres) (red arrow) (G). Calcium precipitates are also common in the amorphous extracellular matrix, where they appear as smooth-bordered (H) (red arrow) or irregular-bordered microdeposits (red arrow) (I). Lipids (red arrow) and residual cell remnants (red arrow) where calcium can precipitate (J). Overall, most microcalcifications result from calcium precipitation in all components of the extracellular matrix.

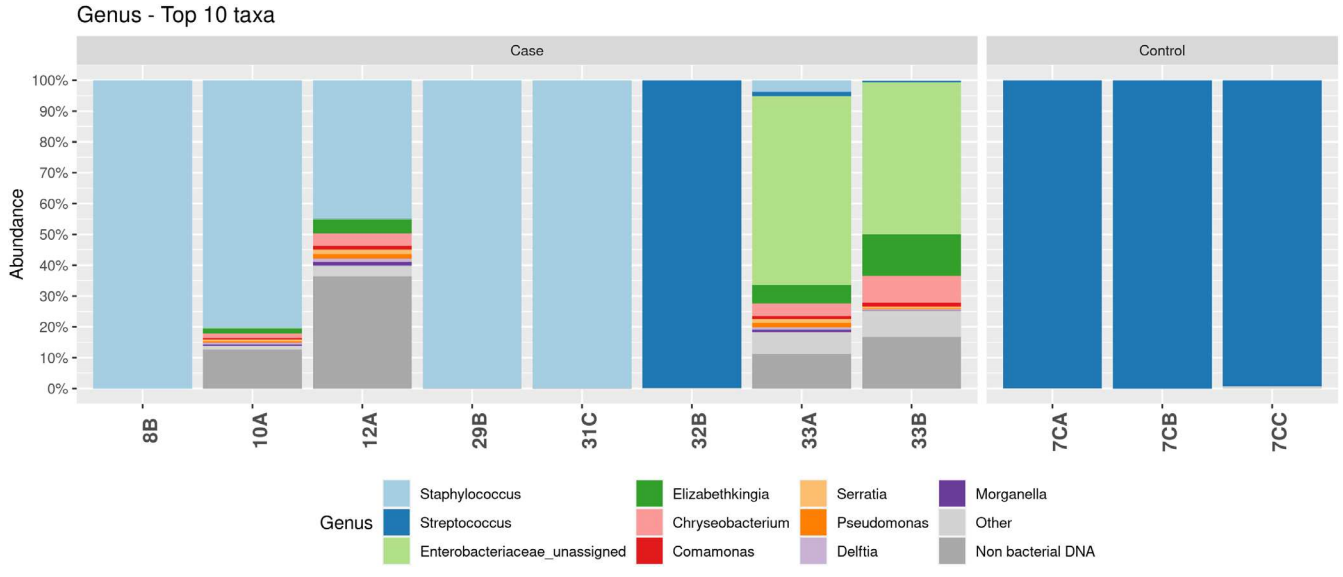


FIGURE 4 Relative abundances of the taxa at the genus level.

In the CAVD subset of patients, valves' calcium content was higher in leaflets with positive microbiologic results, both at univariable (coeff 126, 95% CI 17–236, $p = .024$) and multivariable (coeff 124, 95% CI 4–225, $p = .043$) analysis. The expression of the early osteoblast marker osterix correlates with positive microbiological results.

The expression of early (*osterix* (*OSX*) and *COL1A1*) and late (*ALP, BGLA*P) osteoblast markers were evaluated

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

by qPCR. OSX expression was comparable among leaflets (R: 19.45-fold, L: 40.6-fold, and NC: 19.76-fold) and among CAVD cases and controls (Figure 6). Interestingly, the expression of Osx was significantly higher in leaflets with positive microbiological results both at univariable (coeff. 116, 95% CI 50–182, $p = .001$ 95% CI) and multivariable analysis (coeff. 123, 95% CI 58.5–189.2, $p < .0001$). After controlling for leaflet calcification severity, the association between positive microbiologic results and higher OSX expression remained significant at multivariable analysis (coefficient 133; 95% CI 66–201 $p < .001$).

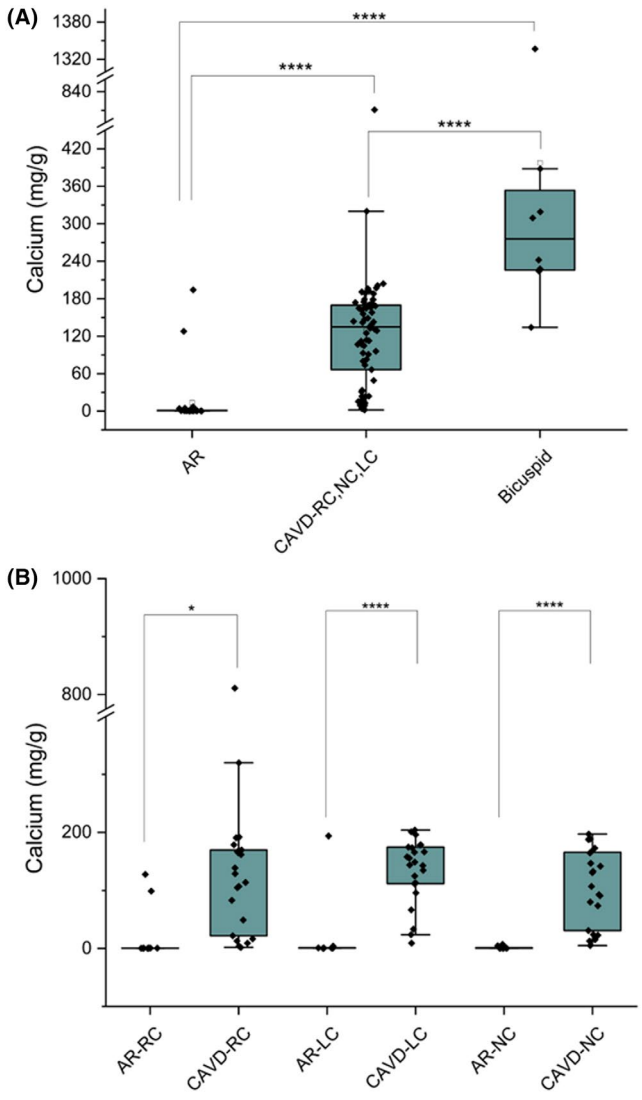


FIGURE 5 Calcium quantification in control and CAVD human aortic valve leaflets. Calcium levels were measured by ICP-OES in pooled AR and CAVD valve leaflets (A) and in individual leaflets (right coronary [RC], left coronary [LC] and non-coronary [NC]) from control and CAVD samples (B). A significant increase in calcium amount is evident in CAVD samples. *$p = .052$, ****$p < .0001$.

No significant differences in COL1A1 expression were observed between CAVD and AR patients, either overall or when analysing individual valve leaflets (RC, LC and NC) separately (Figure 4B). ALP expression was significantly elevated in CAVD leaflets relative to AR, both overall and within each individual valve leaflet (RC, LC and NC) (Figure 6). However, no association was observed with the presence of positive microbiological findings.

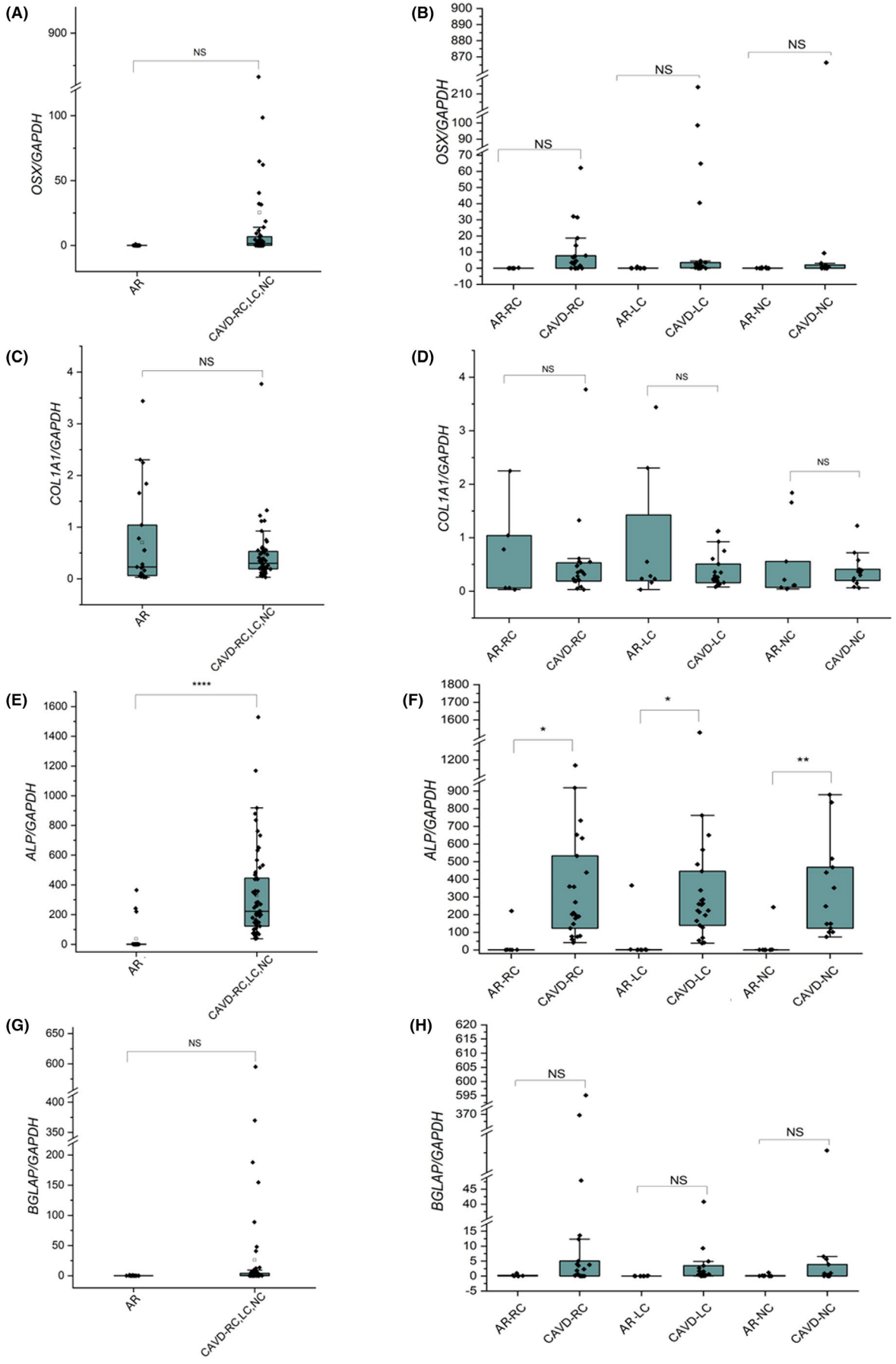
COL1A1 was increased in younger patients, both a univariable (coeff. −.02, 95% CI −.04–.004, $p = .02$) and multivariable analysis (coeff −.02, 95% CI −.04–.0002 $p = .05$). *BGLAP* expression was comparable among the groups (Figure 4D), it was increased in patients with dyslipidemia at multivariable analysis (coeff. 58, 95% CI 16–101, $p = .007$). ALP expression was increased in CAVD patients compared with AR both at univariable (coeff. 348, 95% CI 169–527, $p < .0001$) and multivariable analysis (coeff. 348, 95% CI 169–527, $p < .0001$) also it was increased in bicuspid aorta patients compared with AR at univariable (coeff. 276.7, 95% CI 91.8–461.6, $p = .003$) and multivariable (coeff. 298.3, 95% CI 105–491, $p = .002$).

3.5 | Cell culture VICs staining

Valve interstitial cells (VICs) isolated from four CAVD and four AR samples were cultured under standard growth conditions, without the addition of osteogenic differentiation media. Remarkably, VICs derived from CAVD patients exhibited spontaneous mineral deposition, as evidenced by positive Alizarin Red staining (Figure 7). In contrast, no mineralization was observed in VICs from control valves under the same conditions. Notably, the formation of calcific nodules was exclusively detected in CAVD patient-derived cultures. These findings suggest that VICs from CAVD patients possess an intrinsic osteogenic potential, capable of promoting calcium deposition in vitro even in the absence of mineralizing stimuli.

FIGURE 6 Gene expression analysis of early and late osteoblastic markers. The expression of osterix (*OSX*) (A, B), collagen type I (*COL1A1*) (C, D), alkaline phosphatase (*ALP*) (E, F), osteocalcin (*BGLAP*) (G, H) was analysed in pooled AR and CAVD aortic valve leaflets and in individual leaflets (right coronary [RC], left coronary [LC] and non-coronary [NC]) from control and CAVD samples. Significant increase in OSX and ALP is evident in CAVD samples.

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ect.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

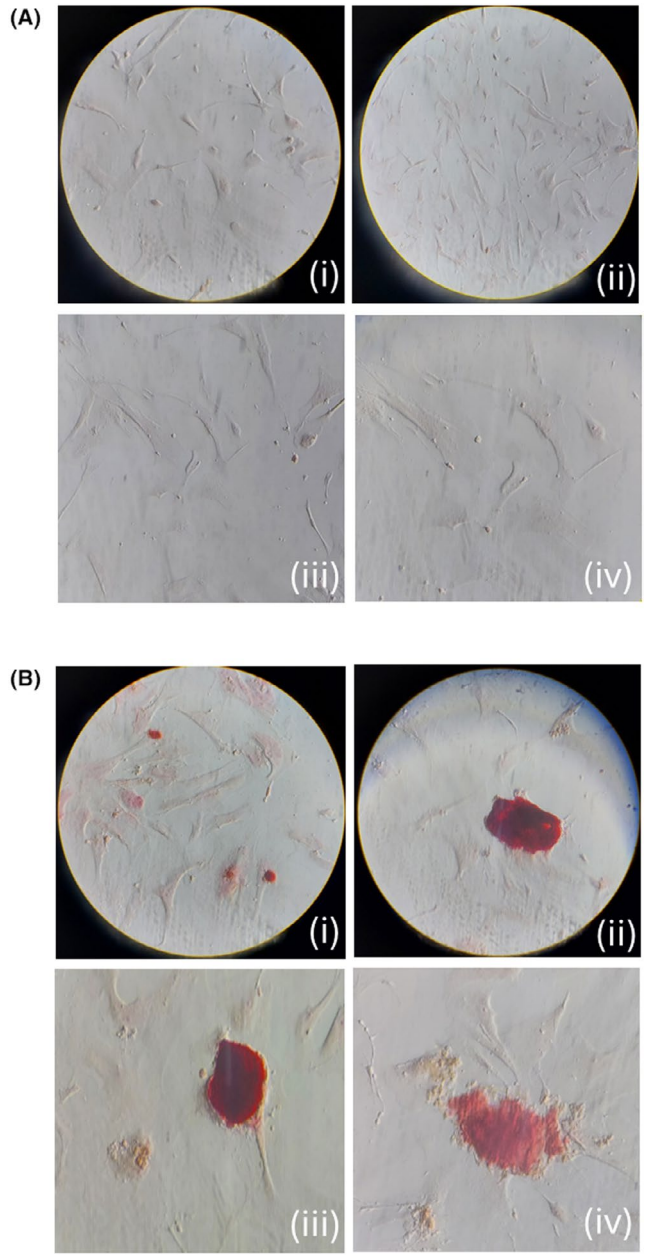


FIGURE 7 Osteogenic transformation in CAVD VICs cultured without differentiation medium. Valve interstitial cells (VICs) from AR (A) and CAVD (B) patients were cultured in standard (non-osteogenic) medium. No evidence of mineralization was observed in control VICs, whereas CAVD-derived VICs exhibited spontaneous osteogenic differentiation, as demonstrated by positive Alizarin Red staining indicating calcium deposition. Ai, Aii 10× magnification; Aiii, Aiv 40× magnification; Bi 10× magnification; Bii, Biii, Biv 40× magnification.

4 | DISCUSSION

The present study investigated the occurrence of bacterial infiltration in CAVD and in AR valves using histopathology, direct and enrichment cultures and molecular-based techniques. Histopathology did not provide evidence of bacterial presence or significant differences in tissue inflammation between enrichment-positive and -negative leaflets. Direct culture of all 108 valve leaflets yielded negative results while enrichment culture detected bacterial presence in 5.95% of CAVD leaflets and 4.16% of non-CAVD leaflets, indicating that some bacteria were viable but at levels too low for direct culturing. In fact, sonication may improve bacterial detection compared to direct culture methods.[14] Other studies of valves from patients with structural valve diseases[7,8] and non-endocarditis cardiac valves[15] did not find bacterial presence through cultural methods, as in both papers sonication step and enrichment culture were not performed. Thus, one of the purposes of the present project was to couple the challenging culture-based methods for bacterial characterization of aortic valves with more sensitive culture-free molecular techniques to improve the efficiency of the diagnosis. Bacteria have been detected only in enrichment culture in roughly 5% of patients, while standard culture yielded negative results. A good correlation was observed between enrichment culture and molecular methods. To attest the sterility of the procedure, negative controls tested negative in the molecular analysis. Bacteria isolated from the enrichment culture included coagulase-negative *Staphylococci* (*S. warneri* and in other two cases *S. pettenkoferi*) and *Streptococcus oralis*. Both *Staphylococcus* species share the features associated with low virulence species and often harbour genes for biofilm production.[16,17] *S. oralis* is one of the most abundant commensal bacteria in the human oral cavity, with relatively low pathogenicity and virulence; it retains the ability to produce biofilm and, under favourable circumstances, can become pathogenic.[18] PCR amplification of the V3-V5 regions of the 16S identified ASVs assigned to bacteria in 22.5% of CAVD patients and 12.5% of AR patients, while the percentage of positive cusps was overlapping among the two groups (roughly 10%). Indeed, ASVs assigned to *Staphylococcus* spp. and *Streptococcus* spp. were found at high relative abundances in 19.5% of CAVD and 12.5% of AR leaflets. Most of these leaflets also tested positive in enrichment cultures, supporting the presence of viable bacteria. On the other hand, ASVs *of Enterobacteriaceae* were detected in a single CAVD case (two leaflets); however, these bacteria did not grow in culture. This discrepancy may indicate the presence of non-viable bacterial DNA possibly associated with a transient valve colonization. However, in this study, the frequencies of bacteria ASVs detected in CAVD samples were lower than those previously reported in the literature.[7,8,15,19] Di Bella and colleagues[15] found bacterial DNA in 44% of the 34 valves which included 20 aortic and 14 mitral valves. In six out of 15 valve samples, DNA of *Escherichia coli* was found, and DNA of *Enterococcus faecalis* and *Gemella sanguinis* was

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

found in three valve samples each. DNA assigned to other bacteria such as *Acinetobacter baumannii*, *Lysinibacillus fusiformis*, *Mycobacterium* spp., *Pseudomonas aeruginosa*, *Staphylococcus* spp., *Streptococcus* spp. and *Stenotrophomonas maltophilia* was found in the remaining samples.

Curini et al.[19] documented the presence of a CAV-associated microbiota in 20 patients enrolled in Germany (eleven patients) and Italy (nine patients), 75% of whom showed severe aortic calcification. Oberbach and colleagues[8] evidenced that 72% of patients with heart valve disease were DNA positive for bacteria commonly associated with IE, such as *C. acnes* (59%), *E. faecalis* (16%), *S. aureus* (15%) and *S. pyogenes* (10%), while bacterial 16S-rDNA could be detected in 94% of the heart valve disease cohort. The detection of bacterial ASVs in the present study is corroborated by culture methods. When comparing our results with those reported in the literature, Di Bella et al. and Oberbach et al.[8,15] did not find in the cultures the same bacterial species that they had detected via 16S sequencing, while Curini et al.[19] did not perform bacterial cultures. However, robust comparisons of our 16S sequencing results with existing literature are challenging due to the lack of harmonization in the methods used. The bacterial genera found through 16S rRNA gene sequencing in our study differed from those reported in previous studies. The identification of distinct bacterial genera in the literature, alongside the detection of *Staphylococcus* and *Streptococcus* species in the valves analysed in our study, raises the question of whether these bacteria play an active role in valve calcification or merely represent incidental colonizers.

Although some genera detected in this study, such as *Comamonas*, *Chryseobacterium* and *Elizabethkingia*, are known to have limited pathogenic potential and have also been reported as reagent contaminants, they were detected in four leaflets (three valves) at low relative abundances. Notably, these taxa were observed in samples characterized by a high relative abundance of a different bacterial taxon. While their biological relevance cannot be entirely excluded, we speculate that bacteria potentially involved in valve calcification would be expected to be present at a predominant biomass. In contrast, low-abundance taxa may reflect transient colonization or technical contamination and should therefore be interpreted with caution. In a rabbit model, Cohen and colleagues[13] evidenced that some low virulence bacteria from the mouth may cause subclinical endocarditis that may be associated with valve calcification. Although our study does not establish a definitive causal link between oral microorganisms and aortic valve calcification in humans, it highlights the potential for such an association. In fact, although bacterial detection rates were low, the early osteogenic marker osterix was upregulated in patients whose valves tested positive for microbial presence. These data suggest a potential contributing role for bacteria in driving cellular differentiation towards an osteoblastic phenotype. The increased expression of the early osteoblastic marker osterix is in line with previous reports that found increased levels of OSX adjacent to calcified nodules.[20] The proximity of Osx-positive VICs to inflammatory immune cells highlights its central role in inflammation-driven calcific remodelling. Pattern recognition receptors (PRRs) require identification of microbe-associated molecular patterns (MAMPs), which include lipopolysaccharide, lipoteichoic acid, lipoprotein and peptidoglycan, in order to trigger host immune responses.[21] Bone cells are modulated by secretory microbial compounds.[22] Bacterial infections can directly influence bone homeostasis through stimulating osteoclast differentiation and activation as well as reducing osteoblast activation and differentiation.[23] Moreover, bacteria and their MAMPs have been reported to influence osteoimmunological responses.[24] Bacterial components can stimulate osteoclastogenesis and alveolar bone resorption by activating pro-inflammatory cytokines, including IL-1β, IL-6 and TNF-α.[25] As a result, immune cells release pro-inflammatory mediators that activate VICs.[26] Immunohistochemical investigations of human diseased valves demonstrate high co-expression of Osx in VICs and inflammatory cells, especially around calcified regions.[20] The close relationship of Osx positive VICs to inflammatory immune cells emphasizes its essential involvement in inflammation-driven calcific remodelling.[27] These findings seem in agreement with the increase of osterix expression in our cohort, in particular in those patients with concomitant bacterial DNA detection.

VICs represent a key factor for progression of CAVD; they can be activated by different stimuli. Previous studies reported that calcified VICs spontaneously develop calcified nodules and exhibit positive staining with Alizarin Red in standard culture, whereas control (healthy) VICs necessitate osteogenic media for this process.[3,28] VICs derived from calcified valves that exhibited substantial Alizarin Red staining at standard growing conditions imply an activated osteogenic phenotype; however, control VICs are stained only following osteogenic induction. Thus, calcified VICs preserve a persistent osteogenic state, whereas healthy VICs need induction.[29,30]

The main limitations of this study are its single-center design and the relatively small sample size, which may limit the generalizability and reproducibility of the findings. In particular, all patients included were Caucasian of Italian origin, which may reflect a specific bacterial cluster unique to this population. Moreover, AR patients included as controls showed leaflet degeneration, a condition that may be associated with altered blood flow and

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

potentially enhanced bacterial adhesion, as described for the case with *Streptococcus oralis*.

The strengths of this study lie mainly in the multidisciplinary approach, which ensured a comprehensive analysis of the potential role of bacteria in CAVD pathogenesis.

5 | CONCLUSIONS

Our study showed that the presence of bacteria in CAVD valves is under the threshold of detection with direct culture and light and electron microscopy study but the combination of enrichment culture with molecular techniques enhances bacterial detection, even at a low rate. The early osteogenic marker osterix was found to be upregulated in patients whose valves tested positive for microbial presence, suggesting a potential role for bacteria in driving cellular differentiation towards an osteoblastic phenotype in CAVD.

AUTHOR CONTRIBUTIONS

A.F., P.G., C.M.S.R., E.S. were involved in conceptualization. A.A.Q.A., I.M., F.A., S.P. were involved in methodology. A.K., M.C.P.C., T.P., M.F., were in-volved in data collection and investigation. G.P., A.D.S. and S.R. were involved in data analysis. B.F., were involved in writing original draft. All authors were involved in writing review and editing. G.Z., F.R. and L.F. were involved in visualization. A.A., V.V. and A.S. were in-volved in supervision. R.B. and E.S. were involved in project administration. E.S. and A.D.S. were involved in funding acquisition.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Antonella Forlino https://orcid.org/0000-0002-6385-1182
Paola Giordani https://orcid.org/0009-0000-5507-3803
Cristina Merla https://orcid.org/0000-0003-2986-6644
Silvia Roda https://orcid.org/0000-0002-0710-174X
Roberta Besio https://orcid.org/0000-0003-1430-2934
Angela Kuka https://orcid.org/0009-0007-5134-5526
Irene Mileto https://orcid.org/0009-0007-5272-3138
Greta Petazzoni https://orcid.org/0000-0003-3537-8006
Marta Corbella https://orcid.org/0000-0002-0969-6484
Baldanti Fausto https://orcid.org/0000-0002-6899-3462
Annalisa De Silvestri https://orcid.org/0000-0003-3128-8441
Totaro Pasquale https://orcid.org/0000-0002-5403-6274
Filippo Amoroso https://orcid.org/0009-0008-5552-8116
Maraschi Federica https://orcid.org/0000-0002-7895-5291
Eloisa Arbustini https://orcid.org/0000-0003-2948-7994
Viviana Vilardo https://orcid.org/0009-0009-0493-0938
Alexandra Smirnova https://orcid.org/0000-0002-8911-8827
Raffaele Bruno https://orcid.org/0000-0002-0235-9207
Seminari Elena https://orcid.org/0000-0001-5246-928X

REFERENCES

1. Kraler S, Blaser MC, Aikawa E, Camici GG, Lüscher TF. Calcific aortic valve disease: from molecular and cellular mechanisms to medical therapy. *Eur Heart J*. 2022;43(7):683-697. doi:10.1093/eurheartj/ehab757
2. Summerhill VI, Moschetta D, Orekhov AN, Poggio P, Myasoedova VA. Sex-specific features of calcific aortic valve disease. *Int J Mol Sci*. 2020;21(16):5620. doi:10.3390/ijms21165620
3. Goody PR, Hosen MR, Christmann D, et al. Aortic valve stenosis: from basic mechanisms to novel therapeutic targets. *Arterioscler Thromb Vasc Biol*. 2020;40(4):885-900. doi:10.1161/ATVBAHA.119.313067
4. Otto CM, Prendergast B. Aortic-valve stenosis--from patients at risk to severe valve obstruction. *N Engl J Med*. 2014;371(8):744-756. doi:10.1056/NEJMra1313875
5. Freeman RV, Otto CM. Spectrum of calcific aortic valve disease: pathogenesis, disease progression, and treatment strategies. *Circulation*. 2005;111(24):3316-3326. doi:10.1161/CIRCULATIONAHA.104.486738
6. Rajamannan NM, Subramaniam M, Rickard D, et al. Human aortic valve calcification is associated with an osteoblast phenotype. *Circulation*. 2003;107(17):2181-2184. doi:10.1161/01.CIR.0000070591.21548.69
7. Oberbach A, Friedrich M, Lehmann S, et al. Bacterial infiltration in structural heart valve disease. *J Thorac Cardiovasc Surg*. 2020;159(1):116-124.e4. doi:10.1016/j.jtcvs.2019.02.019
8. Oberbach A, Schlichting N, Friedrich M, et al. Quantification of multiple bacteria in calcified structural Valvular heart disease. *Semin Thorac Cardiovasc Surg*. 2020;32(2):255-263. doi:10.1053/j.semtcvs.2019.10.003
9. Kwon Y, Park C, Lee J, et al. Regulation of bone cell differentiation and activation by microbe-associated molecular patterns. *Int J Mol Sci*. 2021;22(11):5805. doi:10.3390/ijms22115805
10. Otto CM, Nishimura RA, Bonow RO. Et. Al 2020 ACC/AHA guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. *J Thorac Cardiovasc Surg*. 2021;162(2):e183-e353. doi:10.1016/j.jtcvs.2021.04.002
11. Fowler VG, Durack DT, Selton-Suty C, et al. The 2023 Duke-International Society for Cardiovascular Infectious Diseases Criteria for infective endocarditis: updating the modified Duke criteria. *Clin Infect Dis*. 2023;77(4):518-526. doi:10.1093/cid/ciad271

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

12. Thubrikar MJ, Aouad J, Nolan SP. Patterns of calcific deposits in operatively excised stenotic or purely regurgitant aortic valves and their relation to mechanical stress. *Am J Cardiol*. 1986;58(3):304-308. doi:10.1016/0002-9149(86)90067-6
13. Cohen DJ, Malave D, Ghidoni JJ, et al. Role of oral bacterial flora in calcific aortic stenosis: an animal model. *Ann Thorac Surg*. 2004;77(2):537-543. doi:10.1016/S0003-4975(03)01454-1
14. Gomes A, van Oosten M, Bijker KLB, et al. Sonication of heart valves detects more bacteria in infective endocarditis. *Sci Rep*. 2018;8(1):12967. doi:10.1038/s41598-018-31029-w
15. Di Bella S, Campisciano G, Luzzati R, et al. Bacterial colonization of explanted non-endocarditis cardiac valves: evidence and characterization of the valvular microbiome. *Interact Cardiovasc Thorac Surg*. 2021;32(3):457-459. doi:10.1093/icvts/ivaa285
16. Magnan C, Ahmad-Mansour N, Pouget C, et al. Phenotypic and genotypic virulence characterisation of staphylococcus pettenkoferi strains isolated from human bloodstream and diabetic foot infections. *Int J Mol Sci*. 2022;23(24):15476. doi:10.3390/ijms232415476
17. Alawad MJ, Ali GA, Goravey W. Underrecognized pathogen; Staphylococcus warneri-associated native mitral valve endocarditis in an immunocompetent host: a case report and literature review. *Clin Case Reports*. 2022;10(4):e05591. doi:10.1002/ccr3.5591
18. Ren J, Sun P, Wang M, Zhou W, Liu Z. Insights into the role of *Streptococcus oralis* as an opportunistic pathogen in infectious diseases. *Front Cell Infect Microbiol*. 2024;14:1480961. doi:10.3389/fcimb.2024.1480961
19. Curini L, Alushi B, Christopher MR, et al. The first taxonomic and functional characterization of human CAVD-associated microbiota. *Microb Cell*. 2023;10(2):36-48. doi:10.15698/mic2023.02.791
20. Alexopoulos A, Bravou V, Peroukides Set al. Bone regulatory factors NFATc1 and Osterix in human calcific aortic valves. *Int J Cardiol*. 2010;4:142-149. doi:10.1016/j.ijcard.2008.10.014
21. Chu H, Mazmanian SK. Innate immune recognition of the microbiota promotes host-microbial symbiosis. *Nat Immunol*. 2013;14(7):668-675. doi:10.1038/ni.2635
22. Iwami K, Moriyama T. Effects of short chain fatty acid, sodium butyrate, on osteoblastic cells and osteoclastic cells. *Int J Biochem*. 1993;11:1631-1635.
23. Charles JF, Nakamura MC. Bone and the innate immune system. *Curr Osteoporos Rep*. 2014;12:1-8. doi:10.1007/s11914-014-0195-2
24. D'Amelio P, Sassi F. Gut microbiota, immune system, and bone. *Calcif Tissue Int*. 2018;102:415-425. doi:10.1007/s00223-017-0331-y
25. Sato T, Watanabe K, Kumada H, Toyama T, Tani-Ishii N, Hamada N. Peptidoglycan of Actinomyces naeslundii induces inflammatory cytokine production and stimulates osteoclastogenesis in alveolar bone resorption. *Arch Oral Biol*. 2012;57(11):1522-1528. doi:10.1016/j.archoralbio.2012.07.012
26. Mahler GJ, Butcher JT. Inflammatory regulation of valvular remodeling: the good(?), the bad, and the ugly. *Int J Inf Secur*. 2011;2011:721419. doi:10.4061/2011/721419
27. Towler DA. Molecular and cellular aspects of calcific aortic valve disease. *Circ Res*. 2013;113(2):198-208.
28. Bogdanova M, Zabirnyk A, Malashicheva A, et al. Interstitial cells in calcified aortic valves have reduced differentiation potential and stem cell-like properties. *Sci Rep*. 2019;9(1):12934. doi:10.1038/s41598-019-49016-0
29. Bogdanova M, Kostina A, Zihlavnikova Enayati K, et al. Inflammation and mechanical stress stimulate osteogenic differentiation of human aortic valve interstitial cells. *Front Physiol*. 2018;9:1635. doi:10.3389/fphys.2018.01635
30. Wang Z, Rao Z, Wang Y, Dong N. Establishment and characterization of a novel immortalized human aortic valve interstitial cell line. *Sci Rep*. 2025;15:10917. doi:10.1038/s41598-025-85909-z

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Forlino A, Giordani P, Merla C, et al. Infectious seeds of valve calcification: Exploring the bacterial hypothesis in the pathogenesis of calcific aortic valve disease. *Eur J Clin Invest*. 2026;56:e70188. doi:10.1111/eci.70188

13652362, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eci.70188 by CochraneItalia, Wiley Online Library on [14/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Cytokine 196 (2025) 157042

Contents lists available at ScienceDirect

Cytokine

journal homepage: www.elsevier.com/locate/cytokine

Differential cytokine and chemokine signatures in bronchoalveolar lavage and plasma predict clinical outcomes of COVID-19 patients hospitalized in intensive care unit

Antonio Piralla [a,*,1], Sabrina Gioria [b,1], Giuditta Guerrini [b], Greta Petazzoni [a], Guglielmo Ferrari [a], Alessandro Ferrari [a], Federica Giardina [c], Federica Bergami [a], Anita Orlando [d], Silvia Mongodi [d], Federico Capra Marzani [d], Mara De Amici [e], Giorgia Testa [f], Luigi Calzolai [b], Fausto Baldanti [c,a]

[a] Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[b] European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy
[c] Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy
[d] Anaesthesia and Intensive Care, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[e] Laboratory of Immuno-Allergology of Clinical Chemistry and Pediatric Clinic, Fondazione IRCCS Policlinico San Matteo and University of Pavia, Pavia, Italy
[f] Pediatric Clinic, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

ARTICLE INFO

Keywords:
SARS-CoV-2
Cytokines
Severe infections
ICU
Lower respiratory tract inflammation
Inflammatory markers

ABSTRACT

This study explores the immune response dynamics in critically ill SARS-CoV-2 patients compared to non-SARS-CoV-2 controls both hospitalized in intensive care unit (ICU), focusing on localized and systemic inflammatory profiles to identify patterns linked to clinical outcomes. A cohort of ICU patients (37 SARS-CoV-2 positive, 17 controls) underwent bronchoalveolar lavage (BAL) and plasma analysis to assess inflammatory markers using a multiplex assay. SARS-CoV-2 patients showed distinct clinical and immunological features, including prolonged ICU stays and elevated markers of systemic inflammation. Cluster analysis revealed subgroup-specific immune signatures, with COVID-19 survivors demonstrating coordinated cytokine/chemokine interactions suggestive of adaptive immune regulation, while non-survivors exhibited dysregulated profiles indicative of maladaptive inflammation. The study underscores the spatial segregation of immune responses, with lung-specific inflammation in BAL contrasting with systemic profiles, highlighting the importance of tissue-targeted monitoring. These findings suggest that immune profiling could inform stratification for tailored immunomodulatory therapies, advancing precision approaches in critical care for severe SARS-CoV-2 infections.

1. Introduction

The emergence and global spread of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) have led to a pandemic with a profound impact on public health [1]. Early studies identified a critical association between SARS-CoV-2 infection and localized pulmonary cytokine storms, characterized by hyperinflammation and immune dysregulation [2–4]. The term "cytokine storm," first coined in 1993 to describe the effects of graft-versus-host disease (GVHD) [5], has since then been applied to severe viral infections, including influenza and, notably, SARS-CoV-2 [6,7]. In severe coronavirus disease 2019 (COVID-19), this phenomenon manifests as elevated systemic and pulmonary levels of pro-inflammatory cytokines and chemokines such as IL-6, IL-1β, TNF-α, and IFN-γ–induced protein 10 (IP-10), which correlate with disease severity and mortality [8–11].

While systemic hyperinflammation has been extensively studied [12–15], the localized pulmonary immune response—assessed via bronchoalveolar lavage (BAL) fluid—remains less characterized. BAL specimens provide a direct significant inside into lung-specific immune dynamics, which may diverge significantly from systemic profiles. For instance, studies have shown that local pulmonary inflammation in COVID-19 patients is often more pronounced than systemic responses

* Corresponding author.
E-mail address: antonio.piralla@unipv.it (A. Piralla).
[1] [1]These authors contributed equally to this work.

https://doi.org/10.1016/j.cyto.2025.157042
Received 14 August 2025; Received in revised form 12 September 2025; Accepted 28 September 2025
Available online 1 October 2025
1043-4666/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

[16–19]. However, few studies have simultaneously profiled cytokine levels in both BAL and plasma samples from critically ill patients, leaving gaps in understanding how localized versus systemic inflammation relates to disease progression [3,20]. Thus, our understanding of how localized versus systemic inflammation is related to disease severity and prognosis is still limited. Moreover, comparing cytokine signatures between intensive care unit (ICU) patients COVID-19 and not, is crucial to discern virus-specific immune alterations from general features of critical illness. While overlaps in inflammatory markers exist among ICU populations, the magnitude and combination of cytokines appear distinct in SARS-CoV-2 infection, influencing clinical heterogeneity [21]. For istance, IP-10 and monocyte chemotactic protein-3 (MCP-3) have been identified as robust predictors of severe SARS-CoV-2 disease and mortality, underscoring their potential as biomarkers [18]. Importantly, identifying mediators associated with mortality or recovery could refine risk stratification and therapeutic strategies.

In this study, we aimed to characterize the inflammatory landscape of critically ill SARS-CoV-2 patients by quantifying a broad panel of cytokines and chemokines in both BAL and plasma specimens. By comparing these profiles to those of non-COVID-19 patients, we sought to clarify the unique immune features of severe SARS-CoV-2 infection and examined associations between cytokine patterns, clinical outcomes, and survival. Through integrative analyses—including correlation networks and clustering—we identified key pathways driving disease severity, such as the sustained elevation of IP-10 and GM-CSF, which were linked to poor outcomes in prior work [3,18]. Our findings highlight the interplay between innate and adaptive immune responses and provide insights into potential therapeutic targets for mitigating hyperinflammation in critical SARS-CoV-2 infections.

2. Methods

2.1. Patients and samples

Diagnostic residual samples (bronchoalveolar lavage and plasma samples) were collected from 37 COVID-19-positive and 17 COVID-19-negative (used as control) patients admitted to the ICU at the Fondazione IRCCS Policlinico San Matteo of Pavia (Pavia, Italy) from March 2020 and January 2021. Blood and plasma samples were collected on the same day in most cases, but always within 48 h of each other, and at an early stage after admission, for routine diagnostic purposes. The presence of SARS-CoV-2 RNA in BAL samples was assessed using specific real-time RT-PCR as previously described [22,23]. Residual samples after diagnostic procedures were used for cytokine analysis and were decontaminated by UVC treatment as described in Magri et al. [24]. Information about the patients' age and medical history, were collected from their medical files. This included details about the patients' symptoms and the time between the start of their symptoms and the start of the lung test. Information on treatments at the time of the BAL procedure, oxygen requirements, transfers to the ICU and intra-hospital mortality were also collected. The World Health Organization's Clinical Progression Scale (WHO-CPS) was used to measure disease severity at the time of the BAL procedure. The study complies with the Declaration of Helsinki and the samples and data were collected under the approval of the Institutional Review Board of Policlinico San Matteo (protocol number P_20200029440).

2.2. Cytokine analysis by multiplex ELISA

Diagnostic residual BAL and plasma samples were tested for the presence of cytokines and chemokines by commercial ELISA-based microarrays that simultaneously measure multiple proteins in a single sample aliquot. Multiplex Bio-Plex Pro™ Human Cytokine 27-plex Assay (cat. M500KCAF0Y) was used for assessing the production of IL-1β, IL-1ra, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8 (CXCL-8), IL-9, IL-10, IL-12 (p70), IL-13, IL-15, IL-17 A, Eotaxin (CCL11), FGF basic, G-CSF, GM-CSF, IFN-γ, IP-10 (CXCL10), MCP-1 (CCL2), MIP-1α (CCL3), PDGF-BB, MIP-1β (CCL4), RANTES (CCL5), TNF-α and VEGF. Samples were run according to the manufacturer's instructions. Cytokines were analysed with the Bio-Plex200 System using the Bio-Plex ManagerTM software, and data were analysed by the Bio-Plex Data ProTM software, using five-parametric curve fitting. For each cytokine, assay ranges and LOD were provided by the manufacturer. All reagents and instruments, including Washing Station and Shaker Incubator, were from BIO-RAD Laboratories, Inc. (Hercules, CA, USA). Two repeated measurements were made for each sample. Values reported in the figures are the averages of such repeated measurements.

2.3. Statistical analysis

Statistical analyses were done by using GraphPad Prism 9 software (GraphPad Software Inc., San Diego, Calif). Data were tested for normal distribution by the Shapiro-Wilk test. Data that fit the assumption of normal distribution were compared by using parametric analysis as Student unpaired or paired *t*-tests were used for comparing data from 2 groups. Data that did not fit the assumption of normal distribution were compared by using nonparametric analysis. A nonparametric correlation matrix was calculated by comparing BAL fluid and plasma analytes as well as biochemical variables in a pairwise fashion by using the corr.test function from the pshych CRAN package: the corrplot package was subsequently used to graphically display the correlation matrix. Spearman correlation coefficients were indicated by a heat scale according to which blue indicates a positive linear correlation and red indicates a negative linear correlation. *P* values less than 0.05 were considered significant.

3. Results

3.1. Demographic and clinical data

Baseline characteristics of the study population, management, and outcomes are reported in Table 1. A total of 54 patients were included in the study, of whom 37 were male (68.5 %) and 17 were female (31.5 %). Of these patients, 37 were infected with SARS-CoV-2, while 17 were hospitalized in the ICU with severe neurological, gastrointestinal, oncological, cardiovascular or circulatory conditions, but not respiratory syndrome (see Supplementary Table 1 for a complete list of the clinical syndromes of the control group). Patients positive for SARS-CoV-2 had a more prolonged ICU stay (median 28 days, IQR19.5–49.5 days; $p = 0.002$), Invasive mechanical ventilation (IMV) duration (median 25 days IQR 28.5–62.5 days; <0.001) and hospital stay (median 36 days IQR 28.5–62.5 days; $p = 0.01$) compared to SARS-CoV-2 negative patients. (Table 1). In addition, increased values of C-reactive protein (CRP) (16.2 vs. 2.7; $p = 0.003$), procalcitonin (0.68 vs. 0.14; p = 0.01) and lactate dehydrogenase (LDH) (495 vs. 238 IU/L; $p < 0.001$) were observed in SARS-CoV-2-positive patients compared to the control group (Table 1). Among patients with SARS-CoV-2 infection ($n = 37$), 22 (59.4 %) survived and 15 (40.6 %) died (Table 2). A higher number of ECMO procedures were performed on deceased patients (0 % vs. 26.7 %; $p = 0.01$), and higher levels of LDH were observed in deceased patients (470.5 vs. 542.0 IU/L; $p < 0.001$) compared to SARS-CoV-2 positive patients who survived the disease (Table 2).

3.2. BAL vs. plasma levels of cytokines in SARS-CoV-2 patients and controls

A series of 27 cytokines and chemokine were quantified in 54 BAL of 37 COVID-19-positive patients and 17 controls. Elevated concentrations of IL-7, IP-10 and MCP-1 were observed in patients with SARS-CoV-2 infections compared to the control group, whereas a higher level of IL-8 was observed in control group (Fig. 1A). Elevated concentrations of IP-10 were also observed in plasma of COVID-19 positive patients

Table 1
Clinical characteristics of SARS-CoV-2 positive and negative patients.

Categories		Total ($n = 54$)	SARS-2-CoV + ($n = 37$)	SARS-2-CoV - ($n = 17$)	P value
Median age (range)		62 (1–84)	62 (37–79)	60 (33–84)	0.93
Gender	Female	17 (31.5)	8 (21.6)	9 (52.9)	**0.0298**
	Male	37 (68.5)	29 (78.4)	8 (47.1)	
Comorbidities	No	7 (13.0)	6 (16.2)	1 (5.9)	0.41
	Yes	47 (87.0)	31 (83.8)	16 (94.1)	
Number of comorbidities	1	20 (42.6)	12 (38.7)	8 (50.0)	
	2–3	23 (48.9)	17 (54.8)	6 (37.5)	0.49
	>3	4 (8.5)	2 (6.5)	2 (12.5)	
SOFA score		8 (4–12)	8 (4–12)	9 (6–11)	0.32
ICU stays (days)		23 (1−230)	28 (6–230)	9 (1−120)	**0.002**
IMV duration (days)		21.5 (0−230)	25 (3−230)	9 (0−120)	**<0.001**
Hospital stays (days)		33 (1–230)	36 (6–230)	17 (1–120)	**0.01**
ECMO	No	48 (88.9)	33 (89.2)	15 (88.2)	0.15
	yes	6 (11.1)	4 (10.8)	2 (11.8)	
ICU mortality	No	36 (66.6)	22 (59.5)	14 (82.4)	0.09
	Yes	18 (33.3)	15 (40.5)	3 (17.6)	
Laboratory finding	PaO_2/FiO_2 ratio	118.5 (50–428)	105 (54–218)	270 (50–428)	**0.0014**
	White cell count $x10^3$	10.2 (2.0–24.5)	10.1 (3.3–22.6)	11.9 (2.0–24.5)	0.27
	Lymphocyte count $x10^3$	0.7 (0.1–1.8)	0.65 (0.2–1.3)	0.70 (0.1–1.8)	0.77
	Platelet count	197 (51–503)	208 (71–503)	171.5 (51–467)	0.33
	D-dimer	1333 (14–35,000)	1377 (14–35,000)	1289 (344–4343)	0.99
	CRP	15.7 (0.12–41.1)	16.2 (1.9–40.8)	2.7 (0.1–41.1)	**0.003**
	Procalcitonin	0.53 (0.02–112)	0.68 (0.14–112)	0.14 (0.02–27.6)	**0.01**
	LDH	441 (144–2952)	495 (234–2952)	238 (144–1041)	**<0.001**

SOFA, sequential organ failure assessment; ICU, intensive care unit; IMV, invasive mechanical ventilation, ECMO, extra corporeal membrane oxygenation; CRP, c-reactive protein; LDH, *lactate dehydrogenase.*

Table 2
Characteristics comparison of COVID-19-positive patients survived and not-survived.

Categories		SARS-CoV-2 alive ($n = 22$)	SARS-CoV-2 deceased ($n = 15$)	P value
Median age (range)		61.5 (37–79)	60.0 (50–78)	0.14
Gender	Female	4 (18.2)	4(26.7)	0.37
	Male	18 (81.8)	11 (73.3)	
Comorbidities	No	4 (18.2)	2 (13.3)	0.69
	Yes	18 (81.8)	13 (86.7)	
Number of comorbidities	1	8.0 (44.4)	4 (50.0)	
	2–3	10.0 (55.6)	7 (37.5)	0.10
	>3	0 (0.0)	3 (12.5)	
SOFA score		7.0 (4–12)	8 (6–11)	0.18
ICU stays (days)		24.5 (10–68)	9.0 (1–120)	0.39
IMV stays (days)		21.5 (3–76)	29.0 (5–230)	0.19
ECMO	No	22 (100)	11 (73.3)	**0.01**
	yes	0 (0)	4 (26.7)	
Hospital stay (days)		36 (6–230)	30 (6–230)	0.07
Laboratory finding	PaO_2/FiO_2 ratio	103.5 (54–218)	105.0 (54–188)	0.86
	White cell count $x10^3$	13.3 (3.3–22.6)	10.1 (4.2–16.1)	0.64
	Lymphocyte count $x10^3$	0.70 (0.3–1.3)	0.59 (0.2–1.2)	0.77
	Platelet count	278 (74–503)	187 (71–287)	0.06
	D-dimer	1065 (14–19,990)	2139 (210–35,000)	0.24
	CRP	19.5 (3.4–40.7)	12.8 (1.9–40.3)	0.38
	Procalcitonin	0.56 (0.14–112)	0.79 (0.18–5.7)	0.38
	LDH	471 (234–1402)	542 (403–2952)	**<0.001**

SOFA, sequential organ failure assessment; ICU, intensive care unit; IMV, invasive mechanical ventilation, ECMO, extra corporeal membrane oxygenation; CRP, c-reactive protein; LDH, *lactate dehydrogenase.*

compared to controls (Fig. 1B). In addition, IL-13 and PDGF-bb were slightly higher in patients with COVID-19 with *p*-values close to 0.05. A different pattern of cytokine concentrations was observed comparing BAL and plasma results in COVID-19 patients (Fig. 1C). Elevated concentrations of IL-1β, IL-5, IL-6, IL-8, IL-15, G-CSF, MCP-1, MIP-1α and VEGF were observed in BAL compared to plasma. Conversely, elevated concentrations of IL-1Ra, IL-4, IL-7, IL-9, IL-13, eotaxin, IFN-γ, PDGF-BB and RANTES were observed in the plasma of patients with SARS-CoV-2 infection (Fig. 1C). A similar pattern was observed in the plasma of control patients: seven of the nine increased concentrations observed in patients with SARS-CoV-2 infection (IL-4, IL-7, IL-9, eotaxin, IFN-γ, PDGF-BB and RANTES) were also present in the control group, as shown in Fig. 1D.

3.3. Correlation analysis of inflammatory markers

The relationship between cytokine/chemokine concentrations was investigated and presented as a heatmap showing only significant strong positive (σ value >0.90) or negative (σ value <−0.90) correlations (red box in Fig. 2). When analysing the BAL of patients positive for SARS-CoV-2, only MIP-1α and G-CSF, as well as FGF and IL-17, had significant σ values (0.96 and 0.91, respectively, Fig. 2A). In plasma sample from the same patients, only IL-1Ra and INF-γ had a strong positive correlation (σ value =0.93, Fig. 2B).

Regarding control patient's strong correlations were observed between a series of several cytokines mainly due to the different clinical pictures included in this category. Here are reported only inflammatory markers with at least 3 other markers in correlation. In detail, IL-2 had significate σ values with IL-4, IL-5, and IL-17, while IL-4 was correlated with IL-5, IL-17, FGF, MCP-1 and finally PDGF-bb was correlated with IL-1β, IL-17 and FGF. In plasma samples of control group, IL-4 was correlated with IL-17 and eotaxin and eotaxin was correlated with IL-17.

3.4. Inflammatory markers comparison between survivors and non-survivors from COVID-19

SARS-CoV-2-positive patients were categorised according to their

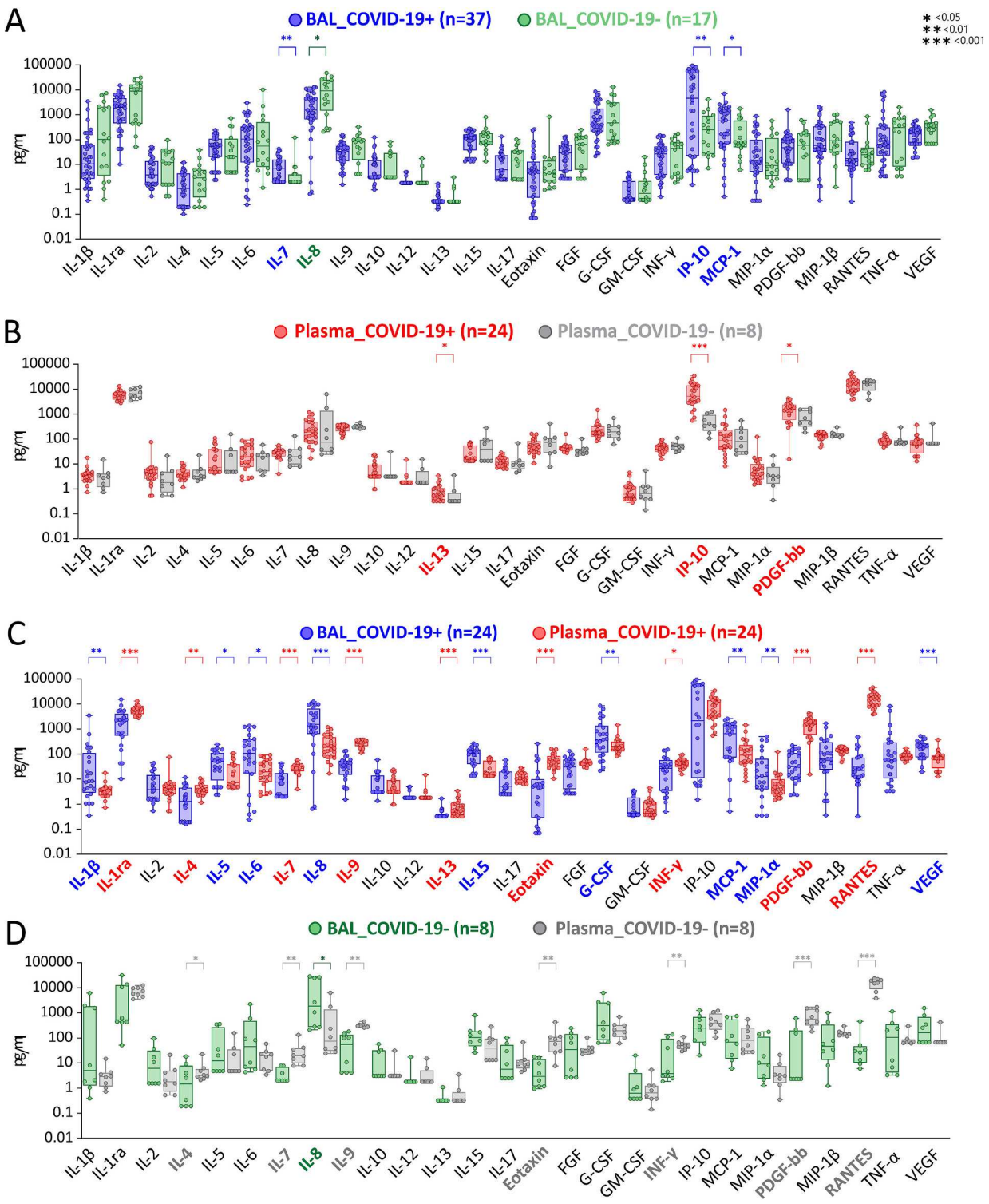


Fig. 1. Cytokine concentrations are measured in pg/mL and plotted on a logarithmic scale. (A) comparison of cytokine concentrations between BAL samples of SARS-CoV-2 positive ICU patients (BAL_COVID-19+, blue bars) and control ICU patients (BAL_covid-19-, green bars). (B) Comparison of cytokine concentrations between plasma samples of SARS-CoV-2 positive ICU patients (Plasma_COVID-19+, red bars) and control ICU patients (Plasma_covid-19-, grey bars). (C) Comparison of cytokine concentrations between BAL samples of SARS-CoV-2 positive ICU patients (BAL_COVID-19+, blue bars) and plasma samples of control ICU patients (Plasma_covid-19-, red bars). (D) Comparison of cytokine concentrations between BAL samples of control ICU patients (BAL_COVID-19-, green bars) and plasma samples of control ICU patients (Plasma_covid-19-, grey bars). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

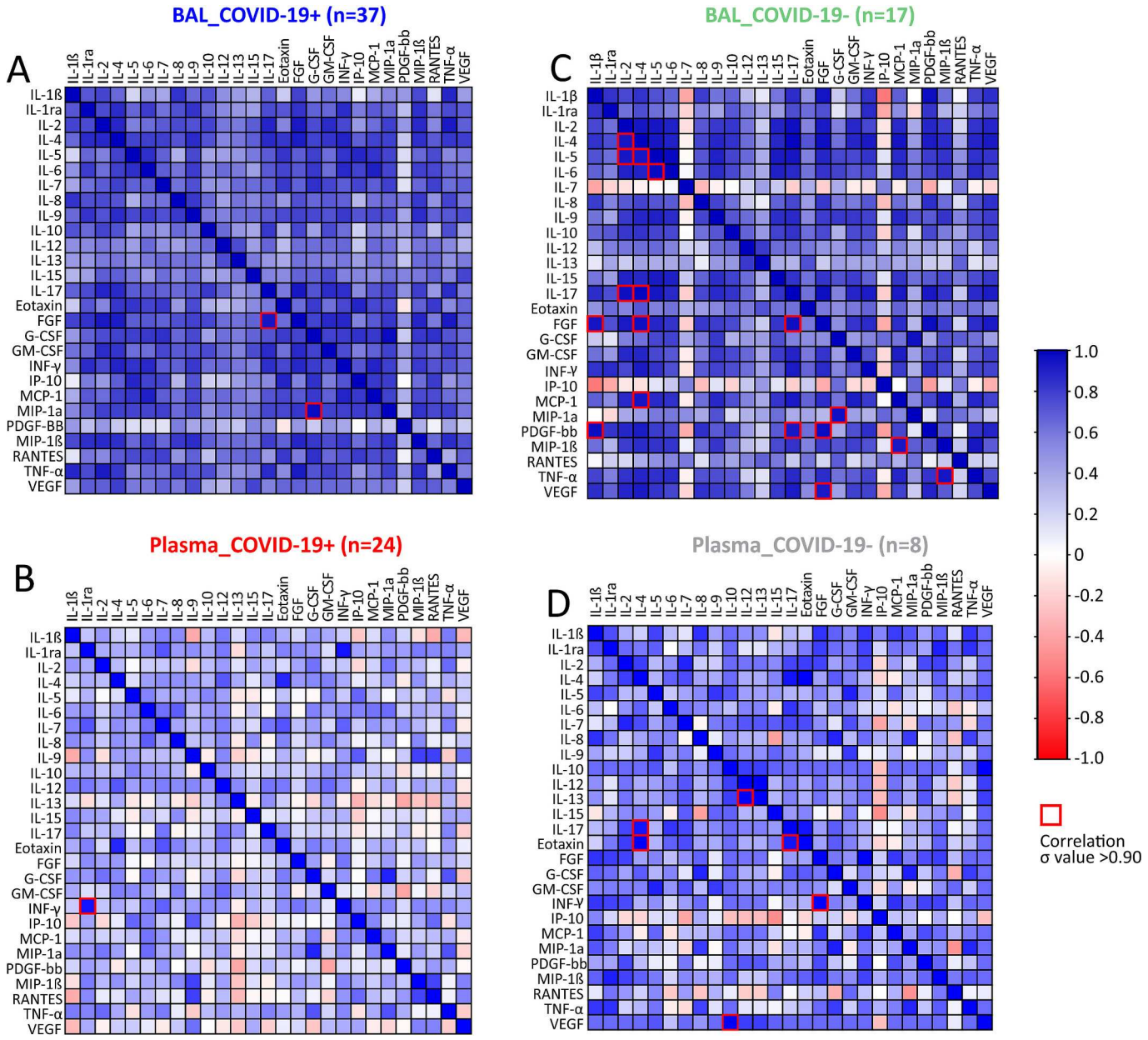


Fig. 2. Heat map of correlations between the various plasma inflammatory markers. As shown on key on the right, blue indicates the strongest positive correlations, while red indicates the strongest negative correlations. (A) Heat map correlation of inflammatory markers in BAL samples of SARS-CoV-2 positive ICU patients (BAL_COVID-19+). (B) Heat map correlation of inflammatory markers in plasma samples of SARS-CoV-2 positive ICU patients (Plasma_COVID-19+). (C) Heat map correlation of inflammatory markers in BAL samples of control ICU patients (BAL_COVID-19-). (D) Heat map correlation of inflammatory markers in plasma samples of control ICU patients (Plasma_COVID-19-). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

outcome, resulting in two groups: one comprising 22 survivors, and the other comprising 15 patients who died. Cytokine/chemokine concentrations were compared, with the results presented in Fig. 3. A pattern of elevated inflammatory markers was observed in BAL samples from surviving patients. Specifically, IL-2, IL-4, IL-6, IL-13, IL-17, FGF, G-CSF, IFN-γ, MCP-1, MIP-1α and TNF-α were elevated in survived patients as compared to deceased patients (Fig. 3A). By contrast, no differences were observed when comparing the levels of cytokines/chemokines in plasma of survivors and non-survivors (Fig. 3B). Interestingly, the heatmap of correlations revealed two distinct cytokine patterns (Fig. 3C-3E). Fig. 3D and F showed a network graph summarising the results of the correlation analysis, clearly demonstrating that the pattern of inflammatory markers differed between these two categories.

Among non-survivor patients a strict correlation (σ value >0.90) was observed between IP-10, Eotaxin and IL-4, and three couple of cytokine/chemokine (IL-6 and IL1ra, MIP-1a and GCSF, IL-5 and INF-γ). Among survived patients, a strict correlation (σ value >0.90) was observed between IL-2, FGF and IL-17 supported by couple between TNF-α and IL-2 and IL-4. Notably, MIP-1α and G-CSF were positively correlated in both categories, suggesting an early innate immune response. None of cytokine/chemokine were negatively correlated with each other (σ value > − 0.90).

3.5. *Exploring cluster analysis of cytokine/chemokine and clinical characteristics*

To analyze the implications of elevated cytokine levels and their relationship with clinical data, we investigated the association of

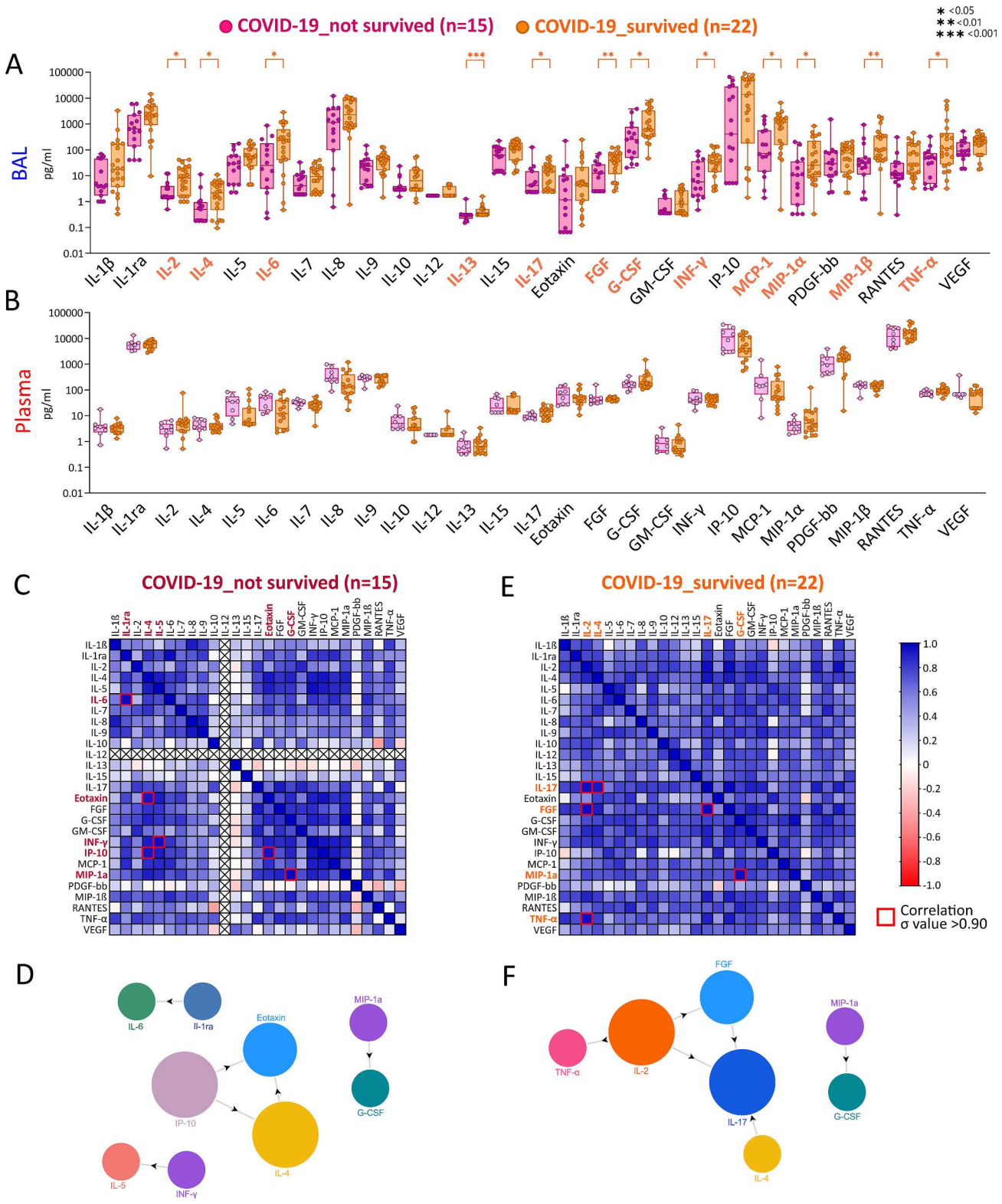


Fig. 3. Comparison of cytokine/chemokine in BAL and plasma of survival vs non-survival patients. (A) Comparison of cytokine/chemokine in BAL samples between SARS-CoV-2 positive ICU patients who did not survive the disease (COVID-19 + _not survived) and SARS-CoV-2 positive ICU patients who survived the disease (COVID-19 + _survived). (B) Comparison of cytokine/chemokine in plasma samples between SARS-CoV-2 positive ICU patients who did not survive the disease (COVID-19 + _not survived) and SARS-CoV-2 positive ICU patients who survived the disease (COVID-19 + _survived). (C) Heat map correlation of inflammatory markers in BAL samples of SARS-CoV-2 positive ICU patients who did not survive the disease (COVID-19 + _not survived). (D) Network graph summarising the results of the correlation analysis of inflammatory markers in BAL samples of SARS-CoV-2 positive ICU patients who did not survive the disease (COVID-19 + _not survived). (E) Heat map correlation of inflammatory markers in BAL samples of SARS-CoV-2 positive ICU patients who survived the disease (COVID-19 + _survived). (F) Network graph summarising the results of the correlation analysis of inflammatory markers in BAL samples of SARS-CoV-2 positive ICU patients who survived the disease (COVID-19 + _survived).

cytokine/chemokine concentrations in BAL to clinical characteristics at ICU admission by hierarchical clustering analysis (Fig. 4). Among chemokines with elevated concentrations in patients with COVID-19 infection (IL-7, IP-10 and MCP-1), only IP-10 and MCP-1 were correlated with a specific cluster (#1) composed by 17 patients (46.0 % of total SARS-CoV-2 patients). A second cluster (#2) was observed, which was made up of six SARS-CoV-2-positive patients and five controls. All patients included in this cluster were found to have survived. A third cluster (#3) composed of 12 patients (five SARS-CoV-2 and seven control) was observed with increasing level of IL-8, IL-1ra (in blue in Fig. 4). In addition, as observed in heatmap correlation analysis MIP-1α and GCS-F clustered together in the same branch of clustering tree (highlighted in red in Fig. 4).

In a subset of 32 patients the cluster analysis was performed in both BAL and plasma samples (n = 22 SARS-CoV-2 positive and n = 10 controls). The main result of this analysis was the clustering of cytokines and chemokines based on samples (Fig. 5). Almost all the inflammatory markers in BAL fluid clustered together, and this was also observed in plasma, confirming a significant difference between the local (BAL) and peripheral (plasma) cytokine/chemokine patterns. In detail, three different clusters (#1–#3) were observed based on the categorisation of patients (Fig. 5). In cluster #1, only SARS-CoV-2 positive patients were included. These patients had significantly high IP-10 levels in BAL and IL-1ra, RANTES and IP-10 in plasma. Cluster #2 included four SARS-CoV-2 positive patients and four control. This cluster was characterized by an increased level of IL-1ra and IL-8 in BAL and IL-1ra, RANTES in plasma. SARS-CoV-2 patients of cluster #2 had also an increasing level of IP-10 in plasma. Cluster #3 included five SARS-CoV-2 positive

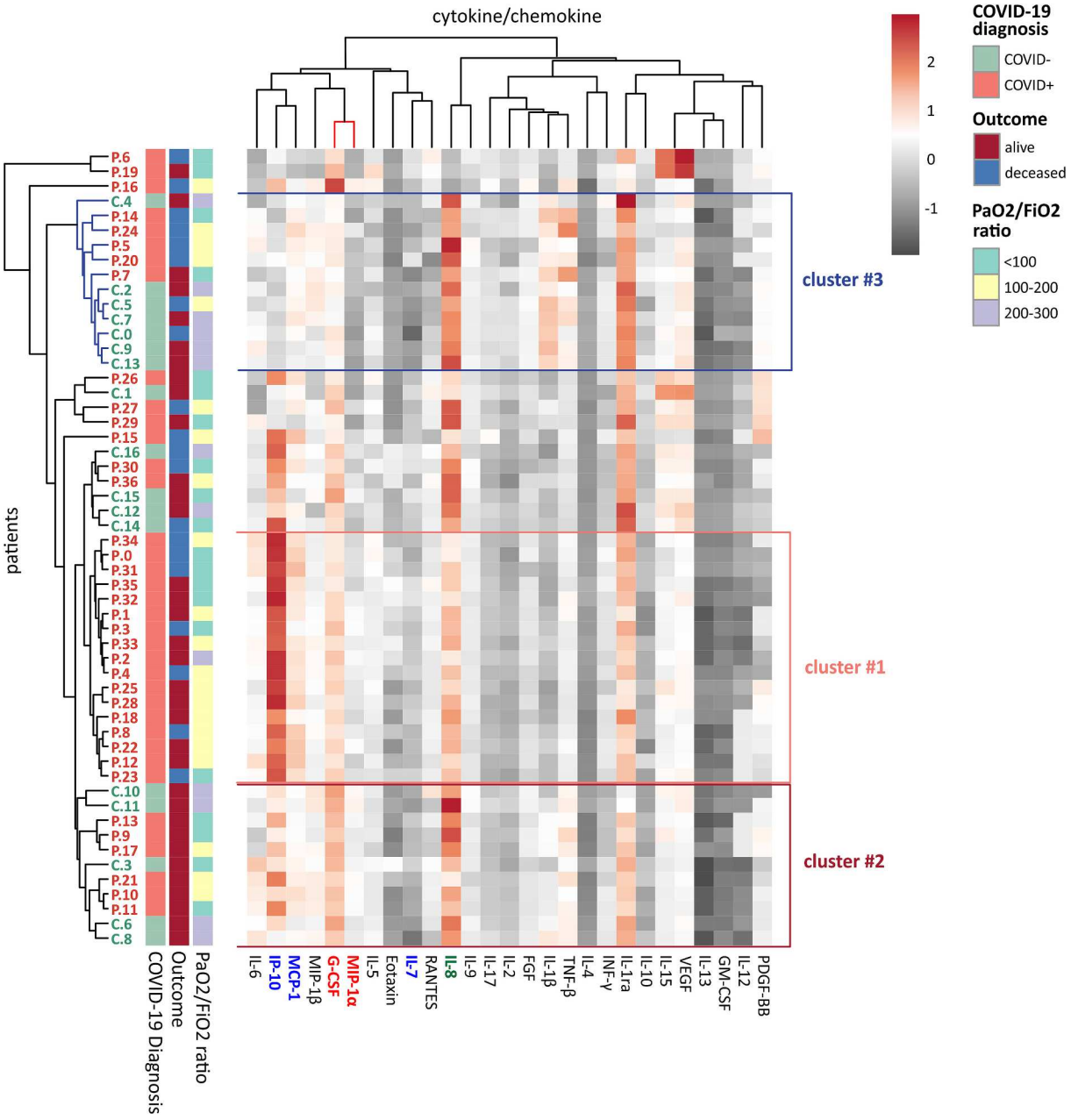


Fig. 4. Euclidean hierarchical clustering analysis of all BAL cytokine measurements in SARS-CoV-2 infected patients (COVID+, red) and controls (COVID-, green). The outcome (alive/deceased) and the PaO_2/FiO_2 ratio are also reported for each patient. The concentration of each cytokine column is normalized by setting the highest value to 1 and the lowest value to 0. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

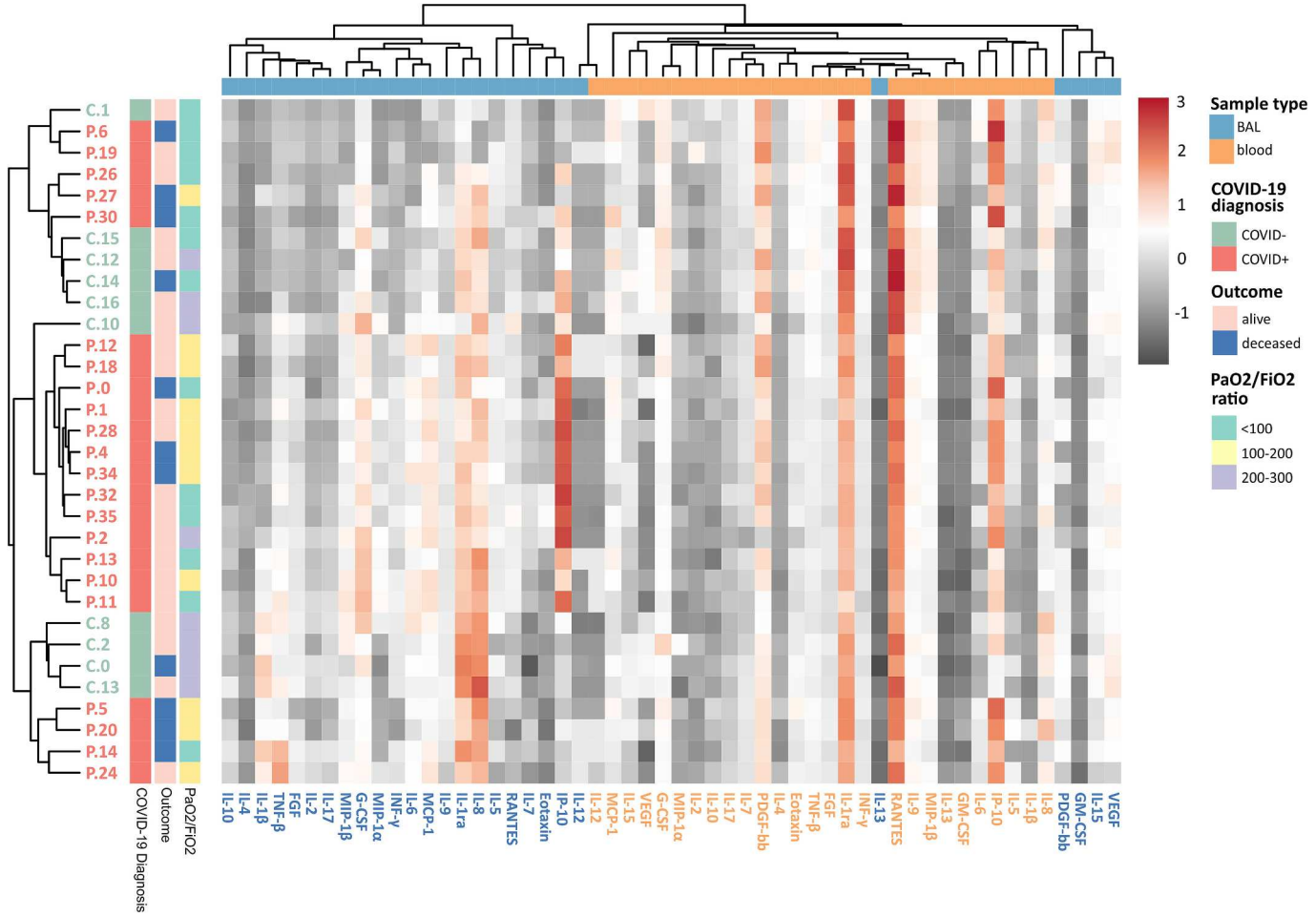


Fig. 5. Euclidean hierarchical clustering analysis of BAL (Blue) and plasma (Orange) cytokine measurements in a subset of 32 patients. SARS-CoV-2 infected patients (n = 22 patients COVID+, red) and controls (n = 10 controls COVID-, green). The outcome (alive/deceased) and the PaO_2/FiO_2 ratio are also reported for each patient. The concentration of each cytokine column was normalized by setting the highest value to 1 and the lowest value to 0. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

patients and six controls. This cluster was characterized by increasing level of PDGF-bb, IL-1ra, RANTES in plasma, and IP-10 in plasma only in SARS-CoV-2 positive patients.

4. Discussion

The primary objective of our study was to characterize the inflammatory profile in critically ill patients with and without SARS-CoV-2 infection, focusing on cytokine/chemokine patterns in BAL and plasma samples. Notably, the majority of our cohort were male patients (68.5 %), a demographic trend consistent with prior studies highlighting increased susceptibility and severity of SARS-CoV-2 infection in males [25]. Our findings align with previously observed more prolonged hospitals and ICU stays, as well as extended durations of IMV, in SARS-CoV-2 patients [26]. These clinical observations were corroborated by elevated inflammatory markers, including CRP, procalcitonin and LDH, which are hallmarks of severe SARS-CoV-2 infection [27]. Of particular significance is the elevated LDH levels observed in cases of SARS-CoV-2 infected patients compared to the control group as also previously reported by other [28,29]. This elevation was further pronounced in deceased patients, whose LDH levels were significantly higher than those of survivors, underscoring its potential as a prognostic indicator. Importantly, Lin et al. demonstrated a strong association between LDH levels above 335 IU/L and increased mortality in ICU patients with cardiac arrest reinforcing the clinical relevance of LDH monitoring in critically ill populations [30]. Elevated IP-10 (CXCL10), MCP-1 and IL-7 in BAL fluid from the patients with COVID-19 support their role in immune cell recruitment (primarily neutrophils and monocytes) and lung inflammation. Furthermore, IP-10 was also significantly elevated in plasma sample from these patients, reinforcing its role as a central mediator of immune dysregulation in COVID-19 [31,32]. Conversely, higher IL-8 levels observed in LRT samples from control patients associated with SARS-CoV-2 infection suggest that IL-8 elevation may not be specific to SARS-CoV-2 infection [33]. These findings align with previous studies demonstrating compartmentalization of the cytokine response in the LRT during SARS-CoV-2 infection [17,34,35]. Specifically, our results show that IL-1β, IL-5, IL-6, IL-8, Il-15, G-CSF, MCP-1, MIP-1α, and VEGF were elevated locally in BAL fluid of COVID-19 positive patients but not systemically, consistent with the concept of a compartmentalized “cytokine storm” confined to the lung microenvironment [36]. Conversely, increased plasma levels of IL-1Ra, IL-4, IL-7, IL-9, IL-13, and IFN-γ likely reflect systemic immune activation or a counter-regulatory anti-inflammatory response [8]. Our results are consistent with those of Voirot et al. [37], who analysed paired BAL and blood samples from 37 critically ill ICU patients with confirmed SARS-CoV-2 infection. This study focused explicitly on 28-day outcomes, which was a clinically relevant endpoint for ICU mortality and recovery. [37]. The study also emphasized the complementary value of BAL and blood biomarkers for risk stratification, indicating that lung-specific inflammation could be a more immediate cause of adverse outcomes than systemic markers. Two additional studies investigated the compartmentalization of immune responses in the lungs of critically ill

patients with SARS-CoV-2 infection [38,39]. BAL revealed distinct inflammatory profiles (neutrophilia, ISGs, and cytokines) that correlated with poor outcomes. Blood markers (e.g., lymphopenia and IL-6) reflect systemic inflammation but are less predictive of severity than lung-specific signals [38,39]. Together, these studies underscore the importance of studying both compartments (BAL and blood) in order to improve our understanding of the pathophysiology of severe SARS-CoV-2 infection and develop more accurate prognostic tools.

In addition, the overlap of certain plasma cytokine patterns observed between SARS-CoV-2 patients and controls underscore that some cytokine elevations may reflect generalized critical illness rather than being specific to COVID-19 [40]. This highlights the importance of distinguishing disease-specific immune signatures from those associated with sepsis or multi-organ failure in critically ill populations. The strong associations between MIP-1α and G-CSF, as well as between FGF and IL-17 in BAL fluid, suggest coordinated regulation of neutrophil and fibroblast responses, potentially contributing to both acute inflammation and tissue remodelling in the lung [41]. In plasma, the positive correlation between IL-1Ra and IFN-γ indicates a complex interplay between pro- and anti-inflammatory mediators at the systemic level. Interestingly, control patients exhibited broader cytokine intercorrelations, likely reflecting the heterogeneity of clinical syndromes and underlying immune responses.

One of the most striking findings of this work, is the elevated cytokine concentrations in BAL fluid from survivors compared to non-survivors from COVID-19. Survivals demonstrated significantly higher levels of IL-2, IL-4, IL-6, IL-13, IL-17, FGF, G-CSF, IFN-γ, MCP-1, MIP-1α, and TNF-α, suggesting that a robust local immune response may be protective by promoting viral clearance and tissue repair [42]. This contrasts with plasma cytokine levels, which did not differ significantly between survivors and non-survivors, underscoring that systemic cytokine measurements alone may not accurately reflect the protective or pathogenic immune dynamics in the lung. Previous work have shown that survivors typically mount an effective, balanced immune response, whereas fatal cases are associated with dysregulated or exhausted immunity [16]. The distinct correlation networks observed in survivors and non-survivors further support differential immune programming. In deceased patients, a strong correlation among IP-10, IL-4, and Eotaxin may represent a dysregulated immune response characterized by excessive chemokine-driven cell recruitment (IP-10, Eotaxin) and an ineffective Th2-biased environment (IL-4) leading to persistent inflammation, immune-mediated tissue injury, and poor viral control [36]. Conversely, survivors exhibited positive correlations among IL-2, FGF, and IL-17, potentially reflecting effective immune activation and tissue repair. The consistent positive correlation of MIP-1α and G-CSF in both groups highlights the importance of innate immune activation during early disease progression [43].

Cluster analysis identified subgroups of patients with distinct cytokine signatures profiles linked to clinical outcomes. A cluster (#1) defined by high IP-10 and MCP-1 among SARS-CoV-2 patients reinforces the association between elevated chemokine levels and a specific inflammatory phenotype. The overlap of cytokines clusters between survivors and controls suggests that cytokine profiles may partially overlap with those seen in other critical illnesses, consistent with reports that cytokine elevations are not unique to COVID-19 but rather reflect generalized inflammation induced by severe disease [21].

While this study has some limitations, including a relatively modest cohort size and heterogeneity in the control group, it provides valuable insights into the immune dynamics of critically ill patients with and without SARS-CoV-2 infection. The use of BAL fluid, though limited to a subset of patients due to its invasive nature, allow for direct assessment of the lung microenvironment, which may not be fully captured by systemic measurements. Longitudinal sampling would be beneficial to understand the temporal dynamics of cytokine responses. The identification of distinct cytokine/chemokine patterns in BAL and plasma underscores the potential of these biomarkers for understanding localized immune mechanisms and informing therapeutic strategies. Future studies incorporating larger, more homogeneous cohorts and longitudinal sampling could further refine these findings and elucidate the temporal evolution of cytokine responses in the lung and bloodstream. Despite its current constraints, this work contributes meaningfully to the growing body of evidence on immune profiling in critical illness and highlights the importance of tissue-specific immune monitoring in guiding precision medicine approaches.

In conclusion, this study confirms that SARS-CoV-2 infection triggers elevated inflammatory markers in both lung tissue and systemic compartments, underscoring the dual nature of immune activation in response to the virus. Furthermore, cytokine/chemokine profile was identified in survivors, characterized by a robust yet balanced local immune response that likely contributes to viral clearance and tissue repair. The differential pattern observed between survivors and non-survivors highlight the potential for these biomarkers to stratify patients for targeted immunomodulatory therapies, enabling more personalized approaches to critical care. The compartmentalization of immune responses—wherein local inflammation in the lung contrasts with systemic immune dynamics—reveals the complexity of SARS-CoV-2 immunopathology. These findings emphasize the need for an integrated strategy that considers both tissue-specific and systemic immune markers to guide treatment decisions. By bridging the gap between localized immune mechanisms and systemic immune dysregulation, this work provides a foundation for refining therapeutic interventions and improving outcomes in severe COVID-19.

CRediT authorship contribution statement

Antonio Piralla: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis. **Sabrina Gioria:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Giuditta Guerrini:** Writing – review & editing, Writing – original draft. **Greta Petazzoni:** Software, Formal analysis. **Guglielmo Ferrari:** Validation, Formal analysis. **Alessandro Ferrari:** Investigation. **Federica Giardina:** Investigation. **Federica Bergami:** Investigation. **Anita Orlando:** Resources. **Silvia Mongodi:** Resources. **Federico Capra Marzani:** Writing – review & editing, Resources. **Mara De Amici:** Writing – review & editing. **Giorgia Testa:** Investigation. **Luigi Calzolai:** Writing – review & editing. **Fausto Baldanti:** Supervision, Funding acquisition, Conceptualization.

Funding

This project was supported by the European Union's Horizon 2020 research and innovation programme (ATAC, No. 101003650). The study was partially supported by the European Commission's Joint Research Centre (JRC) within the Directorate Health and Food through the Portfolio "Strategic technologies".

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Acknowledgments

We thank Mrs. Daniela Sartori for manuscript editing. The authors would like to acknowledge Prof. Arnd Hoeveler and Dr. Rita Maccario for their support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cyto.2025.157042.

Data availability

Data will be made available on request.

References

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, et al., A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (8) (2020) 727–733, https://doi.org/10.1056/NEJMoa2001017.

[2] P. Mehta, D.F. McAuley, M. Brown, E. Sanchez, R.S. Tattersall, J.J. Manson, et al., COVID-19: consider cytokine storm syndromes and immunosuppression, Lancet 395 (10229) (2020) 1033–1034, https://doi.org/10.1016/S0140-6736(20)30628-0.

[3] S. Hue, A. Beldi-Ferchiou, I. Bendib, M. Surenaud, S. Fourati, T. Frapard, et al., Uncontrolled innate and impaired adaptive immune responses in patients with COVID-19 acute respiratory distress syndrome, Am. J. Respir. Crit. Care Med. 202 (11) (2020) 1509–1519, https://doi.org/10.1164/rccm.202005-1885OC.

[4] Q. Ye, B. Wang, J. Mao, The pathogenesis and treatment of the `cytokine storm' in COVID-19, J. Inf. Secur. 80 (6) (2020) 607–613, https://doi.org/10.1016/j.jinf.2020.03.037.

[5] J.L. Ferrara, S. Abhyankar, D.G. Gilliland, Cytokine storm of graft-versus-host disease: a critical effector role for interleukin-1, Transplant. Proc. 25 (1 Pt 2) (1993) 1216–1217 (PMID: 8442093).

[6] B.G. Chousterman, F.K. Swirski, G.F. Weber, Cytokine storm and sepsis disease pathogenesis, Semin. Immunopathol. 39 (5) (2017) 517–528, https://doi.org/10.1007/s00281-017-0639-8.

[7] J.R. Tisoncik, M.J. Korth, C.P. Simmons, J. Farrar, T.R. Martin, M.G. Katze, Into the eye of the cytokine storm, Microbiol. Mol. Biol. Rev. 76 (1) (2012) 16–32, https://doi.org/10.1128/MMBR.05015-11.

[8] D.M. Del Valle, S. Kim-Schulze, H.H. Huang, N.D. Beckmann, S. Nirenberg, B. Wang, et al., An inflammatory cytokine signature predicts COVID-19 severity and survival, Nat. Med. 26 (10) (2020) 1636–1643, https://doi.org/10.1038/s41591-020-1051-9.

[9] C. Lucas, P. Wong, J. Klein, T.B.R. Castro, J. Silva, M. Sundaram, et al., Longitudinal analyses reveal immunological misfiring in severe COVID-19, Nature 584 (7821) (2020) 463–469, https://doi.org/10.1038/s41586-020-2588-y.

[10] Y. Yang, C. Shen, J. Li, J. Yuan, J. Wei, F. Huang, et al., Plasma IP-10 and MCP-3 levels are highly associated with disease severity and predict the progression of COVID-19, J. Allergy Clin. Immunol. 146 (1) (2020) 119–127.e4, https://doi.org/10.1016/j.jaci.2020.04.027.

[11] V.P. Chavda, R. Bezbaruah, N. Ahmed, S. Alom, B. Bhattacharjee, L.V. Nalla, et al., Proinflammatory cytokines in chronic respiratory diseases and their management, Cells 14 (6) (2025) 400, https://doi.org/10.3390/cells14060400.

[12] T.S. Bonny, E.U. Patel, X. Zhu, E.M. Bloch, M.K. Grabowski, A.G. Abraham, et al., Cytokine and chemokine levels in coronavirus disease 2019 convalescent plasma, Open Forum Infect. Dis. 8 (2) (2020) ofaa574, https://doi.org/10.1093/ofid/ofaa574.

[13] A.N. Nguyen, T.S. Kouyate, K. Ryff, A.L. Plotkin, S. Doss-Gollin, S. Thomas, et al., Patients hospitalized with COVID-19 demonstrate distinct plasma cytokine and chemokine concentrations in vivo and TLR-mediated cytokine and chemokine production in whole blood in vitro, J. Innate Immun. 17 (1) (2025) 288–301, https://doi.org/10.1159/000545432.

[14] C. Huang, X. Hu, D. Wang, R. Gong, Q. Wang, F. Ren, et al., Multi-cohort study on cytokine and chemokine profiles in the progression of COVID-19, Sci. Rep. 14 (1) (2024) 10324, https://doi.org/10.1038/s41598-024-61133-z.

[15] S.G. Vanderkamp, M. Niazy, A.A. Stegelmeier, K.J. Stinson, N. Ricker, B.W. Bridle, Cytokine, chemokine, and acute-phase protein profiles in plasma as correlative biomarkers of clinical outcomes for patients with COVID-19, Sci. Rep. 15 (1) (2025) 15397, https://doi.org/10.1038/s41598-025-99248-6.

[16] E.J. Giamarellos-Bourboulis, M.G. Netea, N. Rovina, K. Akinosoglou, A. Antoniadou, N. Antonakos, et al., Complex immune dysregulation in COVID-19 patients with severe respiratory failure, Cell Host Microbe 27 (6) (2020) 992–1000.e3, https://doi.org/10.1016/j.chom.2020.04.009.

[17] M. Liao, Y. Liu, J. Yuan, Y. Wen, G. Xu, J. Zhao, et al., Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19, Nat. Med. 26 (6) (2020) 842–844, https://doi.org/10.1038/s41591-020-0901-9.

[18] X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, et al., Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study, Lancet Respir. Med. 8 (5) (2020) 475–481, https://doi.org/10.1016/S2213-2600(20)30079-5.

[19] C. Bime, N.G. Casanova, J. Nikolich-Zugich, K.S. Knox, S.M. Camp, et al., Strategies to DAMPen COVID-19-mediated lung and systemic inflammation and vascular injury, Transl. Res. 232 (2021) 37–48, https://doi.org/10.1016/j.trsl.2020.12.008.

[20] Y. Liu, C. Zhang, F. Huang, Y. Yang, F. Wang, J. Yuan, et al., Elevated plasma levels of selective cytokines in COVID-19 patients reflect viral load and lung injury, Natl. Sci. Rev. 7 (6) (2020) 1003–1011, https://doi.org/10.1093/nsr/nwaa037.

[21] P. Sinha, M.A. Matthay, C.S. Calfee, Is a "cytokine storm" relevant to COVID-19? JAMA Intern. Med. 180 (9) (2020) 1152–1154, https://doi.org/10.1001/jamainternmed.2020.3313.

[22] F. Rovida, D. Cereda, S. Novati, A. Licari, A. Triarico, G.L. Marseglia, et al., Low risk for SARS-CoV2 symptomatic infection and early complications in paediatric patients during the ongoing CoVID19 epidemics in Lombardy, Clin. Microbiol. Infect. 26 (11) (2020) 1569–1571, https://doi.org/10.1016/j.cmi.2020.06.006.

[23] A. Piralla, M. Ricchi, M.G. Cusi, P. Prati, N. Vicari, G. Scarsi, et al., Residual SARS-CoV-2 RNA in nasal swabs of convalescent COVID-19 patients: is prolonged quarantine always justified? Int. J. Infect. Dis. 102 (2021) 299–302, https://doi.org/10.1016/j.ijid.2020.10.072.

[24] D. Magrì, A. Navarro, F. Bergami, E. Percivalle, A. Ferrari, T. Lettieri, et al., Impact of viral decontamination method on cytokine profile of COVID-19 patients, Biomedicines 9 (10) (2021) 1287, https://doi.org/10.3390/biomedicines9101287.

[25] H. Peckham, N.M. de Gruijter, C. Raine, A. Radziszewska, C. Ciurtin, L. R. Wedderburn, et al., Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission, Nat. Commun. 11 (1) (2020) 6317, https://doi.org/10.1038/s41467-020-19741-6.

[26] G. Grasselli, M. Greco, A. Zanella, G. Albano, M. Antonelli, G. Bellani, et al., Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. JAMA Intern Med. 180 (10) (2020) 1345–1355, https://doi.org/10.1001/jamainternmed.2020.3539.

[27] J. Zhou, J. Sun, Z. Cao, W. Wang, K. Huang, F. Zheng, et al., Epidemiological and clinical features of 201 COVID-19 patients in Changsha city, Hunan, China, Medicine (Baltimore) 99 (34) (2020) e21824, https://doi.org/10.1097/MD.0000000000021824.

[28] B. Fialek, M. Pruc, J. Smereka, R. Jas, M. Rahnama-Hezavah, A. Denegri, et al., Diagnostic value of lactate dehydrogenase in COVID-19: a systematic review and meta-analysis, Cardiol. J. 29 (5) (2022) 751–758, https://doi.org/10.5603/CJ.a2022.0056.

[29] F. Ostadi, B.C. Anzali, H.R. Mehryar, Relationship between serum lactate dehydrogenase levels and prognosis in patients infected with omicron and delta variants of COVID-19: a cross-sectional study, Toxicol. Rep. 11 (2023) 368–373, https://doi.org/10.1016/j.toxrep.2023.10.003.

[30] L. Lin, R. Gao, L. Chen, Z. Wu, X. Wei, Y. Xie, Relationship between serum lactate dehydrogenase and mortality after cardiac arrest: a retrospective cohort study, Medicine (Baltimore) 101 (45) (2022) e31499, https://doi.org/10.1097/MD.0000000000031499.

[31] T. Eichhorn, S. Huber, R. Weiss, M. Ebeyer-Masotta, L. Lauková, R. Emprechtinger, et al., Infection with SARS-CoV-2 is associated with elevated levels of IP-10, MCP-1, and IL-13 in Sepsis patients, Diagnostics (Basel). 13 (6) (2023) 1069, https://doi.org/10.3390/diagnostics13061069.

[32] Y. Chen, J. Wang, C. Liu, L. Su, D. Zhang, J. Fan, et al., IP-10 and MCP-1 as biomarkers associated with disease severity of COVID-19, Mol. Med. 26 (1) (2020) 97, https://doi.org/10.1186/s10020-020-00230-x.

[33] M.C. Cesta, M. Zippoli, C. Marsiglia, E.M. Gavioli, F. Mantelli, M. Allegretti, et al., The role of Interleukin-8 in lung inflammation and injury: implications for the management of COVID-19 and hyperinflammatory acute respiratory distress syndrome, Front. Pharmacol. 12 (2022) 808797, https://doi.org/10.3389/fphar.2021.808797.

[34] Y. Jouan, T. Baranek, M. Si-Tahar, C. Paget, A. Guillon, Lung compartmentalization of inflammatory biomarkers in COVID-19-related ARDS, Crit. Care 25 (1) (2021) 120, https://doi.org/10.1186/s13054-021-03513-9.

[35] J.M. Cavaillon, B.G. Chousterman, T. Skirecki, Compartmentalization of the inflammatory response during bacterial sepsis and severe COVID-19, J Intensive Med. 4 (3) (2024) 326–340, https://doi.org/10.1016/j.jointm.2024.01.001.

[36] D. Blanco-Melo, B.E. Nilsson-Payant, W.C. Liu, S. Uhl, D. Hoagland, R. Møller, et al., Imbalanced host response to SARS-CoV-2 drives development of COVID-19, Cell 181 (5) (2020) 1036–1045.e9, https://doi.org/10.1016/j.cell.2020.04.026.

[37] G. Voiriot, K. Dorgham, G. Bachelot, A. Fajac, L. Morand-Joubert, C. Parizot, et al., Identification of bronchoalveolar and blood immune-inflammatory biomarker signature associated with poor 28-day outcome in critically ill COVID-19 patients, Sci. Rep. 12 (1) (2022) 9502, https://doi.org/10.1038/s41598-022-13179-0.

[38] A. Saris, T.D.Y. Reijnders, E.J. Nossent, A.R. Schuurman, J. Verhoeff, S.V. Asten, et al., Distinct cellular immune profiles in the airways and blood of critically ill patients with COVID-19, Thorax 76 (10) (2021) 1010–1019, https://doi.org/10.1136/thoraxjnl-2020-216256.

[39] P.A. Szabo, P. Dogra, J.I. Gray, S.B. Wells, T.J. Connors, S.P. Weisberg, et al., Longitudinal profiling of respiratory and systemic immune responses reveals myeloid cell-driven lung inflammation in severe COVID-19, Immunity 54 (4) (2021) 797–814.e6, https://doi.org/10.1016/j.immuni.2021.03.005.

[40] M. Kox, N.J.B. Waalders, E.J. Kooistra, J. Gerretsen, P. Pickkers, Cytokine levels in critically ill patients with COVID-19 and other conditions, JAMA 324 (15) (2020) 1565–1567, https://doi.org/10.1001/jama.2020.17052.

[41] F. Coperchini, L. Chiovato, L. Croce, F. Magri, M. Rotondi, The cytokine storm in COVID-19: an overview of the involvement of the chemokine/chemokine-receptor system, Cytokine Growth Factor Rev. 53 (2020) 25–32, https://doi.org/10.1016/j.cytogfr.2020.05.003.

[42] D. Mathew, J.R. Giles, A.E. Baxter, D.A. Oldridge, A.R. Greenplate, J.E. Wu, et al., Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications, Science 369 (6508) (2020) eabc8511, https://doi.org/10.1126/science.abc8511.

[43] J. Hadjadj, N. Yatim, L. Barnabei, A. Corneau, J. Boussier, N. Smith, et al., Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients, Science 369 (6504) (2020) 718–724, 10.1126/.

microorganisms

MDPI

Article

Native Bacterial Communities of Two Italian Salso-Bromo-Jodic and Sulphurous Natural Mineral Waters

Angela Kuka [1,†], Irene Mileto [1,†], Marco Saler [2,3,4], Greta Petazzoni [1,2,*], Marta Corbella [1], Fausto Baldanti [1,2], Angela Faga [3] and Giovanni Nicoletti [2,3,4,5,*]

[1] Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
[2] Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia, 27100 Pavia, Italy
[3] Advanced Technologies for Regenerative Medicine and Inductive Surgery Research Center, University of Pavia, 27100 Pavia, Italy
[4] Integrated Unit of Experimental Surgery, Advanced Microsurgery and Regenerative Medicine, University of Pavia, Via Adolfo Ferrata, 9, 27100 Pavia, Italy
[5] Surgery Unit, Azienda Socio-Sanitaria Territoriale di Pavia, 27100 Pavia, Italy
* Correspondence: g.petazzoni@smatteo.pv.it (G.P.); giovanni.nicoletti@unipv.it (G.N.)
† These authors contributed equally to this work.

Abstract: A correlation between resident non-pathogenic bacterial populations in certain natural mineral waters and their beneficial effects has been established by several research groups. This study aims to characterize the bacterial composition of the Rivanazzano salso-bromo-jodic and sulphurous mineral waters (Pavia, Italy). Water samples were collected from natural sources and dispensing systems. DNA was extracted and subjected to 16S rRNA gene sequencing. Microbial composition, as well as alpha and beta diversity, were analyzed using amplicon sequence variants and compared across sampling sites. The predominant phyla in both waters were *Proteobacteria*, *Campylobacterota*, *Bacteroidota*, and *Desulfobacterota*. However, diversity at the family taxonomic level was recorded. In terms of bacterial diversity, waters collected from the dispensing systems within the spa resort were more similar between them than those from natural sources. The therapeutic properties of the Rivanazzano mineral waters are likely to be related to their combined mineral and biological composition.

Keywords: bacterial communities; 16S; mineral water; sulphurous water; salso-bromo-jodic water; non-pathogenic microflora

check for updates

Received: 27 February 2025
Revised: 23 April 2025
Accepted: 26 April 2025
Published: 30 April 2025

Citation: Kuka, A.; Mileto, I.; Saler, M.; Petazzoni, G.; Corbella, M.; Baldanti, F.; Faga, A.; Nicoletti, G. Native Bacterial Communities of Two Italian Salso-Bromo-Jodic and Sulphurous Natural Mineral Waters. *Microorganisms* **2025**, *13*, 1038. https://doi.org/10.3390/microorganisms13051038

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Natural mineral waters, whose therapeutic properties have been long recognised on empirical basis, are currently being investigated with rigorous experimental methods and have demonstrated antioxidant, wound-healing, skin hydration, and skin barrier recovery activities [1–8].

Over the last 15 years, our research group has been focusing on the regenerative properties of natural mineral waters. In detail, we have demonstrated this potential with human in vitro cell and ex vivo tissue culture experimental models both in calcium magnesium bicarbonate-based natural mineral water (Comano, Trento, Italy) [9–14] and in salso-bromo-jodic-based water (Rivanazzano, Pavia, Italy) [15].

Natural mineral waters are defined as bacteriologically pure, the latter definition meaning that they contain neither pathogenic microorganisms nor microorganisms indicating fecal or other contamination [16]. However, resident non-pathogen bacterial populations

have been identified in some natural mineral waters, and a correlation between these bacterial microflora and the waters' beneficial effects has been demonstrated. Indeed, *Aquaphilus dolomiae*, isolated from Avène thermal water (France), is able to regulate keratinocyte inflammatory and lymphocyte immune responses [17–20]. Similarly, *Vitreoscilla filiformis*, found in LaRoche-Posay thermal water (France), has been demonstrated to activate cutaneous regulatory T cells [21,22]. The native non-pathogenic microflora in Comano (Trento, Italy) natural mineral water display immune-modulating, anti-inflammatory, and even antibiotic effects [23].

A previous study of ours on both in vitro and ex vivo experimental models demonstrated some regenerative effects in a bacterial lysate from a new strain close to *Mesorhizobium erdmanii* (*Proteobacteria* phylum) [14], isolated in Comano mineral water.

Both favourable clinical effects and some experimental evidence [15,24] of Rivanazzano mineral waters are likely to be related to the combination of their well-known chemical–physical composition and their microbiological composition, which is still unknown to date. The aim of this experimental study was to characterize the native bacterial composition of Rivanazzano Terme (Pavia, Italy) salso-bromo-jodic and sulphurous mineral waters.

2. Materials and Methods

2.1. Thermal Bath Setting

Rivanazzano Terme (Pavia, Italy) is an Italian village located at 150 metres above sea level. The Rivanazzano Salus per Aquam (spa resort) has, unusually, two distinctly different water springs in the same site: a salso-bromo-jodic natural mineral water and a sulphurous one [25]. The chemical–physical composition of both waters is reported in Table 1.

Table 1. Physical–chemical composition of the Rivanazzano salso-bromo-jodic and sulphurous mineral waters.

Parameter	Salso-Bromo-Jodic	Sulphurous
Temperature at the spring	13.9 °C	13.2 °C
pH	7.7	7.4
Conductivity at 20 °C	8100 µS/cm	1805 µS/cm
Fixed residue at 180 °C	12,400 mg/L	1334 mg/L
Oxidability	48.0 mg/L (O_2)	39.0 mg/L (O_2)
Silica (SiO_2)	50.5 mg/L	19.9 mg/L
Bicarbonates (HCO_3)	814 mg/L	705 mg/L
Chlorides (Cl)	7300 mg/L	192 mg/L
Sulphates (SO_4)	35.7 mg/L	283 mg/L
Sodium (Na)	4490 mg/L	379.1 mg/L
Potassium (K)	24.5 mg/L	4.7 mg/L
Calcium (Ca)	143 mg/L	75 mg/L
Magnesium (Mg)	78.5 mg/L	31.2 mg/L
Dissolved iron (Fe)	3.37 mg/L	<0.01 mg/L
Ammonium ion (NH_4)	22.4 mg/L	3.15 mg/L
Total phosphorus (P)	0.13 mg/L	<0.02 mg/L
Sulphide level (H_2S)	not detectable	11.9 mg/L
Strontium (Sr)	15.8 mg/L	1.964 mg/L
Lithium (Li)	0.417 mg/L	0.066 mg/L
Aluminum (Al)	<0.01 mg/L	<0.01 mg/L
Bromides (Br)	37.1 mg/L	1.0 mg/L
Iodides (I)	13.9 mg/L	0.5 mg/L
Antimony (Sb)	<0.0005 mg/L	<0.0005 mg/L
Arsenic (As)	0.001 mg/L	<0.001 mg/L
Barium (Ba)	3.6 mg/L	<0.1 mg/L
Boron (B)	54.1 mg/L	3.24 mg/L
Cadmium (Cd)	<0.0001 mg/L	<0.0001 mg/L
Chromium (Cr)	0.001 mg/L	<0.001 mg/L
Copper (Cu)	<0.1 mg/L	<0.1 mg/L
Cyanides (CN)	<0.001 mg/L	<0.001 mg/L
Fluorides (F)	1.4 mg/L	0.26 mg/L
Lead (Pb)	<0.001 mg/L	<0.001 mg/L

Table 1. *Cont.*

Parameter	Salso-Bromo-Jodic	Sulphurous
Manganese (Mn)	0.06 mg/L	0.041 mg/L
Mercury (Hg)	<0.0001 mg/L	<0.0001 mg/L
Nickel (Ni)	<0.001 mg/L	0.001 mg/L
Nitrates (NO_3)	<1.0 mg/L	<1.0 mg/L
Nitrites (NO_2)	<0.002 mg/L	0.007 mg/L
Selenium (Se)	0.005 mg/L	<0.001 mg/L
Anionic surfactant agents	<50 µg/L	<50 µg/L
Mineral oils—dissolved or emulsified hydrocarbons	41 µg/L	34 µg/L
Benzene	<0.1 µg/L	<0.1 µg/L
Benzo[a]pyrene	not detectable	<0.003 µg/L
Benzo[b]fluoranthene	not detectable	<0.003 µg/L
Benzo[k]fluoranthene	not detectable	<0.003 µg/L
Benzo[ghi]perylene	not detectable	<0.003 µg/L
Dibenz[a,h]anthracene	not detectable	<0.003 µg/L
Indeno[1,2,3-cd]pyrene	not detectable	<0.003 µg/L
Acenaphthene	not detectable	<0.003 µg/L
Fluorene	not detectable	<0.003 µg/L
Phenanthrene	not detectable	<0.003 µg/L
Anthracene	not detectable	<0.003 µg/L
Fluoranthene	not detectable	<0.003 µg/L
Pyrene	not detectable	<0.003 µg/L
Benzo[a]anthracene	not detectable	<0.003 µg/L
Chrysene	not detectable	<0.003 µg/L
Chloroform	not detectable	<0.1 µg/L
Dibromochloromethane	not detectable	<0.1 µg/L
Bromodichloromethane	not detectable	<0.1 µg/L
Bromoform	not detectable	<0.1 µg/L
Trichloroethylene	not detectable	<0.1 µg/L
Tetrachloroethylene	not detectable	<0.1 µg/L
1,2-Dichloroethane	not detectable	<0.1 µg/L
1,1-Dichloroethylene	not detectable	<0.1 µg/L
1,2-Dichloroethylene	not detectable	<0.1 µg/L
1,1,1-Trichloroethane	not detectable	<0.1 µg/L
Carbon tetrachloride	not detectable	<0.1 µg/L
Total hardness	59.6 °F	31.6 °F

The salso-bromo-jodic water is drawn up at a temperature of 13.9 °C from an approximately 350-metre-deep artesian well, consisting of a perforated 316 stainless steel "jacket" surrounded by draining gravel and a 316 stainless steel suction tube. A 316 stainless steel extraction pump is placed at a depth of roughly 60 m. The water is stored in a series of three underground 10,000 L settling tanks interconnected by 316 stainless steel pipes. The tanks are built in reinforced concrete covered with water-repellent plaster. After passing through the last tank, the water is fed through a long polyethylene pipeline until it reaches the spa resort 2.5 km away. Here, the water is shifted to two different storage tanks built in reinforced concrete covered with water-repellent plaster. One is filled with water at the native extraction temperature, and the other one is filled with water heated to 70 °C. The waters are then delivered to the individual bath tubs through a 316 stainless steel distribution network. In the bath tubs, the heated water is blended with water at the native temperature to provide adequate comfort for patients.

The sulphurous water is extracted at 13.2 °C from an approximately 12-metre-deep artesian well with the same structural features as the salso-bromo-jodic water well. The water is delivered to an underground storage fiberglass tank within the spa resort through a dedicated polyethylene pipeline, which is almost 2 km long. Through a 316 stainless steel pipeline, the sulphurous water is then delivered to inhalation dispensing systems and to a single individual bath tub, where it is extemporaneously warmed according to the patient's requirements (Figure 1).

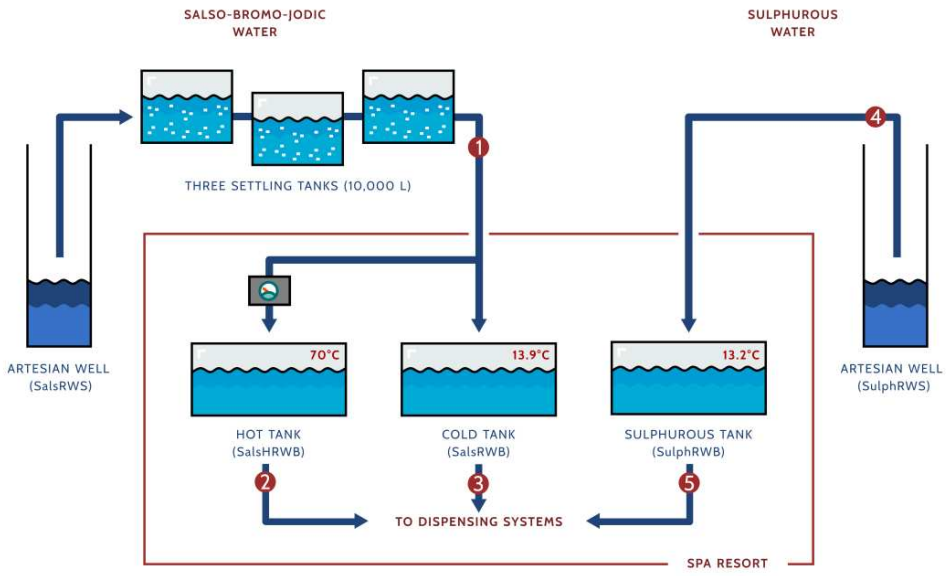


Figure 1. Schematic representation of the Rivanazzano Terme water network. Collection sites are indicated by numbers in red circles: (1) salso-bromo-jodic water from the natural source collected at the exit of the last settling tank (SalsRWS); (2) heated salso-bromo-jodic water collected at the exit of its storage tank within the spa resort (SalsHRWB); (3) salso-bromo-jodic water at the natural source temperature collected at the exit of its storage tank within the spa resort (SalsRWB); (4) sulphurous water collected at the natural source (SulphRWS); and (5) sulphurous water collected at the exit of its storage tank within the spa resort (SulphRWB).

The polyethylene pipelines from the natural sources to the spa resort are regularly cleaned with a reverse-flow rinse with tap water on a 3-monthly basis. The storage tanks in the spa resort are maintained every year with sediment removal and tap water rinsing. The stainless steel pipelines from the storage tanks to the different dispensing systems and to the individual bath tubs in the spa resort are rinsed with reverse-flow tap water for 90 min every night.

2.2. Study Design

2.2.1. Samples Collection

The water samples were collected at three different time points: May, June, and July 2022. Waters were collected from five sample sites to obtain a comprehensive overview of the bacterial communities (Figure 1):

1. The exit of the third settling tank at the site of the salso-bromo-jodic natural source (SalsRWS).
2. The exit of the salso-bromo-jodic storage tank at 70 °C within the spa resort (Hot tank; SalsHRWB).
3. The exit of the salso-bromo-jodic storage tank at the natural source temperature within the spa resort (Cold tank; SalsRWB).
4. The sulphurous natural source (SulphRWS).
5. The exit of the sulphurous storage tank at the natural source temperature within the spa resort (Sulphurous tank; SulphRWB).

For each site and collection period, water samples were gathered by a single individual using an aseptic procedure, as reported in a previous study [10]. In each site, water samples were collected in two 2-litre sterile pots. Then, they underwent centrifugation to remove macroscopic sediment. Filtration was performed using cellulose filters with a porosity of 0.1 µm (Whatman, GE HealthCare Life Sciences, Chicago, IL, USA), and filters were vortexed in an aliquot of water before starting with sample processing. Each sample

was then concentrated in 2 mL tubes to proceed with the genomic DNA extraction. Two technical replicates were carried out for each water sample.

2.2.2. Library Preparation and Sequencing

Bacterial DNA was extracted using NucleoSpin Tissue (Macherey–Nagel®, Düren, Germany) following the manufacturer's instructions, and the rRNA 16S region was sequenced by BMR Genomics s.r.l. (Padova, Italy). The V3 and V4 hypervariable regions were targeted using Pro341F (5′-CCTACGGGNBGCASCAG-3′) and Pro805R (Rev 5′-GACTACNVGGGTATCTAATCC-3′) primers for Bacteria and Archaea [26], modified with Illumina sequencing tails (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNBGCASCAG-3′ and Rev 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACNVGGGTATCTAATCC-3′, respectively) and sequenced with a paired-end set up (2 × 300 bp, Illumina, San Diego, CA, USA).

2.2.3. Bioinformatic Analysis

The raw sequence data were processed in QIIME2 (vr. 2022.8) [27] using the DADA2 algorithm [28] to trim reads (last 5 nt for forward and 30 nt for reverse reads), remove low-quality reads, denoise, estimate run-specific error rates, and infer amplicon sequence variants (ASVs). Within the algorithm, quality control included chimaera detection and removal, sequence error elimination, and singleton exclusion. Taxonomy was assigned on ASVs training a naive Bayes classifier with the SILVA 138 database [29]. Multiple sequence alignments were performed using MAFFT (vr. 7.505) [30] and FastTree 2 (vr. 2.1.11) [31] was used to infer a rooted phylogenetic tree.

2.2.4. Diversity Analyses

The α-diversity metrics (i.e., richness and Shannon diversity index) of bacterial communities were computed at all taxonomic levels on even rarefied data to analyse the within-site diversity using the phyloseq package (version 1.40.0) [32]. The β-diversity was computed at all taxonomic levels. First, it was computed to assess the diversity within sites for the three different water sample collection times using UniFrac weighted distances and then to compare sampling sites using Bray–Curtis distances on even rarefied data (three from salso-bromo-jodic water and two from sulphurous water). Principal coordinate analysis (PCoA) was used for visualisation, and a permutational multivariate analysis of variance (PERMANOVA) was performed with 999 permutations on distances using the vegan R package (version 2.6-4) [33]. Once the homogeneity between time points within the same site had been assessed, data were pooled by the median of the ASV count to evaluate the bacterial composition. To compare relative abundance distributions between sites in salso-bromo-jodic water and sulphurous water, we used the Mann–Whitney U-test corrected with the Bonferroni method in the case of the first water type. Statistical significance was assumed with a p-value < 0.05.

3. Results

3.1. Bacterial Community Structure of the Waters

The filtering raw sequence quality process retained a total of 287,570 sequences in the final dataset. Each sample was represented by a range of between 4877 and 17,048 high-quality sequences, which were assigned to 1114 ASVs across 30 water samples. Analysis of β-diversity within sites at three different time points revealed no statistically significant differences at any taxonomic level. Therefore, the samples from three time points were gathered to compare the bacterial composition of the different sites for each water type.

3.1.1. Bacterial Community Structure in Salso-Bromo-Jodic Waters

The most representative phyla detected in salso-bromo-jodic waters were *Proteobacteria* (57.57%; IQR: 12.70%), *Campylobacterota* (19.21%; IQR: 14.12%), and *Bacteroidota* (10.23%; IQR: 3.32%), with traces of *Desulfobacterota* (6.06%; IQR: 4.49%). The top four phyla accounted for 93.07% of the total ASVs. Interestingly, larger amounts of *Proteobacteria* were found in both the water at natural source temperature and the heated water samples collected within the spa resort (57.57% and 60.65%, respectively), compared to the samples collected at the natural source (35.25%; p-value < 0.05). With *Bacteroidota*, a statistical difference was observed in the samples collected at the natural source vs. the ones at the spa resort at natural source temperature (5.70% and 10.23% respectively; p-value < 0.05). Conversely, no statistically significant difference was detected comparing the samples collected at the natural source vs. the heated ones at the spa resort (12.34%). *Campylobacterota* and *Desulfobacterota* were the most prevalent in the water collected at the natural source (37.96% and 12.20%, respectively), vs. the heated water (19.21% and 6.06%) and at the natural source temperature (9.73% and 3.22%) within the spa resort. Other phyla were *Firmicutes* and *Deinococcota*. These were detected exclusively in heated water (5.78% and 7.03%, respectively) (Figure 2a).

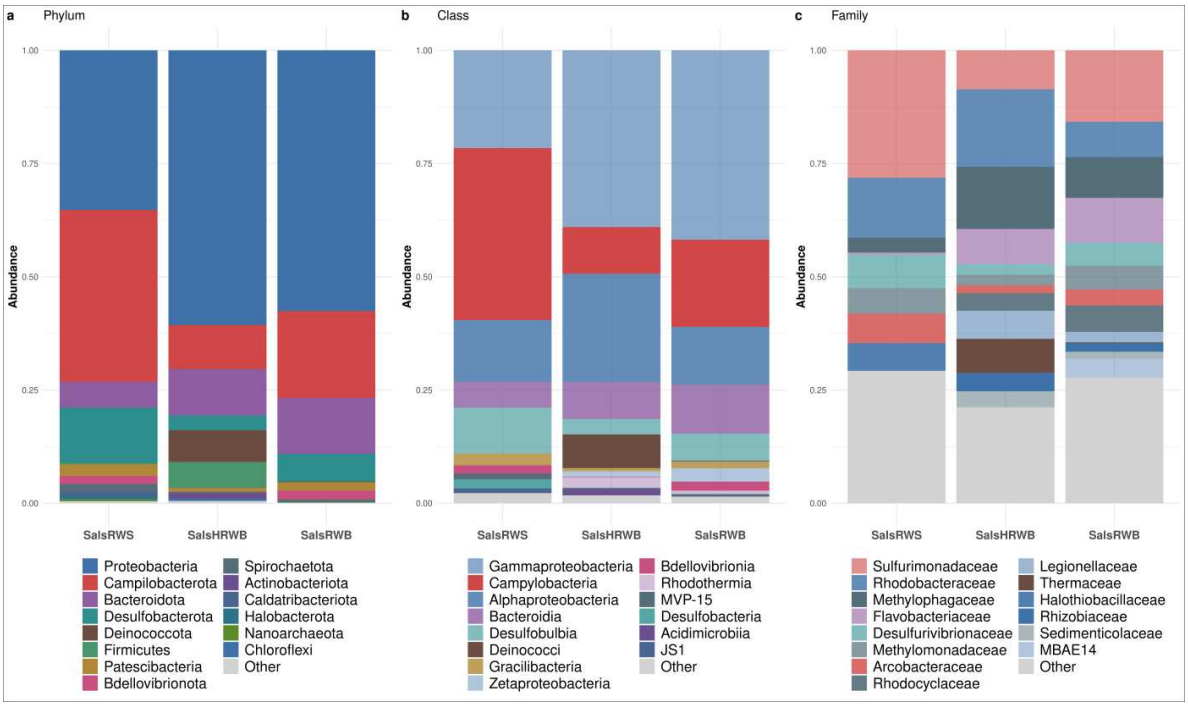


Figure 2. Most representative taxa at the (**a**) phylum, (**b**) class, and (**c**) family level of bacterial communities in salso-bromo-jodic waters divided by sample site: salso-bromo-jodic water collected at the natural source (SalsRWS); heated salso-bromo-jodic water (SalsHRWB); and salso-bromo-jodic water at natural source temperature (SalsRWB) collected at the exit of the respective storage tanks within the spa resort. For the barplots, samples of each site were pooled by median ASV count.

The *Proteobacteria* phylum includes *Gammaproteobacteria* (39.03%; IQR: 10.81%), *Alphaproteobacteria* (13.65%; IQR: 7.94%), and *Zetaproteobacteria* (1.23%; IQR: 1.91%) classes. *Gammaproteobacteria* was the most representative class in the waters collected at spa resort, both at the natural temperature (41.83%) and after heating to 70 °C (39.03%). However, in the samples collected at the natural source, this was not the case (21.60%; p-value < 0.05). *Alphaproteobacteria* and *Zetaproteobacteria* displayed a higher prevalence in the waters collected at the spa resort, both after heating (23.96% and 1.23%, respectively) and at the

natural source temperature (12.80% and 2.95%, respectively) vs. the water collected at the natural spring source (13.65% and 0.00%, respectively; Figure 2b).

Regarding *Gammaproteobacteria*, the *Methylomonadaceae* presented an equal abundance both in the waters collected at the natural source (5.63%) and the ones at the spa resort at the natural source temperature (5.17%), while a lesser percentage was found in the heated water (2.39%).

Furthermore, *Halothiobacillaceae*, and *Rhodanobacteraceae* were relatively plentiful in the water at the natural source (6.10% and 2.08%, respectively), while only traces were observed in the waters collected at the other sites (p-value < 0.05 for the *Halothiobacillaceae*). Conversely, *Rhodocyclaceae*, *Legionellaceae*, *Sedimenticolaceae*, *Gallionellaceae*, and MBAE-14 were found in the waters collected at the spa resort, both after heating (3.86%, 6.22%, 3.45%, 0.94% and 0.14%, respectively) and at the natural source temperature (5.84%, 2.31%, 1.51%, 3.13% and 4.22%, respectively), while the aforementioned taxa were absent from water at the natural source.

Moreover, the *Methylophagaceae* family was prevalent in the heated water (13.72%), vs. the water at the natural source temperature collected at the spa resort (9.01%) and at the natural source (3.26%). A higher abundance of *Methylococcaceae* and *Vibrionaceae* (4.68% and 1.47%, respectively) was also observed in heated water. In the water collected at the spa resort at the natural source temperature, *Pseudomonadaceae* and *Alteromonadaceae* (2.12% and 2.82%) were the most represented families compared to at the other water collection sites (Figure 2c).

Alphaproteobacteria were well represented in the heated water. As for the abundance of the *Rhodobacteraceae* family (17.11%), it was significantly higher in this water vs. the water collected at the spa resort at the natural source temperature (7.81%, p-value < 0.05). However, its abundance was not statistically significant in the water collected at the natural source (13.26%). Moreover, the *Defluviicoccales* (1.28%) and *Rhizobiaceae* (4.00%) families were found only in the heated water. Finally, the *Zetaproteobacteria* class was represented mainly by the *Mariprofundaceae* family, which was detected only in the waters collected at the spa resort, both at the natural source temperature (2.97%) and after heating at 70 °C (1.24%). This family was absent from the water collected at the source (Figure 2c).

Concerning the *Campylobacterota* phylum, the *Campylobacteria* was the most abundant class in the water from the natural source (37.96%) vs. the waters collected at the spa resort, both after heating (19.21%) and at the natural source temperature (10.30%; Figure 2b). Within the *Campylobacterales* order, *Sulfurimonadaceae* accounted for 28.13% of total family abundance in the water from the natural source but only for 15.77% and 8.62% in the waters collected at the spa resort after heating and at the natural source temperature, respectively. Other *Campylobacterales*, such as *Arcobacteraceae* and *Sulfurospirillaceae*, were also prevalent in the water at the natural source (6.53% and 3.31%, respectively), vs. the waters collected at the other sites (Figure 2c).

The *Bacteroidota* phylum was mainly represented by *Bacteroidia* and *Rhodothermia* classes (Figure 2b). Within the *Bacteroidia* class, the *Bacteroidetes* VC2.1 Bac22 accounted for 3.90% of the total family abundance in the water from the natural source, while only a few traces were detected in the other water collection sites (0.68% at the spa resort at the natural source temperature and 0.30% in the heated water, p-value < 0.05). Furthermore, the *Flavobacteriaceae* family was particularly abundant in the water collected at the spa resort at the natural source temperature (9.92%) vs. the heated water (7.80%) and the water from the natural source (0.49%, p-value < 0.05). Finally, the *Rhodothermaceae* family, part of the *Rhodothermia* class, was mainly represented in the heated water (2.33%), vs. the water from the natural source and water collected at the spa resort at the natural source temperature (0.00% and 0.76%, respectively; Figure 2c).

Desulfobulbia (*Desulfobacterota* phylum) was the most abundant class in the water from the natural source (10.12%) vs. the water collected at the spa resort, both at the natural source temperature (6.02%) and after heating at 70 °C (3.40%; Figure 2b). This was confirmed for the *Desulfobulbales* order too, whose *Desulfurivibrionaceae* and *Desulfocapsaceae* families accounted for 7.38% and 2.74%, respectively, of the total class abundance in the water from the natural source (Figure 2c).

Finally, the *Deinococci* class (*Deinococcota* phylum) was more represented in the heated water (7.44%) vs. the water collected at the other sites, in which only a few traces were observed (p-value < 0.05). At the family taxonomic level, *Thermaceae* accounted for 7.48% of the total family abundance in the heated water (Figure 2c).

3.1.2. Bacterial Community Structure in Sulphurous Waters

The most representative phyla in sulphurous waters were *Campylobacterota* (52.30%; IQR: 0.17%), *Proteobacteria* (33.28%; IQR: 3.72%), *Desulfobacterota* (6.27%; IQR: 0.91%), and traces of *Bacteroidota* (4.96%; IQR: 2.20%). As for *Firmicutes*, these were only observed in the sulphurous water collected at the spa resort (1.32%, $p < 0.05$; Figure 3a).

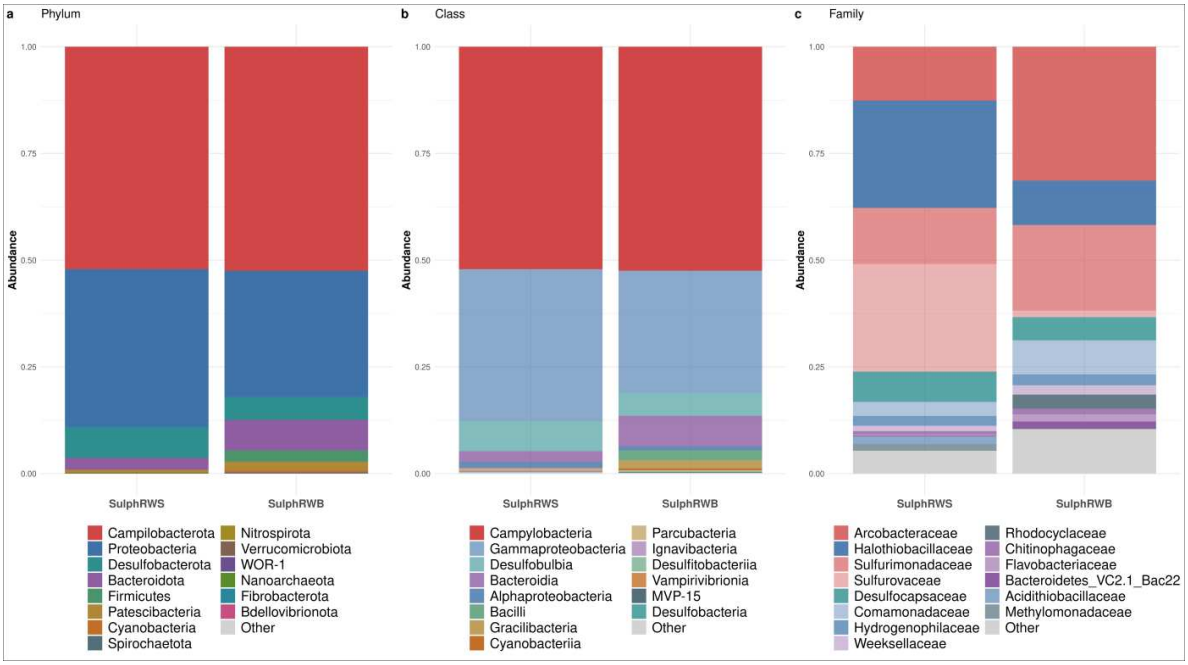


Figure 3. Most representative taxa at the (**a**) phylum, (**b**) class, and (**c**) family level of communities in sulphurous waters divided by sample site: sulphurous water collected at natural source (SulphRWS) and sulphurous water collected at the natural source temperature (SulphRWB) within the spa resort. For the barplots, samples of each site were pooled by median ASV count.

All *Campylobacterota* were assigned to the *Campylobacteria* class. In the water from the natural source, most *Campylobacterales* were detailed in *Sulfurovaceae* (25.18%) vs. the sulphurous water collected at the spa resort where they accounted for only 1.52% ($p < 0.05\%$). Other *Campylobacterales* were represented by *Sulfurimonadaceae* (13.18% at the spring and 20.07% at the spa resort) and *Arcobacteraceae* (12.63% at the spring and 31.37% at the spa resort, $p < 0.05$) (Figure 3b,c).

As for *Proteobacteria*, they were mainly composed of *Gammaproteobacteria* (32.10%; IQR: 3.47%; Figure 3b). The prevalence of *Halothiobacillaceae* was higher in the water collected at the source (25.11%) vs. that collected at the spa resort (10.39%, p-value < 0.05); this was also noticed for *Acidithiobacillaceae* and *Methylomonadaceae* (1.83% and 1.53% at the spring). Conversely, *Comamonadaceae*, *Rhodocyclaceae* and *Sulfuricellaceae* were most prevalent in the

water collected at the spa resort (8.01%, 3.29% and 1.48%, respectively). *Hydrogenophilaceae*, instead, were equally abundant at both sites (2.54% in the water at the spa resort and 2.30% at the source; Figure 3c).

The *Desulfobacterota* phylum was represented only by the *Desulfobulbia* class. The *Desulfocapsaceae* family was more prevalent in the water at the source (7.10%) vs. that at the spa resort (5.40%; Figure 3b,c).

The *Bacteroidota* phylum was mainly represented by the *Bacteroidia* class, and it was particularly abundant in the water collected at the spa resort. In particular, the most represented families were *Weeksellaceae* (2.17%), *Bacteroidetes* VC2.1 Bac22 (Candidatus *Sulfidibacteriales*, 1.71%), *Flavobacteriaceae* (1.65%), and *Chitinophagaceae* (1.37%; Figure 3b,c).

Finally, the *Firmicutes* phylum was mainly represented by *Erysipelatoclostridiaceae*. This was observed only in the water collected at the spa resort, with a prevalence of 1.66%.

3.2. Diversity Analyses

Alpha and beta diversity were computed to compare diversity in different water types.

3.2.1. Diversity Analysis in Salso-Bromo-Jodic Water

Within the salso-bromo-jodic water, at the family taxonomic level, a more diverse community was observed, based on richness, in the water collected at the spa resort at the natural source temperature vs. the water collected at the source ($p < 0.05$) and also the heated water, even though no significant difference was observed for the latter comparison. The Shannon index was higher for the salso-bromo-jodic water collected at the spa resort at the natural source temperature (median: 2.73%; IQR: 0.06%) and the heated one (median: 2.62%; IQR: 0.51%) when compared to the water from the source (median: 1.88%; IQR: 1.01%). Regarding beta diversity, a diversification in the microbial community was observed at the different sites of water collection ($p < 0.05$); from the PcOA, more homogeneity was recorded between the waters collected at the spa resort both at the natural source temperature and after heating at 70 °C (Figure 4).

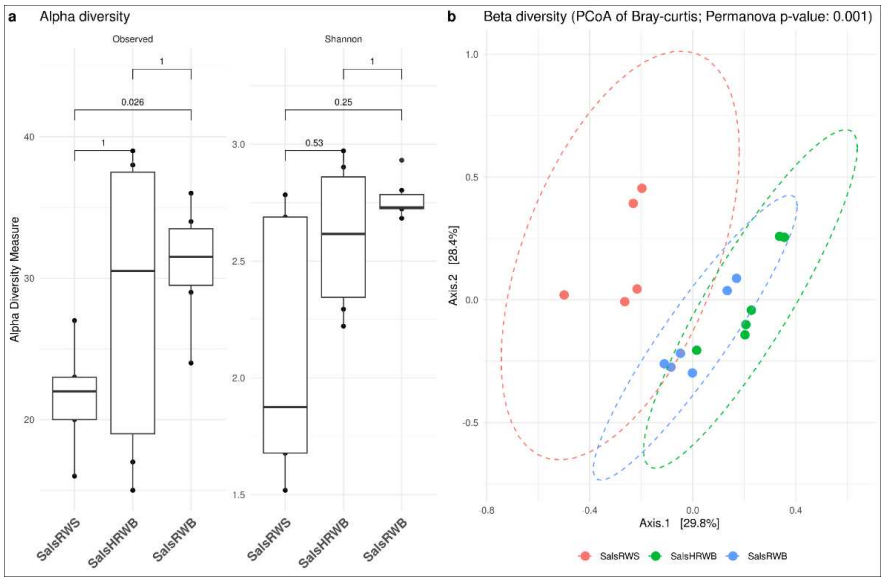


Figure 4. Diversity analyses. (**a**) Alpha and (**b**) Beta diversity at the family taxonomic level for salso-bromo-jodic waters at the different collection sites: salso-bromo-jodic water collected at the natural source (SalsRWS); heated salso-bromo-jodic water (SalsHRWB); and salso-bromo-jodic water at natural source temperature (SalsRWB) collected at the exit of the respective storage tanks within the spa resort.

3.2.2. Diversity Analysis in Sulphurous Water

In the sulphurous waters, greater diversity was found in the water collected at the spa resort vs. that from the source. Indeed, the Shannon index was 2.25 (IQR: 0.63%) for the spa resort and 1.59 (IQR: 0.77%) for the source. This diversity also expressed itself in the beta diversity, where a statistically significant difference was recorded between the two sites (Figure 5).

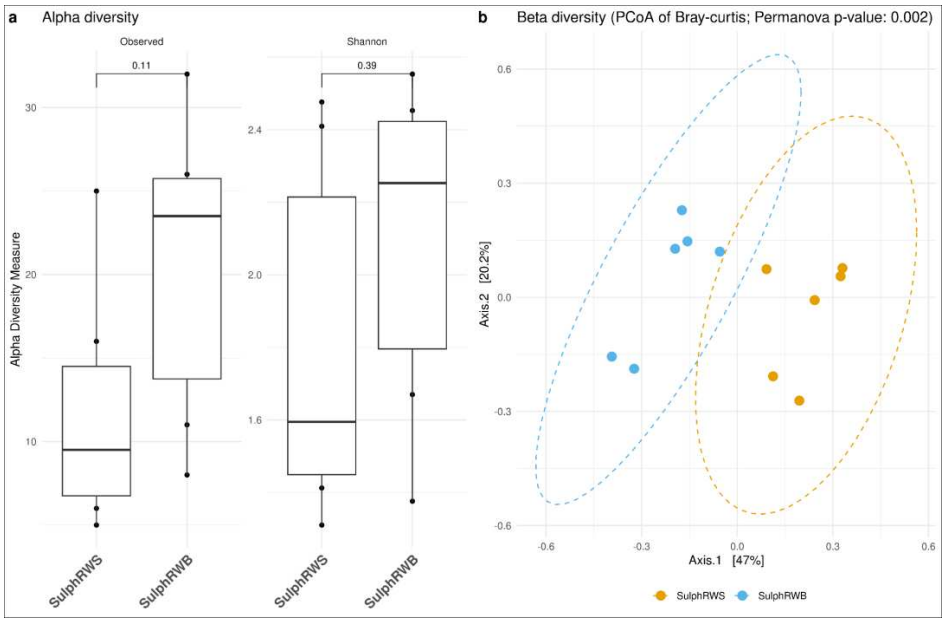


Figure 5. Diversity analyses. (**a**) Alpha and (**b**) Beta diversity at the family taxonomic level for sulphurous waters divided by sample site: sulphurous water collected at natural source (SulphRWS) and sulphurous water collected at the natural source temperature (SulphRWB) within the spa resort.

4. Discussion

The sequential sampling of both Rivanazzano salso-bromo-jodic and sulphurous mineral waters revealed a constant and specific microbiota composition for each water type. The bacterial populations observed are typical of deep-water sources in a marine-derived geological environment. In fact, the Rivanazzano spa resort lies within Oltrepò Pavese (Pavia, Italy), in the foothills of the northernmost portion of the Apennine range. It was formed on the bottom of the Ligurian–Piedmontese oceanic basin and is made up of vast surfaces of sedimentary blankets of ancient alluvial origin (Pleistocene era) and more recent marine origin (Miocene–Upper Pliocene era). The geology of this territory includes a gypsum–sulphur formation formed at the end of the Miocene during the "Messinian salinity crisis" and oil-related deposits which usually accompany the salso-bromo-jodic springs [34].

At the natural sources, the salso-bromo-jodic and sulphurous waters show marked differences in mineralization and composition. While both have similar natural source temperatures (13.9 °C vs. 13.2 °C), the salso-bromo-jodic water exhibits significantly higher conductivity (8100 µS/cm) and fixed residue (12,400 mg/L), indicating greater total dissolved solids. Chemically, the salso-bromo-jodic water is rich in Na^+ (4490 mg/L), Cl^- (7300 mg/L), Mg^{2+} (78.5 mg/L), and Sr^{2+} (15.8 mg/L). In contrast, sulphurous water is characterized by its high H_2S concentration (11.9 mg/L)—which is absent from the salso-bromo-jodic type—and higher SO_4^{2-} content (283 mg/L vs. 35.7 mg/L).

Actually, the natural mineral waters' favourable effects are traditionally attributed to their chemical–physical composition, which is unique to every spring and specifically active on the skin [7,8,15,35], even though, for 20 years now, synergy with some specific

non-pathogenic bacterial populations has been recognized [4,5,13,14,17–22,36]. We believe that, like other natural mineral waters, the therapeutic properties of the Rivanazzano waters might too be related to their combined mineral and microbiological composition.

The general bacterial composition was not homogeneous at the different sites of collection. A relevant change in the native non-pathogenic microflora, although not statistically significant, was appreciated in both waters after the delivery from the natural sources to the spa resort. This could indicate that the bacterial phyla displayed a variable individual decreasing or increasing trend in both waters across the different sites of collection. Furthermore, some phyla completely disappeared, and some new ones appeared, after the water delivery from the natural sources to the spa resort and/or the heating at 70 °C.

In the Rivanazzano salso-bromo-jodic water, samples collected at the natural source tended to display a higher prevalence of bacteria from the *Sulfurimonadaceae*, the the *Arcobacteriaceae*, the *Desulfurivibrionaceae*, and the *Halothiobacillaceae* families.

The *Sulfurimonadaceae* and the *Arcobacteriaceae* families are from the *Campylobacteria* class (*Campylobacterota* phylum). *Sulfurimonadaceae* are known to be involved in the process of reducing nitrates, oxidising both sulphur and hydrogen, and they are coupled to autotrophy [37]. They are characterised by the ability to survive in extreme conditions and hypoxic environments, including deep-sea hydrothermal vents. The *Arcobacteriaceae* family includes non-pathogenic bacteria capable of forming dense solid sulphur-filament mats. These mats serve as essential substrates around hydrothermal vents, facilitating the colonization and settlement of various organisms by providing a stable habitat and attracting diverse marine life [38].

Desulfovibrionaceae is a family within the order *Desulfovibrionales*. They are all strictly anaerobic, mesophilic or moderately psychrophilic; most species are chemo-organoheterotrophic, but some others can be chemolithoheterotrophic. They are found in freshwater, brackish water and marine sediments, but also in biofilms, soil, sewage sludge, and animals [39].

The *Halothiobacillaceae* are a family of halotolerant, mesophilic, and obligate chemolithoautotrophic organisms. As halotolerant organisms, they tend to live in areas such as hypersaline lakes, coastal dunes, saline deserts, salt marshes, and inland salt seas and springs [40,41].

In the Rivanazzano sulphurous water collected at the natural source; the highest bacterial prevalence measured was from the *Sulfurovaceae*, *Desulfocapsaceae* and *Halothiobacillaceae* families (this was also true of *Halothiobacillaceae* in the salso-bromo-jodic water).

The family *Sulfurovaceae* of the *Campylobacterota* phylum [42] can be found in various habitats, including mangroves [43], nearshore sediments [44], shallow hydrothermal zones [45], glaciers [46], seagrass beds [47], and deep-sea hydrothermal vents [48–50].

The *Desulfocapsaceae* are a strict anaerobe family of *Thermodesulfobacteriota*. The family is considered one of the major sulphate-reducing bacterial groups in anoxic marine sediments, thereby accounting for one of the most important metabolic processes of anaerobic degradation of organic matter in the sea [51,52]. Interestingly, in both waters, all of these relevant bacterial families share an active biological role in chemo-auto-trophic processes [53]. Indeed, they belong to the most ancient bacterial populations since the origin of life [54].

We believe that these results are intriguing and require further rigorous investigation to explore the potential correlation between the waters' therapeutic effects and their native non-pathogenic bacterial communities. In our opinion, the next step in this research line might be the lysate extraction from these selected bacterial families and their assessment on in vitro and ex vivo experimental cell models to investigate their potential favourable properties.

5. Limitations of the Study

This study represents an initial exploration of the microbiota present in the mineral waters of Rivanazzano; therefore, it was limited to the characterization of bacterial communities. A more comprehensive analysis would also include Archaea and Fungi. Future studies are desirable to extend these preliminary findings by comprehensively investigating the Rivanazzano waters' microbiota.

Additionally, the sampling period was restricted to a three-month period (May to July), which did not adequately capture the potential seasonal fluctuations in bacterial community composition. Nevertheless, given the preliminary nature of this study, samples were collected during the peak working time of the spa resort to provide a relevant starting point for investigating potential correlations between microbial profiles and the therapeutic properties of the waters.

6. Conclusions

The Rivanazzano mineral waters revealed a rich and distinctive bacterial composition specific to the geographical area. It is reasonable to suggest that these bacterial communities, along with other microorganisms, may play a key role in the therapeutic effects of these waters, due to their ability to interact with the skin microbiota [6]. Actually, there is growing evidence that the skin microbiota is actively involved in the regulation of various diseases, as well as in wound healing, which is archetypal of all regenerative processes [55–57].

Reflecting on these considerations, a novel classification of natural mineral waters based on both their physicochemical characteristics and specific microbiota might be proposed.

Author Contributions: Conceptualization, G.N. and A.F.; methodology, A.K., I.M., G.P. and M.C.; software, G.P.; validation, A.K., I.M., M.S. and G.P.; formal analysis, A.K., I.M., G.P., M.C. and M.S.; investigation, A.K., I.M., G.P. and M.S.; resources, G.N. and M.C.; data curation, G.P.; writing—original draft preparation, A.K., I.M., M.S., G.P., M.C., G.N., A.F. and F.B.; writing—review and editing, G.N., A.K., I.M., M.S., G.P., A.F. and F.B.; visualization, G.P.; supervision, G.N., A.F. and F.B.; project administration, M.S.; funding acquisition, G.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Acknowledgments: The authors wish to thank (1) Morag McGhee and Eleanor Susan Lim for their contribution to the submission of this study; (2) Marco Mario Tresoldi for collecting the water samples; (3) Stefano Gaiarsa for generating the graphical representation of the Rivanazzano Terme water network; and (4) The Terme di Rivanazzano s.r.l. for providing the Rivanazzano salso-bromo-jodic and sulphurous natural mineral waters.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zeichner, J.; Seite, S. From probiotic to prebiotic using thermal spring water. *J. Drugs Dermatol.* **2018**, *17*, 657–662.
2. Staquet, M.J.; Peguet-Navarro, J.; Richard, A.; Schmitt, D.; Rougier, A. In vitro effect of a spa water on the migratory and stimulatory capacities of human Langerhans calls. *Eur. J. Dermatol.* **2002**, *12*, LIX–LXI.

3. Seite, S.; Benech, F.; Berdah, S.; Bayer, M.; Veyrat, S.; Segot, E.; Sakalikova, M.; Gibejova, L.; Zelenkova, H. Management of rosacea-prone skin: Evaluation of a skincare product containing Ambophenol, Neurosensine, and La Roche-Posay Thermal spring water as monotherapy or adjunctive therapy. *J. Drugs Dermatol.* **2013**, *12*, 920–924. [PubMed]
4. Gueniche, A.; Liboutet, M.; Cheilian, S.; Fagot, D.; Juchaux, F.; Breton, L. *Vitreoscilla filiformis* Extract for Topical Skin Care: A Review. *Front. Cell Infect. Microbiol.* **2021**, *11*, 747663. [CrossRef] [PubMed]
5. Mahé, Y.F.; Martin, R.; Aubert, L.; Billoni, N.; Collin, C.; Pruche, F.; Bastien, P.; Drost, S.S.; Lane, A.T.; Meybeck, A. Induction of the skin endogenous protective mitochondrial MnSOD by *Vitreoscilla filiformis* extract. *Int. J. Cosmet. Sci.* **2006**, *28*, 277–287. [CrossRef] [PubMed]
6. Woolery-Lloyd, H.; Andriessen, A.; Day, D.; Gonzalez, N.; Green, L.; Grice, E.; Henry, M. Review of the microbiome in skin aging and the effect of a topical prebiotic containing thermal spring water. *J. Cosmet. Dermatol.* **2023**, *22*, 96–102. [CrossRef]
7. Liang, J.; Kang, D.; Wang, Y.; Yu, Y.; Fan, J.; Takashi, E. Carbonate ion-enriched hot spring water promotes skin wound healing in nude rats. *PLoS ONE* **2015**, *10*, e0117106. [CrossRef]
8. Mourelle, M.L.; Gómez, C.P.; Legido, J.L. Unveiling the role of minerals and trace elements of thermal waters in skin health. *Appl. Sci.* **2024**, *14*, 6291. [CrossRef]
9. Faga, A.; Nicoletti, G.; Gregotti, C.; Finotti, V.; Nitto, A.; Gioglio, L. Effects of thermal water on skin regeneration. *Int. J. Mol. Med.* **2012**, *29*, 732–740. [CrossRef]
10. Nicoletti, G.; Corbella, M.; Jaber, O.; Marone, P.; Scevola, D.; Faga, A. Non-pathogenic microflora of a spring water with regenerative properties. *Biomed. Rep.* **2015**, *3*, 758–762. [CrossRef]
11. Nicoletti, G.; Saler, M.; Malovini, A.; Faga, A.; Scalise, A.; Riva, F. Effects of a spring water on human skin fibroblasts in vitro cultures: Preliminary results. *Acta Vulnologica* **2016**, *14*, 196–201.
12. Nicoletti, G.; Saler, M.; Pellegatta, T.; Tresoldi, M.M.; Bonfanti, V.; Malovini, A.; Faga, A.; Riva, F. Ex vivo regenerative effects of a spring water. *Biomed. Rep.* **2017**, *7*, 508–514.
13. Nicoletti, G.; Saler, M.; Tresoldi, M.M.; Faga, A.; Benedet, M.; Cristofolini, M. Regenerative effects of spring water-derived bacterial lysates on human skin fibroblast in in vitro culture: Preliminary results. *J. Int. Med. Res.* **2019**, *47*, 5777–5786. [CrossRef]
14. Nicoletti, G.; Saler, M.; Tresoldi, M.M.; Villani, L.; Tottoli, E.M.; Jousson, O.; Faga, A. Effects of Comano Spring Water-derived Bacterial Lysates on Skin Regeneration: An Ex-vivo Study. *Vivo* **2023**, *37*, 2498–2509. [CrossRef] [PubMed]
15. Saler, M.; Ferraro, O.; Faga, A.; Sansotta, D.; Villani, S.; Nicoletti, G. Enhancement of human native skin fibroblast proliferation in natural salso-bromo-iodic mineral water added to in vitro culture. *Adv. Clin. Exp. Med.* **2024**. [CrossRef] [PubMed]
16. Éditeur Officiel du Québec: Regulation Respecting Bottled Water: Food Products Act. Chapter P-29, r.2. Updated to 1 June 2016. Available online: https://www.legisquebec.gouv.qc.ca (accessed on 20 February 2025).
17. Nocera, T.; Rossi, A.B.; Mengeaud, V. Clinical development program of a new dermocosmetic range of products containing I-modulia (*Aquaphilus dolomiae* extract) in atopic dermatitis. *J. Am. Acad. Dermatol.* **2014**, *70*, AB62.
18. Patrizi, A.; Bacquey, A.; Schmitt, A.M.; Decoster, C.J.; Phulpin, C.; Theunis, J.; Mengeaud, V. Clinical and biometrologic evaluation of a novel emollient balm containing an *Aquaphilus dolomiae* extract in 1- to 4-year-old children suffering from atopic dermatitis: International, multicenter, randomized versus control group study. *J. Am. Acad. Dermatol.* **2014**, *70*, AB62.
19. Aries, M.-F.; Delga, H.; Levêque, M.; Castex-Rizzi, N.; Bessou-Touya, S.; Nguyen, T. Antiinflammatory and immunomodulatory effects of I-modulia, an *Aquaphilus dolomiae* extract, on atopic dermatitis in vitro. *J. Am. Acad. Dermatol.* **2014**, *70*, AB61.
20. Aries, M.-F.; Duplan, H.; Hernandez-Pigeon, H.; Galliano, M.F.; Castex-Rizzi, N.; Bessou-Touya, S.; Nguyen, T. I-modulia, an *Aquaphilus dolomiae* extract, stimulates innate immune response through Toll-like receptor activation. *J. Am. Acad. Dermatol.* **2014**, *70*, AB63.
21. Mahe, Y.F.; Perez, M.-J.; Tacheau, C.; Fanchon, C.; Martin, R.; Rousset, F.; Seite, S. A new *Vitreoscilla filiformis* extract grown on spa water-enriched medium activates endogenous cutaneous antioxidant and antimicrobial defenses through a potential Toll-like receptor 2/protein kinase C, zeta transduction pathway. *Clin. Cosmet. Investig. Dermatol.* **2013**, *6*, 191–196.
22. Gueniche, A.; Knaudt, B.; Schuck, E.; Volz, T.; Bastien, P.; Martin, R.; Röcken, M.; Breton, L.; Biedermann, T. Effects of nonpathogenic gram-negative bacterium *Vitreoscilla filiformis* lysate on atopic dermatitis: A prospective, randomized, double-blind, placebo-controlled clinical study. *Br. J. Dermatol.* **2008**, *159*, 1357–1363. [CrossRef] [PubMed]
23. Cristofolini, M.; Jousson, O. Una componente ignorata delle acque termali: Il microbiota nascosto. *La Clin. Termal.* **2017**, *64*, 1–4.
24. Saler, M.; Ferraro, O.; Faga, A.; Sansonetta, D.; Villani, S.; Nicoletti, G. *Natural Sulphurous Mineral Water Induced Proliferation En-Hancement of Native Human Skin Fibroblast In-Vitro Cultures*; In-Vivo, Stanford University Highwire Press: Stanford, CA, USA, 2025; *Original Article, Submitted*.
25. Terme di Rivanazzano, Pavia (PV), Lombardia. Terme di Rivanazzano. Available online: https://www.federterme.it/spa/Lombardia/TermediRivanazzano(PV)Lombardia/106/ (accessed on 20 February 2025).
26. Takahashi, S.; Tomita, J.; Nishioka, K.; Hisada, T.; Nishijima, M. Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS ONE* **2014**, *9*, e105592. [CrossRef]

27. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [CrossRef] [PubMed]
28. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods.* **2016**, *13*, 581–583. [CrossRef]
29. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef]
30. Nakamura, T.; Yamada, K.D.; Tomii, K.; Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **2018**, *34*, 2490–2492. [CrossRef]
31. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. [CrossRef]
32. McMurdie, P.J.; Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef]
33. Oksanen, J.; Simpson, G.L.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.R.; O'Hara, R.; Solymos, P.; Stevens, M.H.H.; Szoecs, E.; et al. Vegan: Community Ecology Package. R Package Version 2.6-4. 2022. Available online: https://CRAN.R-project.org/package=vegan (accessed on 20 February 2025).
34. Oltresentieri. Available online: https://www.oltresentieri.com/ (accessed on 20 February 2025).
35. Oliveira, A.S.; Vaz, C.V.; Silva, A.; Correia, S.; Ferreira, R.; Breitenfeld, L.; Martinez-De-Oliveira, J.; Pereira, C.; Cruz, M.T. In vitro evaluation of potential benefits of a silica-rich thermal water (Monfortinho Thermal Water) in hyperkeratotic skin conditions. *Int. J. Biometeorol.* **2020**, *64*, 1957–1968. [CrossRef]
36. Logue, J.B.; Robinson, C.T.; Meier, C.; der Meer, J.R.V. Relationship between sediment organic matter, bacteria composition, and the ecosystem metabolism of alpine streams. *Limnol. Oceanogr.* **2004**, *49*, 2001–2010. [CrossRef]
37. Han, Y.; Perner, M. The globally widespread genus *Sulfurimonas*: Versatile energy metabolisms and adaptations to redox clines. *Front. Microbiol.* **2015**, *6*, 989. [CrossRef]
38. Madigan, M.T.; Martinko, J.M.; Bender, K.S.; Buckley, D.H.; Stahl, D.A.; Brock, T. *Brock Biology of Microorganisms*, 14th ed.; Brock, T., Ed.; Pearson: London, UK, 2015.
39. Kuever, J. The Family Desulfovibrionaceae. In *The Prokaryotes*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 107–133.
40. Kelly, D.P.; Wood, A.P. Halothiobacillaceae fam. nov. In *Bergey's Manual of Systematics of Archaea and Bacteria*; John Wiley & Sons, Ltd.: Chichester, UK, 2015; pp. 1–9. [CrossRef]
41. Boden, R. Reclassification of *Halothiobacillus hydrothermalis* and *Halothiobacillus halophilus* to *Guyparkeria* gen. nov. in the *Thioalkalibacteraceae* fam. nov., with emended descriptions of the genus *Halothiobacillus* and family *Halothiobacillaceae*. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 3919–3928.
42. Waite, D.W.; Vanwonterghem, I.; Rinke, C.; Parks, D.H.; Zhang, Y.; Takai, K.; Sievert, S.M.; Simon, J.; Campbell, B.J.; Hanson, T.E.; et al. Comparative Genomic Analysis of the Class *Epsilonproteobacteria* and Proposed Reclassification to *Epsilonbacteraeota* (phyl. nov.). *Front. Microbiol.* **2017**, *8*, 682. [CrossRef] [PubMed]
43. Lin, X.; Hetharua, B.; Lin, L.; Xu, H.; Zheng, T.; He, Z.; Tian, Y. Mangrove sediment microbiome: Adaptive microbial assemblages and their routed biogeochemical processes in yunxiao mangrove national nature reserve, China. *Microb. Ecol.* **2019**, *78*, 57–69. [CrossRef]
44. Marziah, Z.; Mahdzir, A.; Musa, M.N.; Jaafar, A.B.; Azhim, A.; Hara, H. Abundance of sulfur-degrading bacteria in a benthic bacterial community of shallow sea sediment in the off-Terengganu coast of the South China Sea. *Microbiologyopen* **2016**, *5*, 967–978. [PubMed]
45. Wang, L.; Cheung, M.K.; Liu, R.; Wong, C.K.; Kwan, H.S.; Hwang, J.-S. Diversity of Total Bacterial Communities and Chemoautotrophic Populations in Sulfur-Rich Sediments of Shallow-Water Hydrothermal Vents off Kueishan Island, Taiwan. *Microb. Ecol.* **2017**, *73*, 571–582. [CrossRef]
46. Wright, K.E.; Williamson, C.; Grasby, S.E.; Spear, J.R.; Templeton, A.S. Metagenomic evidence for sulfur lithotrophy by Epsilonproteobacteria as the major energy source for primary productivity in a sub-aerial arctic glacial deposit, Borup Fiord Pass. *Front. Microbiol.* **2013**, *4*, 63.
47. Sun, Y.; Song, Z.; Zhang, H.; Liu, P.; Hu, X. Seagrass vegetation affect the vertical organization of microbial communities in sediment. *Mar. Environ. Res.* **2020**, *162*, 105174.
48. Nakagawa, T.; Takai, K.; Suzuki, Y.; Hirayama, H.; Konno, U.; Tsunogai, U.; Horikoshi, K. Geomicrobiological exploration and characterization of a novel deep-sea hydrothermal system at the TOTO caldera in the Mariana Volcanic Arc. *Environ. Microbiol.* **2006**, *8*, 37–49. [CrossRef]
49. Dahle, H.; Roalkvam, I.; Thorseth, I.H.; Pedersen, R.B.; Steen, I.H. The versatile in situ gene expression of an Epsilonproteobacteria-dominated biofilm from a hydrothermal chimney. *Environ. Microbiol. Rep.* **2013**, *5*, 282–290. [PubMed]

50. Meier, D.V.; Pjevac, P.; Bach, W.; Hourdez, S.; Girguis, P.R.; Vidoudez, C.; Amann, R.; Meyerdierks, A. Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. *ISME J.* **2017**, *11*, 1545–1558.
51. Galushko, A.; Kuever, J. Desulfocapsaceae. In *Bergey's Manual of Systematics of Archaea and Bacteria*; John Wiley & Sons, Ltd.: Chichester, UK, 2015; pp. 1–12. [CrossRef]
52. Song, J.; Hwang, J.; Kang, I.; Cho, J.-C. A sulfate-reducing bacterial genus, *Desulfosediminicola* gen. nov., comprising two novel species cultivated from tidal-flat sediments. *Sci. Rep.* **2021**, *11*, 19978. [CrossRef] [PubMed]
53. Wang, S.; Jiang, L.; Xie, S.; Alain, K.; Wang, Z.; Wang, J.; Liu, D.; Shao, Z. Disproportionation of inorganic sulfur compounds by mesophilic chemolithoautotrophic campylobacterota. *mSystems* **2023**, *8*, e0095422. [PubMed]
54. Cofield, C. Simulating Early Ocean Vents Shows Life's Building Blocks Form Under Pressure. *NASA*. 2020. Available online: https://astrobiology.nasa.gov (accessed on 20 February 2025).
55. Manos, J. The human microbiome in disease and pathology. *Apmis* **2022**, *130*, 690–705.
56. Canchy, L.; Kerob, D.; Demessant, A.; Amici, J.-M. Wound healing and microbiome, an unexpected relationship. *J. Eur. Acad. Dermatol. Venereol.* **2023**, *37* (Suppl. S3), 7–15. [CrossRef]
57. Whiting, C.; Abdel Azim, S.; Friedman, A. The skin microbiome and its significance for dermatologists. *Am. J. Clin. Dermatol.* **2024**, *25*, 169–177.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Journal of Infection and Public Health 18 (2025) 102701

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: www.elsevier.com/locate/jiph

Original article

Streptococcus pyogenes bloodstream infections in an Italian hospital: A ten-year genomic picture

Marta Corbella [a,1], Cristina Merla [a,1], Angela Kuka [a,b], Irene Mileto [a,b], Greta Petazzoni [a,c], Chiara Rebuffa [a], Vincenzo Brunco [a], Stefano Gaiarsa [a,*], Fausto Baldanti [a,c], Patrizia Cambieri [a]

[a] *Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy*
[b] *Specialization School of Microbiology and Virology, University of Pavia, Pavia, Italy*
[c] *Department of Medical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy*

ARTICLE INFO

Article history:
Received 17 October 2024
Received in revised form 7 February 2025
Accepted 9 February 2025

Keywords:
Streptococcus pyogenes
bloodstream infection
genomic epidemiology

ABSTRACT

Background: During 2022 and 2023, a large upsurge of cases of Group A *Streptococcus* (GAS) infection has struck many countries worldwide. This large epidemic event was caused by a diminished level of immunity after COVID-19 social restrictions and was fostered mainly by the *emm*1 genotype of the bacterium.
Methods: We characterized the genomes of the GAS isolates that caused bloodstream infections in the last ten years in a 900-bed hospital in Northern Italy. We obtained short-read genomes, which we used for *emm* typing and to analyze the antimicrobial resistance and virulence gene content. Moreover, we contextualized the isolates in an epidemiological point of view using both ortholog-based and SNP-based phylogeny.
Results: In the last ten years, we registered two upsurges of GAS cases in the study hospital: one in 2023 (in line with the global epidemiological situation), and one in 2016 (local, but observed also in another recent Italian study). The genomic analysis of the bloodstream infection isolates showed the circulation of multiple *emm* types, of which *emm*1, *emm*12, and *emm*4 characterized the two large epidemic events. Notably, the 2016 peak was mainly fostered by the $M1_{Global}$ and *emm*4 genotypes, in contrast with the rise of $M1_{UK}$ observed in the UK in the pre-pandemic years. Phylogeny analysis showed the presence of multiple monophyla of local strains. Despite our genomes being obtained from blood-culture isolates, we did not observe any increased presence of virulence or antimicrobial resistance genes.
Conclusions: Our results further confirm the role of *emm*1 as the most prevalent genotype worldwide. However, we also unveiled the central role of genotype *emm*4 in the 2016 epidemic event, described in the study hospital. Finally, our results suggest the presence of multiple locally persistent strains. Among these, we underlined the presence of a multi-drug resistant strain of *emm*92.
© 2025 The Authors. Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Streptococcus pyogenes, also known as Group A *Streptococcus* (GAS), is a human-restricted pathogen. GAS infections lead to a variety of diseases with differing symptoms and severity levels [1]. The primary infection sites are the upper respiratory tract and skin. In young children, GAS infections can manifest as scarlet fever. Non-invasive infections can spread to other body districts and become more severe, resulting in invasive diseases (iGAS) such as empyema, streptococcal toxic shock syndrome, necrotizing fasciitis, and sepsis [2]. These are associated with high morbidity and mortality, with death rates that could reach 23 % [3]. Molecular typing of GAS is based on the *emm* gene, encoding the M-protein, a surface protein that forms fibrils on the bacterial surface [4].

Among the over 250 known *emm* types, strains of *emm*1 genotype are most commonly linked to invasive infections in high-income countries [5]. During the 1980s, *emm*1 has become a major cause of invasive infections due to several genomic alterations, including the acquisition of virulence factors through phage integration and expression of gene *slo* (encoding streptolysin O) [6]. These resulted in a new clone that spread worldwide. Frequently, isolates

Abbreviations: BSI, Bloodstream Infection; GAS, Group A Streptococcus; IGAS, invasive GAS; MFS, Major facilitator superfamily; SIC, Streptococcal Inhibitor of Complement; SNP, Single Nucleotide Polymorphism

* Correspondence to: SC Microbiologia e Virologia, Fondazione IRCCS Policlinico San Matteo, Viale Golgi 19, Pavia 27100, Italy.
E-mail address: s.gaiarsa@smatteo.pv.it (S. Gaiarsa).
[1] These authors contributed equally

https://doi.org/10.1016/j.jiph.2025.102701
1876-0341/© 2025 The Authors. Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

belonging to the *emm*1 genotype are characterized by inactivating mutations in the *covRS* two-component regulatory system, which causes the overexpression of crucial virulence factors during invasive infections [7].

In the 2014–2019 period, a sub-variant of the *emm*1 genotype ($M1_{UK}$) emerged in multiple countries, causing seasonal upsurges of cases of both non-invasive and invasive infections. $M1_{UK}$ is phylogenetically distinguishable from other *emm*1 sub-variants ($M1_{Global}$) and was initially characterized by the presence of 27 point mutations [8]. One of these revealed to be the key to $M1_{UK}$ success: i.e. an intergenic mutation which removed negative regulation from the expression of phage-encoded virulence gene *speA*, which is responsible for the synthesis of streptococcal pyrogenic exotoxin A [9].

In late 2022, several countries in Europe observed a significant rise in scarlet fever and iGAS infections [10]. This increase, which lasted until 2023, is currently linked to an early start to the GAS infection season and reduced immunity following two years of stringent social distancing measures during the COVID-19 pandemic [11]. Although *emm*1 variants thrived in multiple countries, investigations revealed that this surge was common to GAS strains of multiple *emm* types [11–13].

Fondazione IRCCS Policlinico San Matteo (San Matteo Hospital) is a 900-bed teaching hospital in Pavia, Italy. In the present work, we aimed to characterize our ten-year collection (2015–2024) of Bloodstream infection-causing GAS isolates using genomics, in light of the recent European epidemic events and of the availability of local genomes from a recent study focused on other hospitals in our geographic area [12].

2. Materials and methods

The GAS isolates described in this work were obtained from the blood cultures analyzed within the routine diagnostic pipeline at the San Matteo Hospital during the study period (January 2015 - July 2024); no selection for hospital ward or patient profile was operated. If multiple samples were available for the same patient, only the first one in chronological order was included. Blood cultures were grown into BD BACTEC™ blood culture media incubated in BD Bactec FX automated system (Becton Dickinson and Company, Franklin Lakes, New Jersey, United States). Positive cultures were streaked both on Columbia agar + 5 % sheep blood and incubated at 37 °C in aerobic condition, on Chocolate agar +PolyViteX and incubated at 37°C in 5 % CO_2, and on Schaedler agar + 5 % sheep blood and incubated at 37°C in anaerobic conditions (bioMérieux SA, Marcy-l'Etoile, France). Colonies identification was performed using MALDI-TOF mass spectrometry (Bruker Daltonics GmbH) with BioTyper version 3.0. Antimicrobial susceptibility was tested by Kirby Bauer disk diffusion susceptibility and concentration gradient diffusion (Liofilchem) assay on Mueller–Hinton Fastidious (Liofilchem) and interpreted following the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoints v13.0 (www.eucast.org). Isolates obtained from blood culture were stored at −80°C.

Genomic DNA was extracted using DNeasy blood and tissue kit (QIAGEN, Milan, Italy) following manufacturer instructions after a lysis step obtained through an incubation at 56°C overnight with lysozyme. Whole DNA sequencing was performed using the Illumina technology and the Nextera XT library preparation kit to obtain 2x250bp reads. Genomes were obtained using the Shovill pipeline (github.com/tseemann/shovill), which relies on Spades [14] for the assembly and includes a read trimming step with Trimmomatic [15]. Assemblies are available from the NCBI Genome database under BioProject ID PRJNA1170563. Sequencing reads from the study by Mangioni and colleagues [12] were assembled with the same method (ENA: PRJEB63359). Genomes were scanned for the presence of antimicrobial resistance genes using Abricate (github.com/tseemann/abricate) on the CARD database [16]. Virulence genes were searched using BLAST on an *ad-hoc* created reference database (including the following genes: *fbaA, fbp54, fctA, fctB, prtF2, srtC2, grab, hasA, hasB, hasC, hylA, lepA, lmb, sagA, scpA, sic, ska, slo, smeZ, speB, speC, speG, speH, speI, speJ, ideS, sda1, sdn, sfb1, sfbX, sla, sof, speA, speK, ssa, speL, speM, speQ, speR, spd, spd1, spd3, srtC1*). Alignments below 70 % identity and 70 % coverage were filtered out. Mutations in the genes of the *covRS* regulatory system were assessed by BLAST. Multilocus sequence types and *emm* types were obtained in silico using the programs mlst (github.com/tseemann/mlst) and emm-typer (github.com/MDU-PHL/emmtyper).

A species-wide phylogeny reconstruction was performed by comparing the genomes obtained in this study with the 100 most diverse publicly available genomes. In detail, all genomes available from the BV-BRC database (www.bv-brc.org, consulted in July 2024; N = 2180) were obtained. Assembly-dereplicator (github.com/rrwick/Assembly-Dereplicator) was then applied to select the 100 genomes with the most diverse genome content. The 34 novel genomes from the San Matteo Hospital were added to the dataset, together with the 27 of the Milan study [12]. Single-copy orthologous genes of the dataset were obtained using Panaroo, after calling ORFs with Prodigal. The aligned concatenate of the orthologs was used as input to IQtree [17] to infer the phylogeny using a Maximum Likelihood approach. To characterize putative local outbreaks, we used the P-DOR pipeline [18], selecting the 20 most similar genomes from the BV-BRC collection (see above) for each genome of interest. $M1_{UK}$ strains were identified using phylogeny and confirmed by identifying the 27 characterizing SNPs, by mapping reads to the genome of strain MGAS5005 (CP000017.2) using Snippy (github.com/tseemann/snippy).

3. Results and discussion

To describe the epidemiology of GAS at the San Matteo Hospital (Pavia, Italy), we monitored the monthly incidence of GAS isolations from January 2015 to July 2024 (Fig. 1A). As reported in multiple other studies (e.g. [11–13]), GAS cases followed a similar pattern over the years, showing the yearly maximum peak in spring and a minor peak in autumn. This pattern has been repeating consistently until 2020. Then, the establishment of social distancing measures to combat the COVID-19 pandemic stopped the spread of multiple other infectious diseases, including GAS infections [19]. In the last months of 2022 and in the following spring of 2023, we observed a steep rise in the number of cases, which reached values similar to those observed before the global pandemic. This result is in agreement with the large upsurge observed worldwide, which was associated with the presence of a large population of young children that had not been exposed to GAS in the pandemic period and had not developed any form of immunity. We must underline that we did not observe unprecedented incidence values as in other countries: in the UK the 2022–2023 peak reached values up to five times the average of the previous years [11]. On the other hand, our observation is concordant with the situation described in the Milan area (~30 Km from the study hospital) [12]. Strikingly, the highest peak value of GAS incidence (Fig. 1A) was recorded in 2016.

The number of bloodstream infections (BSI) in the ten-year period of the study is shown in Fig. 1B. The trend observed follows that of overall GAS cases, including a large increase in the post-pandemic period. This observation is also consistent with previous observations [11–13]. Interestingly, six BSI were registered in early 2016, a case number very close to that of the upsurge in 2023.

Of the total 45 bloodstream isolates of the study period, 34 were available (isolation dates and antimicrobial resistance profiles are reported in Supplemental File 1). The isolates belonged to 14 different *emm* types (Fig. 1C), *emm*1 (M-Cluster: A-C3) being the most represented (n = 11). MLST is almost completely concordant with *emm* typing. In fact, 10/11 *emm*1 genomes belonged to ST28, the

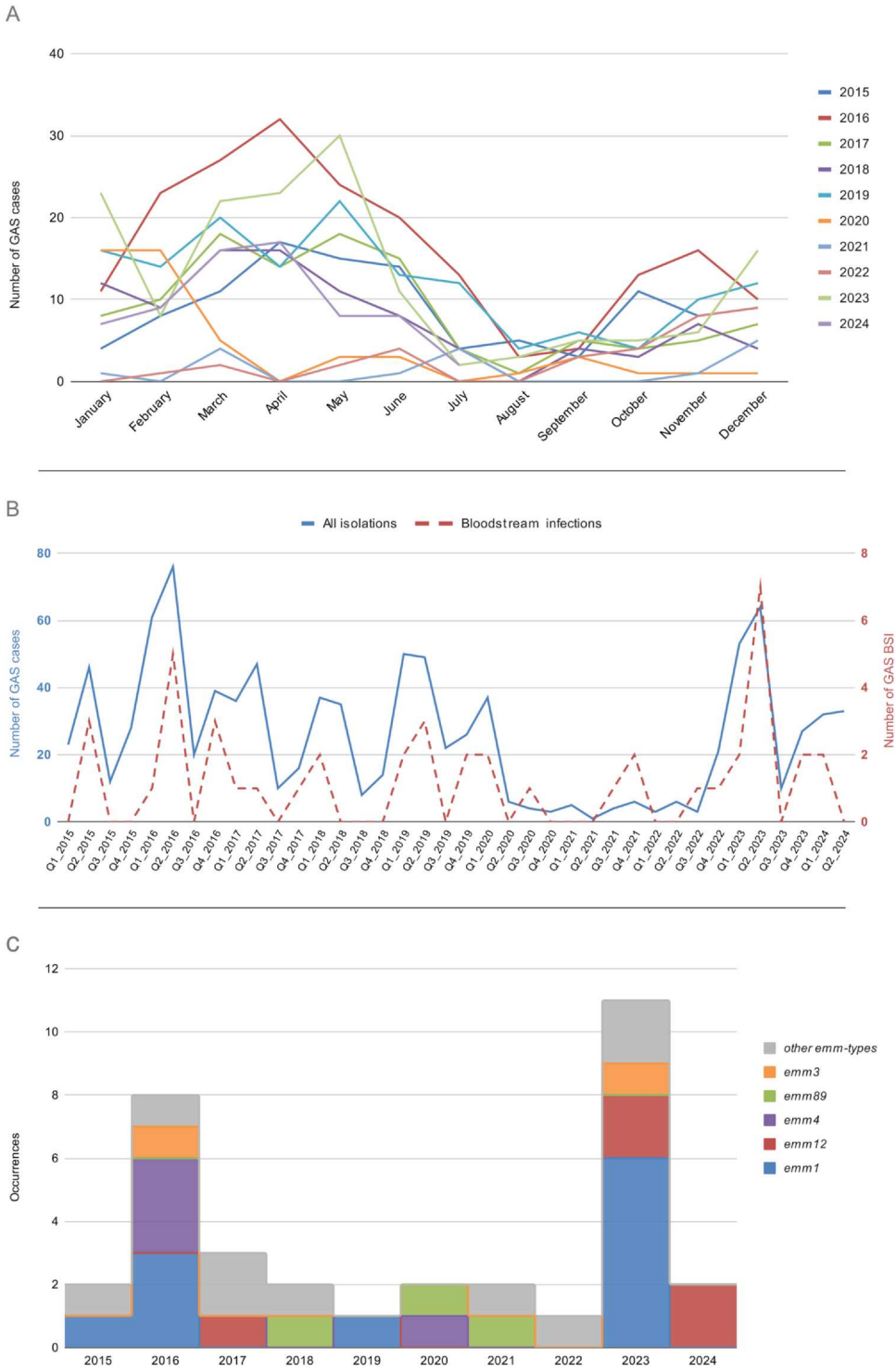


Fig. 1. A - Monthly isolations of GAS at the San Matteo Hospital from January 2015 to July 2024 (all isolation sources / infection types). **B -** GAS isolations at the San Matteo Hospital from January 2015 to July 2024 (left axis) and bloodstream infections (BSI; right axis) grouped in trimesters. **C -** Occurrences of emm-types over the years in the 34 bloodstream infection associated isolates characterized in this work using genomics.

remaining one being ST1319, which is part of the same clonal group as ST28. Full results of both typing methods are embedded in Fig. 2. The 2023 case peak is characterized mainly by *emm*1 isolates (N = 6, 55 %) and *emm*12 (N = 2, 18 %; Fig. 1C). Besides the worldwide dominance of *emm*1, high incidence of *emm12* was observed in many countries in the period 2022–2023. Indeed, in a study conducted in Spain [20], *emm12* predominated in 2022 in invasive infections and was surpassed by *emm1* in 2023, and in Portugal $M1_{UK}$ and *emm12* were the most isolated lineages in 2022–2023 [21]. On the other hand, the 2016 peak is characterized by high numbers of *emm*1 (N = 3, 38 %) and *emm*4 (N = 3, 38 %) isolations (Fig. 1C). The latters were recorded almost exclusively in 2016 and are absent from the 2023 upsurge in the study dataset.

Regarding the antibiotic susceptibility profile, all the strains of this study encoded genes *lmrP* and *mefE,* which encode two Major facilitator superfamily (MFS) antibiotic efflux pumps involved in

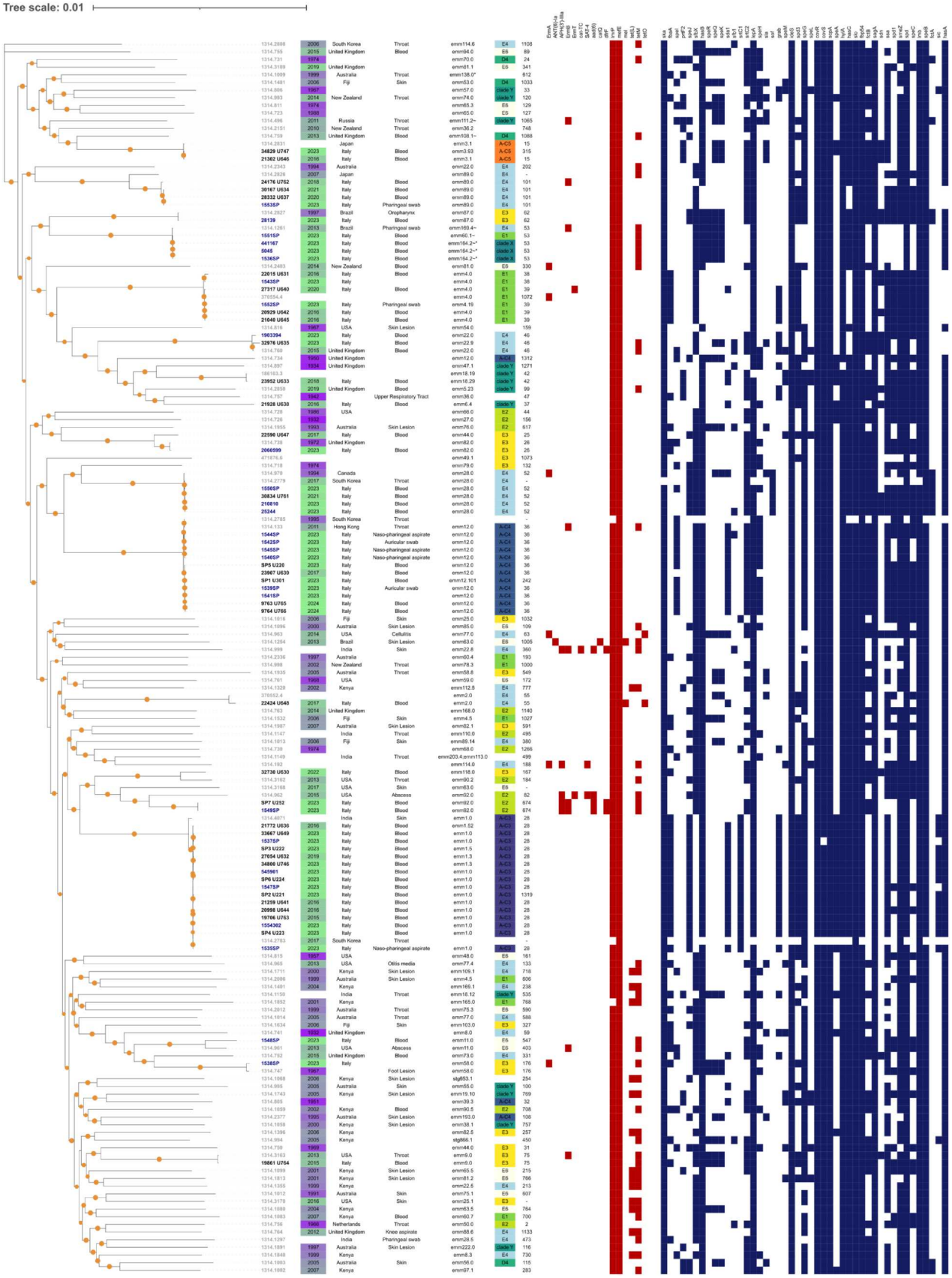


Fig. 2. Species-wide phylogeny of *Streptococcus pyogenes*, including genomes of the present study (in black), of the Milan study (in blue), and of background database genomes (in grey). Columns indicate (from left to right) year of isolation, country of isolation, source of isolation, *emm* type, *emm* Cluster, MLST, antimicrobial resistance gene presence (in red), virulence gene presence (in blue). Bootstrap values from 70 to 100 are indicated as orange dots.

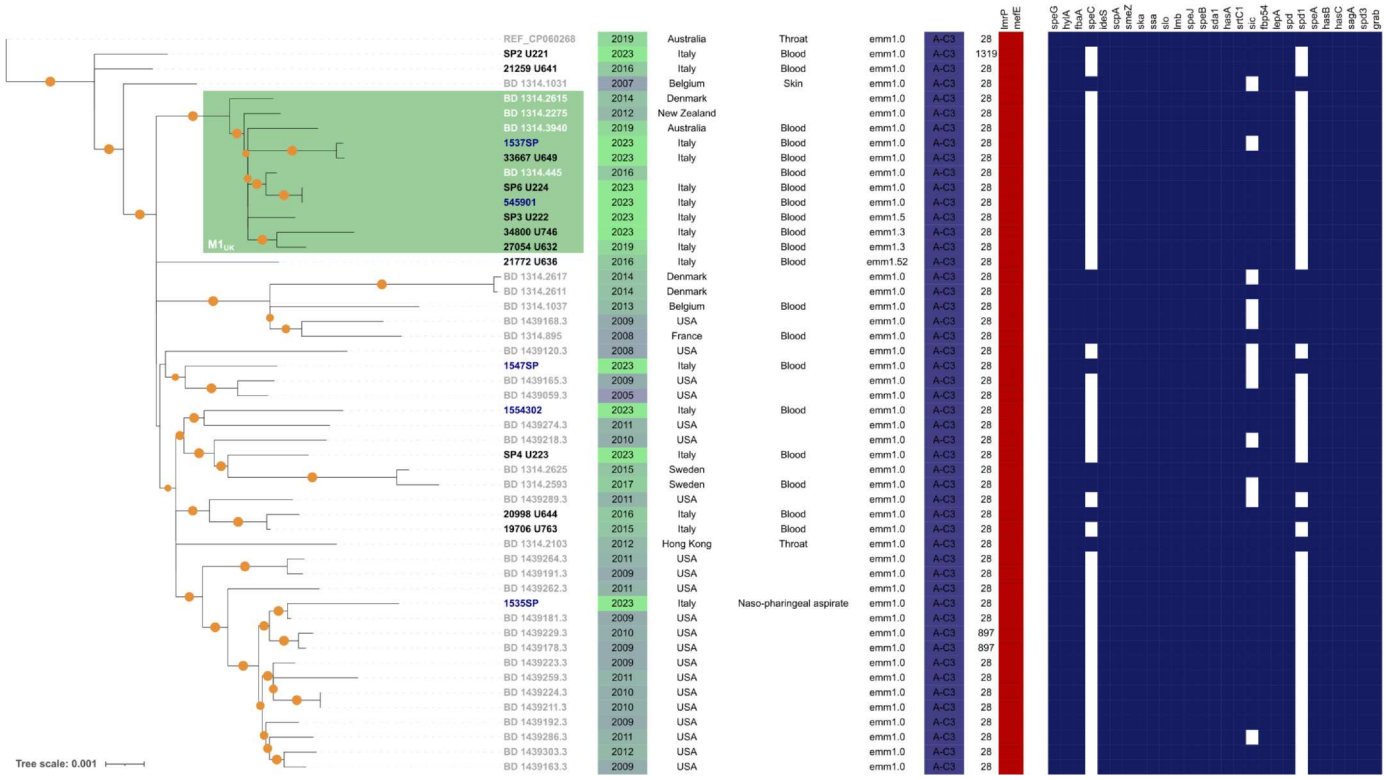


Fig. 3. CoreSNP phylogeny of the *emm*1 isolates of the present study (marked in black) and the Milan study (blue), with a background of 83 database genomes (in grey). Columns indicate (from left to right) year of isolation, country of isolation, source of isolation, *emm* type, *emm* Cluster, MLST, antimicrobial resistance gene presence (in red), virulence gene presence (in blue). The $M1_{UK}$ cluster is marked in green. Bootstrap values between 70 and 100 are indicated with orange dots.

macrolides resistance [22–24]. The presence of the *mefE* gene alone is not sufficient to grant resistance without the synergistic effect of *msrD* (or *mel*) gene [25]. The *mel* gene was only found in genome 22424_U648, which, however, retains susceptibility to macrolides. Five strains had also other resistance genes, such as *tetM* and *tetO* for tetracycline resistance (strains 19861_U764, 32976_U635 and 22424_U648), and *ermB*, and *ermT* for macrolides resistance (strains 24176_U762, 27317_U640). In concordance with their genotype, the two latter strains resulted to be resistant to macrolides with the phenotypic tests (Supplementary File 1). Only one strain (SP7_U252) possesses a higher number of resistance genes: *aph(3′)lla* and *aad(6)* genes for aminoglycoside resistance, *dfrF* gene for trimethoprim resistance, *ermB, and tetM.*

To put the 34 genomes of this study in the correct evolutionary context, a core-genome phylogeny was inferred using also 100 database genomes, representative of the global species-wide diversity, and the 27 genomes of the Milan study (Fig. 2). The *emm*1 cluster is characterized by high consistency in virulence and antimicrobial resistance gene content with one gene being present uniquely in this clade. In fact, the *sic* gene was detected in 16/18 of the *emm*1 genomes: in literature, the gene has been associated with the *emm*1 cluster and its protein product (protein SIC, Streptococcal Inhibitor of Complement) was linked to a diminished antimicrobial response of the host [26].

We further explored the evolutionary relationships using SNP-based phylogeny inferral on the most represented *emm* type clusters. First, we investigated the *emm*1 cluster. The topology of the resulting tree revealed the presence of multiple distinct subclades (Fig. 3), one of which was identified as $M1_{UK}$ (marked in green in Fig. 3). In fact, all genomes in the cluster present the distinctive 27 point mutations described in literature, including that inducing the overexpression of exotoxin-encoding gene *speA* [9]. Five genomes of this study are included within the $M1_{UK}$ cluster (5/11 *emm*1 genomes). Four $M1_{UK}$ isolates of the Pavia dataset were isolated within the 2023 upsurge, while the remaining one in 2019. Moreover, two of the 2023 ones are closely related to isolates from the Milan dataset. These results suggest the presence of a local transmission network of $M1_{UK}$ strains that has been ongoing since before the COVID-19 pandemic. Notably, none of the *emm*1 isolates carries any inactivating mutation in the *covRS* regulation system. In fact, only five isolates out of the 34 of the study present nonsense mutations in the *covS* gene (different among them): 21040_U645 (*emm*4, 2016), 22424_U648 (*emm*2, 2017), 23907_U639 (*emm*12, 2017), 27317_U640 (*emm*4, 2020), and 32730_U630 (*emm*118, 2022).

Similarly to what was observed in *emm*1, five more cases suggest the presence of endemic strains which persisted in the study area over the years (Supplemental File 2). In detail, *emm*4 strain 22015_U631, isolated in Pavia in 2016 shares the same monophylum with a strain of the Milan study (1543SP, SNP distance=277). *emm*12 isolates 9763_U765 and 9764_U766 (Pavia, 2024) are phylogenetically associated (SNP distance=8) and both closely related with Milan strain 1541SP (SNP distances=80; 84). *emm*28 isolate 30834_U761 is associated with Milan strains 210810 and 25244 (SNP distances=10; 22). Finally, *emm*89 strain 28332_U637 (2020) is associated with 1553SP (SNP distance=21), and *emm*92 isolate SP7_U252 with 1549SP (SNP distance=1). Moreover, the two *emm*92 genomes shared the same multi-resistant genotype (Fig. 2). This observation reveals the presence of a highly antimicrobial resistant local clone, which should be further investigated and monitored in surveillance initiatives.

4. Limitations

This study is partially limited by sample numerosity. The epidemiological investigation was focused on bloodstream infections,

which usually have a low incidence. On the other hand, the strength of the study resides in the wide timespan it covers.

5. Conclusions

We conducted a retrospective analysis of GAS infections in a ten-year period. Despite observing incidence patterns similar to other reports worldwide, we also reported a peak in 2016 (observed also in [12]), which was similar to that of the 2022/23 global upsurge and was characterized by a high incidence of invasive cases. One might hypothesize that the 2016 peak could be related to the pre-pandemic GAS upsurge in the UK, which has been linked to the emergence of the $M1_{UK}$ genotype [8]. However, our observations question this hypothesis. Indeed, *emm*1 strains were highly prevalent among the 2016 isolates of our study (Fig. 1C), but none of them was part of the $M1_{UK}$ clade (Fig. 3), which contained mainly genomes of 2022/23. Notably, the 2016 peak also featured a strong presence of *emm*4 strains, which did not have enhanced virulence or antimicrobial resistance profiles.

Among the *emm* types that we observed (Fig. 1C), the most prevalent was *emm*1, including the widespread $M1_{UK}$ subtype. None of the genomes of the *emm*1 strains showed any inactivating mutation in the *covRS* regulation system, which was reported to remove negative regulation in the production of exotoxins in highly virulent strains [7].

Most importantly, we observed multiple cases of possible local transmission clusters, which highlight the persistence of GAS in the general population, also across periods of confinement and highly restricted social interactions such as the very stringent measures introduced by the Italian government during the first waves of the COVID-19 pandemic. Among these, we detected the local transmission of a multidrug resistant strain of *emm*92, which should be monitored in future surveillance efforts.

Ethical statement

This study was designed and conducted in accordance with the Helsinki Declaration and was approved by the Ethics Committee of Policlinico San Matteo in Pavia (internal project code: 2023–3.11/105).

CRediT authorship contribution statement

MC, SG, FB and PC conceptualized the project. AK, IM, CR, VB, and CM conducted *in vitro* analyses. GP and SG conducted *in silico* analyses. MC, CM and SG interpreted results. MC, CM and SG wrote the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank the staff members of the Microbiology and Virology Unit at Fondazione IRCCS Policlinico San Matteo, who contributed to the collection of all isolates over the years. This work was supported by the Ricerca Corrente project 1004-rcr2022–43 (Gaiarsa), by the 5 × 1000 fundings by the Italian Ministry of Health (Fondazione IRCCS Policlinico San Matteo; Gaiarsa), and by EU funding within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jiph.2025.102701.

References

[1] Brouwer S, Rivera-Hernandez T, Curren BF, Harbison-Price N, De Oliveira DMP, Jespersen MG, et al. Pathogenesis, epidemiology and control of Group A Streptococcus infection. Nat Rev Microbiol 2023;21:431–47. https://doi.org/10.1038/s41579-023-00865-7
[2] Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus Pyogenes: Basic Biology to Clinical Manifestations. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016.
[3] Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, et al. Disease manifestations and pathogenic mechanisms of Group A *Streptococcus*. Clin Microbiol Rev 2014;27:264–301. https://doi.org/10.1128/CMR.00101-13
[4] Ghosh P. The nonideal coiled coil of M protein and its multifarious functions in pathogenesis. Adv Exp Med Biol 2011;715:197–211. https://doi.org/10.1007/978-94-007-0940-9_12
[5] Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. Lancet Infect Dis 2009;9:611–6. https://doi.org/10.1016/S1473-3099(09)70178-1
[6] Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long SW, et al. Evolutionary pathway to increased virulence and epidemic group A Streptococcus disease derived from 3,615 genome sequences. Proc Natl Acad Sci USA 2014;111:E1768–76. https://doi.org/10.1073/pnas.1403138111
[7] Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog 2006;2:e5. https://doi.org/10.1371/journal.ppat.0020005
[8] Lynskey NN, Jauneikaite E, Li HK, Zhi X, Turner CE, Mosavie M, et al. Emergence of dominant toxigenic M1T1 Streptococcus pyogenes clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. Lancet Infect Dis 2019;19:1209–18. https://doi.org/10.1016/S1473-3099(19)30446-3
[9] Davies MR, Keller N, Brouwer S, Jespersen MG, Cork AJ, Hayes AJ, et al. Detection of Streptococcus pyogenes M1UK in Australia and characterization of the mutation driving enhanced expression of superantigen SpeA. Nat Commun 2023;14:1051. https://doi.org/10.1038/s41467-023-36717-4
[10] W H O. Increased Incidence of Scarlet Fever and Invasive Group A Streptococcus Infection—Multi-Country. WHO; 2022.
[11] Vieira A, Wan Y, Ryan Y, Li HK, Guy RL, Papangeli M, et al. Rapid expansion and international spread of M1UK in the post-pandemic UK upsurge of Streptococcus pyogenes. Nat Commun 2024;15:3916. https://doi.org/10.1038/s41467-024-47929-7
[12] Mangioni D, Fox V, Saltini P, Lombardi A, Bussini L, Carella F, et al. Increase in invasive group A streptococcal infections in Milan, Italy: a genomic and clinical characterization. Front Microbiol 2023;14:1287522. https://doi.org/10.3389/fmicb.2023.1287522
[13] Wolters M, Berinson B, Degel-Brossmann N, Hoffmann A, Bluszis R, Aepfelbacher M, et al. Population of invasive group A streptococci isolates from a German tertiary care center is dominated by the hypertoxigenic virulent M1UK genotype. Infection 2024;52:667–71. https://doi.org/10.1007/s15010-023-02137-1
[14] Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes de novo assembler. Curr Protoc Bioinforma 2020;70:e102. https://doi.org/10.1002/cpbi.102
[15] Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20. https://doi.org/10.1093/bioinformatics/btu170
[16] McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 2013;57:3348–57. https://doi.org/10.1128/AAC.00419-13
[17] Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32:268–74. https://doi.org/10.1093/molbev/msu300
[18] Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics 2023;39. https://doi.org/10.1093/bioinformatics/btad571
[19] Kajihara T, Yahara K, Kamigaki T, Hirabayashi A, Hosaka Y, Kitamura N, et al. Effects of coronavirus disease 2019 on the spread of respiratory-transmitted human-to-human bacteria. J Infect 2024;89:106201. https://doi.org/10.1016/j.jinf.2024.106201
[20] Ramírez de Arellano E, Saavedra-Lozano J, Villalón P, Jové-Blanco A, Grandioso D, Sotelo J, et al. Clinical, microbiological, and molecular characterization of pediatric invasive infections by Streptococcus pyogenes in Spain in a context of global outbreak. MSphere 2024;9:e0072923. https://doi.org/10.1128/msphere.00729-23
[21] Gouveia C, Bajanca-Lavado MP, Mamede R, Araújo Carvalho A, Rodrigues F, Melo-Cristino J, et al. Sustained increase of paediatric invasive Streptococcus pyogenes infections dominated by M1UK and diverse emm12 isolates, Portugal,

September 2022 to May 2023. Eur Surveill 2023;28. https://doi.org/10.2807/1560-7917.ES.2023.28.36.2300427

[22] Putman M, van Veen HW, Degener JE, Konings WN. The lactococcal secondary multidrug transporter LmrP confers resistance to lincosamides, macrolides, streptogramins and tetracyclines. Microbiology 2001;147:2873–80. https://doi.org/10.1099/00221287-147-10-2873

[23] Martini CL, Silva DNS, Viana AS, Planet PJ, Figueiredo AMS, Ferreira-Carvalho BT. Streptococcus pyogenes lineage ST62/emm87: The International Spread of This Potentially Invasive Lineage. Antibiotics 2023;12. https://doi.org/10.3390/antibiotics12101530

[24] Iannelli F, Santoro F, Santagati M, Docquier J-D, Lazzeri E, Pastore G, et al. Type M resistance to macrolides is due to a two-gene efflux transport system of the ATP-binding cassette (ABC) superfamily. Front Microbiol 2018;9:1670. https://doi.org/10.3389/fmicb.2018.01670

[25] Tatsuno I, Isaka M, Masuno K, Hata N, Matsumoto M, Hasegawa T. Functional Predominance of msr(D), which is more effective as mef(A)-associated than mef (*E*)-associated, over mef(A)/mef(E) in macrolide resistance in streptococcus pyogenes. Micro Drug Resist 2018;24:1089–97. https://doi.org/10.1089/mdr.2017.0277

[26] Westman J, Chakrakodi B, Snäll J, Mörgelin M, Bruun Madsen M, Hyldegaard O, et al. Protein SIC secreted from streptococcus pyogenes forms complexes with extracellular histones that boost cytokine production. Front Immunol 2018;9:236. https://doi.org/10.3389/fimmu.2018.00236

MICROBIAL DRUG RESISTANCE
Volume 30, Number 1, 2024
© Mary Ann Liebert, Inc.
DOI: 10.1089/mdr.2023.0054

Open camera or QR reader and scan code to access this article and other resources online.

Concomitant Resistance to Cefiderocol and Ceftazidime/Avibactam in Two Carbapenemase-Producing *Klebsiella pneumoniae* Isolates from Two Lung Transplant Patients

Greta Bellinzona,[1] Cristina Merla,[2] Marta Corbella,[2] Elizabeth Nagy Iskandar,[2,3] Elena Seminari,[4] Angela Di Matteo,[4] Stefano Gaiarsa,[2] Greta Petazzoni,[2] Davide Sassera,[1] Fausto Baldanti,[2,5] Aurora Piazza,[6] and Patrizia Cambieri[2]

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

In this study, we present two cases of *Klebsiella pneumoniae*, one KPC-33- and one NDM-1-producing, showing resistance to cefiderocol and ceftazidime/avibactam, collected in the intensive care unit of a hospital in Northern Italy from two patients who had recently undergone lung transplantation. Whole-genome sequencing was performed to investigate the molecular features of these strains.

Keywords: cefiderocol, cefiderocol resistance, ceftazidime/avibactam, ceftazidime/avibactam resistance, *Klebsiella pneumoniae*, co-resistance, virulence factors, KPC mutated, *bla*KPC-33, *bla*NDM-1

Cefiderocol (CFDC) is a novel synthetic conjugate siderophore cephalosporin that can exploit the bacterial iron transport system to enter the periplasm. It shows good activity on Gram-negative bacteria, mainly those belonging to *Enterobacteriaceae*, and displays great stability against degradation by metallo-β-lactamases.[1,2]

Despite their recent introduction into clinical practice, reduced susceptibility to CFDC, both *in vitro* and *in vivo*, has already been reported on multiple occasions in carbapenemase-producing *Enterobacteriaceae*.[3–7] Nonetheless, this phenomenon is still considered rare, and the underlying mechanisms have not been clarified.[6] The new combination of ceftazidime/avibactam (AVI/CZA) was recently commercialized representing another solution against highly drug-resistant bacteria.[8]

Likely, the concomitant resistance between AVI/CZA and CFDC is very rare in *Klebsiella pneumoniae*; however, this worrying phenomenon was previously described.[2]

The objective of this study was to investigate by whole-genome sequencing the molecular features of two *K. pneumoniae* clinical isolates showing reduced susceptibility to both AVI/CZA and CFDC.

Case Description

Case 1: isolate 5844

A 61-year-old male patient was admitted to the cardiopulmonary intensive care unit (ICU) of the IRCCS Policlinico San Matteo Hospital, Pavia, Italy, on January 29, 2021 (day 1) for a single lung transplant (LTx). His medical history showed an anti-synthetase syndrome associated with progressive interstitial lung disease for which he underwent several ICU admissions before his LTx. The patient was placed on mechanical ventilation upon admission and, following LTx, prophylactic antibiotic treatment was initiated with piperacillin/tazobactam and vancomycin. Nevertheless,

Departments of [1]Biology and Biotechnology and [6]Clinical, Surgical, Diagnostic, and Paediatric Sciences, University of Pavia, Pavia, Italy.
[2]Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.
[3]Specialization School of Microbiology and Virology, Microbiology and Virology Department, University of Pavia, Pavia, Italy.
[4]Division of Infectious Diseases I, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.
[5]Molecular Virology Unit, Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.

throughout the period between his admission till day 41, the patient has experienced repeated episodes of low-grade fever and frequent alterations in the inflammatory markers with negative blood culture results. In addition, several bronchial aspirations were performed and sampled during that period resulting positive for New Delhi metallo-β-lactamase (NDM)–producing *K. pneumoniae* (NDM-Kp).

Upon clinical stabilization, the patient was discharged to a rehabilitation center on day 80. Ten days later, he was readmitted to our institute and a bronchoalveolar lavage (BAL) sample was collected. NDM-Kp, isolate 5844, was isolated from the BAL and demonstrated resistance to CFDC, AVI/CZA, and aztreonam (ATM) in the antimicrobial susceptibility testing report. However, AVI/CZA and ATM showed a synergic effect with 0.375 fractional inhibitory concentration (FIC) index. As a result, CFDC was suspended and treatment with AVI/CZA and ATM was initiated on day 93. In addition, blood cultures collected on day 133 during a septic episode were reported positive for NDM-Kp. Owing to further worsening of his clinical conditions and inflammatory state, treatment with continuous venovenous hemodiafiltration was started. In the following days, NDM-Kp was repeatedly isolated from rectal swabs and respiratory samples. The patient died after 177 days from his initial ICU admission.

Case 2: isolate 220

A 57-year-old male patient underwent a recent bilateral LTx in January 2021 following a severe COVID-related acute respiratory distress syndrome. He was admitted to the ICU on March 19, 2021 (day 1) with the purpose of stabilizing his cardiorespiratory status, resolving a post-transplant macrophage activation syndrome and tracheostomy and nasogastric (NG) tube weaning. Upon admission (day 1), a rectal swab sample was sent to the microbiology laboratory for screening purposes. On day 1, the patient had several episodes of bronchial secretions that were sent to the microbiology laboratory as a bronchial aspirate sample (BAS). The BAS resulted positive for *Candida utilis* and *Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae* (KPC-kp), isolate 220, which resulted resistant to CFDC and showed a synergic effect between AVI/CZA and ATM (FIC index: 0.250). Conversely, the KPC-kp isolated from the rectal swab sample was susceptible to CFDC (isolate 214).

On day 5, the patient was transferred to the department of infectious diseases and immediately started on inhaled liposomal amphotericin B, AVI/CZA, and ATM. Inflammatory markers started to drop; however, the patient started showing signs of liver fibrosis on echography examination performed on day 11, which was compatible with a recent severe COVID infection.

The patient's clinical conditions improved and antibiotic treatment with AVI/CZA and ATM was suspended on day 21. Between day 20 and day 41, the patient received 2 units of total blood transfusion following a drop in hemoglobin levels (Hb = 7.7 g/dL). On day 41, the patient experienced a febrile episode with significant drop in blood pressure and, as a result, blood culture bottles were collected. KPC-kp was isolated from blood cultures the following day and he was started on meropenem/vaborbactam (MEV). During his hospital stay, he developed macrophage activation syndrome and acute renal failure for which he received appropriate treatment. On day 48 the patient was started on ganciclovir owing to a cytomegalovirus reactivation (3,420 DNA copies/mL), which was later suspended because of severe pancytopenia. The patient's clinical conditions gradually declined, he later became cachectic and severely dysphagic despite the NG tube feeding; therefore, a percutaneous endoscopic gastrostomy was placed. He died on June 7, 2021.

Microbiological Investigations

BAL and BAS samples for both patients were plated by automatic streaking WASP® (Copan, Brescia, Italy) on nutritive and selective media following internal laboratory procedures. Identification of the bacterial colonies at species level was performed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Bruker Daltonics, Bremen, Germany) using the software Bruker Biotyper 3.1. Susceptibility was determined for all the strains using the NMIC-402 panel of the Phoenix M50 BD automated system (Becton Dickinson, Franklin Lakes, NJ) and Sensititre™ Gram-Negative EUMDROXF Plate (ThermoFisher Scientific, Rodano, Italy). Resistance to CFDC was confirmed using disk diffusion with CFDC 30 μg (Liofilchem, Roseto degli Abruzzi, Italy) and E-test (Liofilchem, Roseto degli Abruzzi, Italy). Susceptibility results were interpreted according to EUCAST clinical breakpoint (version 11.0).

Genomic DNA was extracted for both isolates with blood and tissue kit (Qiagen, Düsseldorf, Germany) following the manufacturer's instructions for short-read sequencing and with MagAttract HMW DNA Kit (Qiagen) following the manufacturer's instructions for long-read sequencing. Short reads were obtained on an Illumina Miseq platform with a 2×250 paired-end run, after a Nextera XT library preparation step (Illumina Inc., San Diego, CA), whereas long reads were obtained on MinION following manufacturer's recommendations.

Paired-end Illumina and Nanopore reads were subjected to quality control checks before downstream analysis using FastQC v.0.11.9 and NanoPlot v.1.20.0,[9] respectively. Long-reads sequence data were combined with short-reads Illumina data to generate hybrid assemblies with Unicycler v.0.5.0[10] in bold mode (-m bold), to obtain complete sequences.

Generated assembly files for each isolate were visually inspected to identify plasmids by constructing an assembly graph in Bandage[11] and checking for circularity and higher depth coverage. The Plasmid Database (PLSDB) web server[12] was used for comparing our plasmids with the PLSDB database, and mash distances were registered to estimate sequence distances.

Multilocus sequence typing, virulence loci, capsule (K) and lipopolysaccharide (O) loci were identified using Kleborate v.2.0.1.[13] Additional in-depth investigations about the presence of antimicrobial resistance genes, virulence genes and plasmids were conducted using ABRicate v.1.0.1 using Comprehensive Antibiotic Resistance Database (CARD) database,[14] VirulenceFinder Database[15,16] and PlasmidFinder,[15] respectively. The presence of mobile genetic elements has been checked using

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

MGEFinder[17] and a complete genome annotation has been obtained using prokka v.1.11.[18] Finally, the presence of already known CFDC resistance determinants (mutations/deletions) have been checked using Basic Local Alignment Search Tool.

Results

Case 1: isolate 5844

The 5844 isolate resulted multidrug-resistant (MDR) being resistant to all the antibiotics tested including CFDC, AVI/CZA and colistin with the exception of trimethoprim/sulfamethoxazole, and amikacin, which retained their activity (Table 1). The results of the three methods used (disk diffusion, E-test and broth dilution) for CFDC testing were concordant and categorized the strain as resistant (Table 2).

The whole-genome sequence (WGS) analyses showed that the complete genome of the 5844 isolate is composed of a 5.37 Mb chromosome and five plasmids, that is, p5844_1 (112,869-bp), p5844_2 (54,593-bp), p5844_3 (54,064-bp), p5844_4 (39,711-bp) and, p5844_5 (4,167-bp). This strain was found to belong to ST147, an emerging high-risk clone. KL64 and O2v1 serotype were detected based on capsule and lipopolysaccharide synthesis loci, respectively.

The predicted resistance profile generated with ABRicate using CARD database[14] is given in Fig. 1. Carbapenem resistance could be associated with the presence of the *bla*NDM-1 gene encoded on the p5844_3 plasmid.

In addition to the *bla*NDM-1 gene, several further resistance determinants have been detected on both the chromosome and the plasmids, including other β-lactam resistance genes (*bla*CTX-M-15, *bla*OXA-1, *bla*OXA-9, *lptD*, *ompK37*), genes for quinolones (*qnrS1*), aminoglycosides (*aph(3')-VI*), and sulfonamide (*sul1*) resistance.

Of interest, the 5844 isolate also showed a large number of virulence factors, comprising *fim* and *mrk* loci encoding type 1 and type 3 fimbriae (*fimAICDFGHK, fimE, fimB, mrkABCD*), and the yersiniabactin cluster (*ybtAEPQSTUX*).

Mash searches were performed against the PLSDB database[12] to explore whether similar plasmids have already been described in the literature (Table 3). Of interest, p5844_1 was found to be identical to NZ_CP074089.1 plasmid and p5844_3 identical to NZ_CP074090.1 plasmid, both recovered from an NDM-Kp inside a protracted outbreak started in Tuscany (Italy) in 2018 and continued until 2021.[19]

A search for mutations or deletions in genes encoding porins (*ompK35, ompK36, ompK37*), siderophore receptors (*cirA, fhuA, fepA*), and penicillin-binding protein (PBP) previously associated to CFDC resistance *in vitro* and *in vivo* yielded no results.[20]

Case 2: isolate 220

The 220 isolate showed an MDR phenotype, being resistant to all the antibiotics tested including CFDC, AVI/CZA and colistin with the exception of MEV, trimethoprim/sulfamethoxazole and gentamicin, which retained their activity (Table 1).

The disk diffusion and broth dilution methods for CFDC testing were concordant and categorized the strain

Table 1. Antibiotic Susceptibility Profiles of the Strains 5844 (Case 1) and 220 (Case 2)

Strain	*AMC*	*AMP*	*p/T*	*FEP*	*FOT*	*CZA*	*AVI/CZA*	*C/T*	*COL*	*CIP*	*LVX*	*ERT*	*IMP*	*MEM*	*MEV*	*AK*	*CN*	*ATM*	*SXT*
	S≤8	*S*≤8	*S*≤8	*S*≤1	*S*≤2	*S*≤1	*S*≤8	*S*≤2	*S*≤2	*S*≤0.25	*S*≤0.5	*S*≤0.5	*S*≤2	*S*≤2	*S*≤8	*S*≤8	*S*≤2	*S*≤1	*S*≤2
	R>8	*R*>8	*R*>8	*R*>4	*R*>2	*R*>4	*R*>8	*R*>2	*R*>2	*R*>0.5	*R*>1	*R*>0.5	*R*>4	*R*>8	*R*>8	*R*>8	*R*>2	*R*>4	*R*>4
5844	>32/2 (R)	>8 (R)	>32/4 (R)	>8 (R)	>4 (R)	>8 (R)	>16/4 (R)	>8/4 (R)	>16 (R)	>1 (R)	>2 (R)	>1 (R)	>8 (R)	>16 (R)	>256 (R)	16 (I)	>4 (R)	>32 (R)	≤1/19 (S)
220	>32/2 (R)	>8 (R)	>32/4 (R)	>8 (R)	>4 (R)	>8 (R)	>16/4 (R)	>8/4 (R)	>16 (R)	>1 (R)	>2 (R)	>1 (R)	>8 (R)	>16 (R)	1/8 (S)	>32(R)	2 (S)	>32 (R)	≤1/19 (S)

Interpretation according to EUCAST breakpoints (version 11, 2021) are given in brackets.
AK, amikacin; AMC, amoxicillin–clavulanate; AMP, ampicillin; ATM, aztreonam; AVI/CZA, ceftazidime/avibactam; C/T, ceftolozane/tazobactam; CIP, ciprofloxacin; CN, gentamycin; COL, colistin; CZA, ceftazidime; ETP, Ertapenem; FEP, cefepime; FOT, cefotaxime; IMP, imipenem; LVX, levofloxacin; MEM, meropenem; MEV, meropenem/vaborbactam; P/T, piperacillin/tazobactam; SXT, trimethoprim/sulfamethoxazole.

Table 2. Susceptibility to Cefiderocol Tested with Disk Diffusion, E-Test and Broth Dilution Methods

Strain	Sample	*Disk diffusion with CFDC 30 μg (Liofilchem) (diameter in mm) S ≥ 22 mm; R < 22 mm*	*E-TEST (Liofilchem) S ≤ 2; R > 2*	*Broth microdilution (Sensititre Gram negative EUMDROXF Plate) S ≤ 2; R > 2*
5844	BAL	8 mm (R)	32 (R)	>8 (R)
220	BAS	19 mm (R)	1 (S)	4 (R)

Susceptibility to CFDC was interpreted based on EUCAST breakpoints (version 11, 2021).
BAL, bronchoalveolar lavage; BAS, bronchial aspirate sample; CFDC, cefiderocol.

as resistant, whereas the E-test reported the isolate as susceptible to CFDC, with an minimum inhibitory concentration (MIC) value of 1 μg/mL (Table 2).

According to WGS results, the isolate harbored a 5.32 Mb chromosome and three plasmids, that is, p220_1 (112,610 bp), p220_2 (10,122 bp), and p220_3 (5,598 bp). Isolate 220 was found to belong to ST101 and KL17 and O1v1 serotype based on capsule and lipopolysaccharide synthesis loci. Colistin resistance could be associated with mutations in *mrgB* according to the Kleborate result.

The predicted resistance profile interestingly includes AVI/CZA resistance (Fig. 1). This resistance may be genomically associated with the presence of *bla*KPC-33 gene, which belongs to the KPC-3 variants group, known to confer resistance to AVI/CZA.[20] In the present isolate, the *bla*KPC-33 gene is present in two copies, one located on the chromosome and another on an IncR_1/IncFIA(HI1)_1/IncFII(pKP91)_1 plasmid (p220_1).

Besides *bla*KPC-33, several resistance determinants have been identified: the chromosome-encoded *oqxA* (conferring

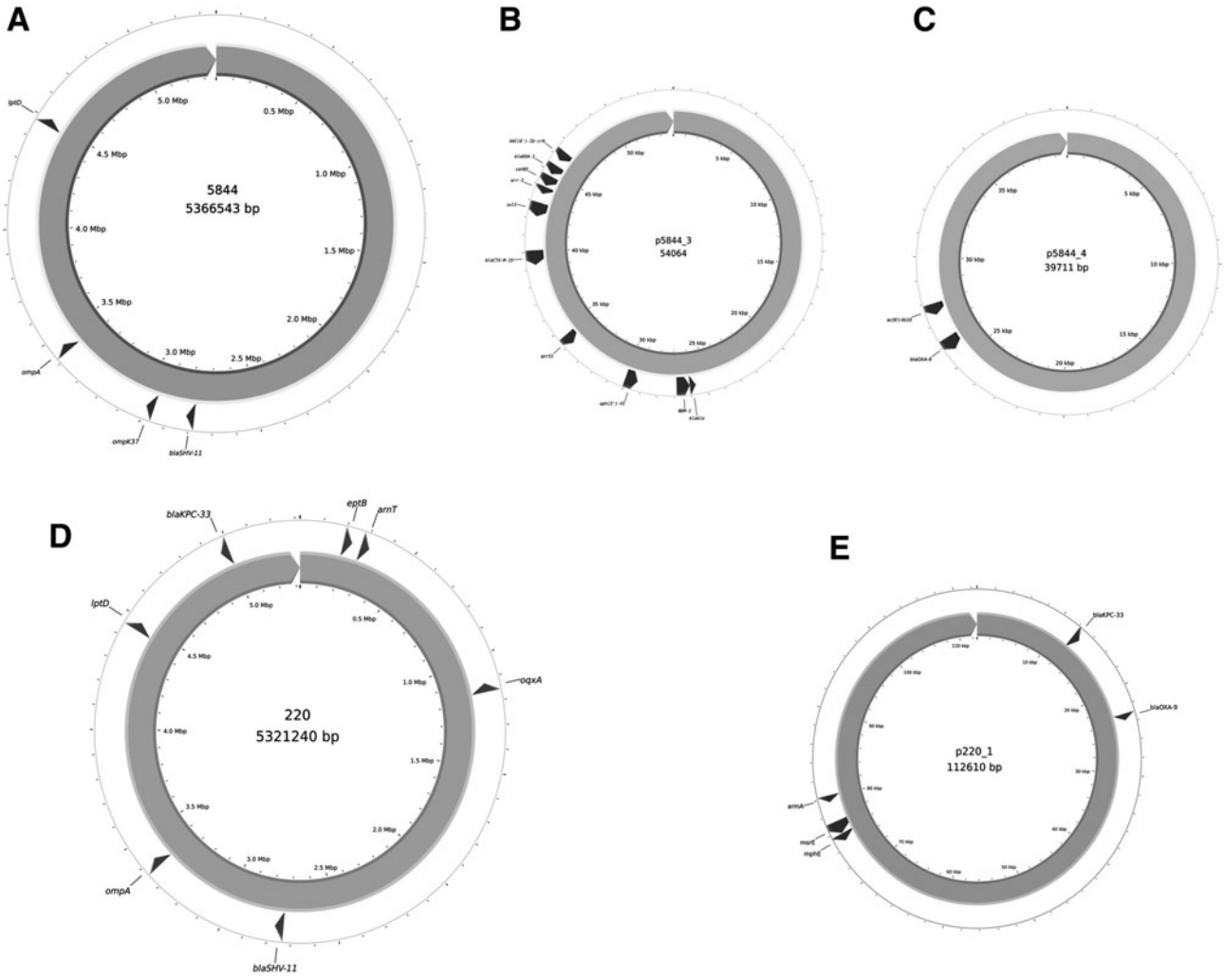


FIG. 1. Graphical representation of the chromosomes and plasmids of main interest for the two *Klebsiella pneumoniae* isolates, showing genes of antibiotic resistance indicated by *arrowheads*. **(A)** Chromosome of 5844; **(B)** p5844_3; **(C)** 5844_4; **(D)** Chromosome of 220; **(E)** p220_3.

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

Table 3. Mash Search Performed Against the PLSDB Database (12) to Isolates 5844 and 220 Plasmids

		PLSDB best match		
	Plasmid size (bp)	*Plasmid accession*	*Mash distance*	p
p5844_1	112,869	NZ_CP074089.1	0	0
p5844_2	54,593	NZ_CP076011.1	0.000191626	0
p5844_3	54,064	NZ_CP074090.1	0	0
p5844_4	39,711	NZ_CP021948.1	0.000981871	0
p5844_5	4,167	NZ_CP069949.1	0.000167546	0
p220_1	112,610	NZ_CP033628.1	0.000508006	0
p220_2	10,122	NZ_CP070526.1	0.0000476906	0
p220_3	5,598	NZ_CP070601.1	0.0000955247	0

resistance to fluoroquinolones), *ompA* and *lptD* (conferring resistance to β-lactams), and the plasmid-borne *mphE* (macrolide resistance), *msrE* (erythromycin resistance), and *armA* (aminoglycoside resistance).

The virulence factors content of the 220 isolate included *fimAICDFGHK, fimE,* and *fimB* cluster for type 1 fimbriae production, *mrkABCD* and *mrkFJIH* loci encoding type 3 fimbriae, *fepBDGC* cluster for iron transport, and the yersiniabactin cluster (*ybtAEPQSTUX*).

Each of the three plasmids was searched against the PLSDB database and the results are given in Table 3.

A similar case of *K. pneumoniae* harboring a KPC-3 variant enzyme, resistant to both CFDC and AVI/CZA, was isolated from an Italian hospital in June 2020.[5] However, it is unlikely that this variant could explain CFDC resistance, as another study showed that of 51 KPC-3-producing *K. pneumoniae* strains, all were susceptible to CFDC.[21] Further investigations about the resistance mechanism to CFDC in KPC-producing *K. pneumoniae* are needed.

As for isolate 5844, none of the mutations or deletions in genes already linked to CFDC resistance have been detected.

Discussion

In this study we report two cases of resistance to CFDC and AVI/CZA in two patients who had recently undergone a lung transplant. Genomics detected a KPC-33 and NDM-1 enzyme in the *K. pneumoniae* isolates, respectively. The 5844 isolate, *bla*NDM-1 positive, showed a higher CFDC MIC value (or a larger inhibition zone in case of disk diffusion test) in comparison with the 220 isolate, *bla*KPC-33 positive. It is known that NDM facilitates the emergence of high level of CFDC resistance in *Enterobacterales*,[4] and KPC-3 variants have been demonstrated *in vitro* to have a possible impact on CFDC resistance. However, the effect of β-lactamases has been shown to be insufficient in causing CFDC resistance, which is more likely to be the result of a range of mechanisms including mutations altering porins function, efflux pumps, and PBPs, generally operating in concert.[20] None of these recognized mechanisms contributing to CFDC resistance were found in the current investigation, raising the possibility of the presence of a still unidentified resistance determinant.

Our study highlights some concerns. First, the two *K. pneumoniae* isolates showed an MDR phenotype, harboring several resistance genes. The hyper-resistant phenotype of these strains poses a serious therapeutic problem because they remained fully susceptible only to few antibiotics. The concomitant resistance to the recently introduced AVI/CZA and CFDC represents a major concern. A second important point is that these isolates, in addition to the MDR phenotype, exhibit a hypervirulent phenotype, which includes the *fim* and *mrk* loci encoding type 1 and type 3 fimbriae, and the yersiniabactin cluster. The latter is harbored on the integrative conjugative elements (ICEs) ICEKp1-like, ICEKpnHS11286-1, known to strongly influence the pathogenicity of *K. pneumoniae* strains. Finally, the occurrence of resistance to AVI/CZA and CFDC simultaneously in two high-risk clones as *K. pneumoniae* ST147 and 101, in addition to being isolated in the ICU, points out a very serious concern.

Conclusion

Although we could not identify a new mechanism underlying the CFDC resistance observed in the two *K. pneumoniae* isolates, the discovery of CFDC-resistant strains in patients not previously treated with CFDC is clinically significant given the auspicious expectations of this drug in treating severe infections caused by carbepenemase-producing enterobacteria. Moreover, the convergence of such broad resistome and virulome in the ST147 and ST101 *K. pneumoniae* high-risk clones can represent a very critical issue.

Authors' Contributions

Conceptualization: P.C. and D.S.; E.S., A.D.M., C.M., and M.C. gathered the clinical specimens, strains, and clinical data; Microbiological analyses have been performed by C.M., M.C., E.N.I., and P.C.; Bioinformatic analyses by G.B., S.G., and G.P.; All the authors contributed to the interpretation of the data; writing—original draft preparation, G.B. and A.P.; writing—review and editing, M.C., E.N.I., P.C., A.P., and D.S.; supervision, F.B., D.S., and P.C. All authors have read, reviewed the draft article, and submitted contributions and amendments and agreed to the present version of the article.

Data Availability

The genomes analyzed during this study are publicly available at NCBI with the following accession numbers: Strain 220: CP098423; Strain 5844: CP098427.

Ethics Approval

The anonymized case description has been carried out in accordance with the Declaration of Helsinki, as revised in 2013.

Disclosure Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding Information

The present work was funded by Ricerca Corrente no. 08022218 ''Real-Time bioinformatic surveillance of nosocomial infections and multidrug resistant bacteria to improve infection control.''

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

Downloaded by Greta Petazzoni from www.liebertpub.com at 01/23/26. For personal use only.

References

1. Sato T, Yamawaki K. Cefiderocol: Discovery, chemistry, and in vivo profiles of a novel siderophore cephalosporin. Clin Infect Dis 2019;69(Suppl 7):S538–S543; doi: 10.1093/cid/ciz826
2. Hobson CA, Cointe A, Jacquier H, et al. Cross-resistance to cefiderocol and ceftazidime-avibactam in KPC β-lactamase mutants and the inoculum effect. Clin Microbiol Infect 2021;27(8):1172.e7–e1172.e10; doi: 10.1016/j.cmi.2021.04.016
3. Klein S, Boutin S, Kocer K, et al. Rapid development of cefiderocol resistance in carbapenem-resistant *Enterobacter cloacae* during therapy is associated with heterogeneous mutations in the catecholate siderophore receptor cirA. Clin Infect Dis 2022;74(5):905–908; doi: 10.1093/cid/ciab511
4. Nurjadi D, Kocer K, Chanthalangsy Q, et al. New Delhi metallo-beta-lactamase facilitates the emergence of cefiderocol resistance in *Enterobacter cloacae*. Antimicrob Agents Chemother 2022;66(2):e0201121; doi: 10.1128/AAC.02011-21
5. Tiseo G, Falcone M, Leonildi A, et al. Meropenem-vaborbactam as salvage therapy for ceftazidime-avibactam-, cefiderocol-resistant ST-512 *Klebsiella pneumoniae*-producing KPC-31, a D179Y variant of KPC-3. Open Forum Infect Dis 2021;8(6):ofab141; doi: 10.1093/ofid/ofab141
6. Yao J, Wang J, Chen M, et al. Cefiderocol: An overview of its in-vitro and in-vivo activity and underlying resistant mechanisms. Front Med (Lausanne) 2021;8:741940; doi: 10.3389/fmed.2021.741940
7. Liao Q, Deng J, Feng Y, et al. Emergence of ceftazidime-avibactam resistance due to a novel *bla*KPC-2 mutation during treatment of carbapenem-resistant *Klebsiella pneumoniae* infections. *J Infect Public Health*. 2022 May;15(5):545–549. doi: 10.1016/j.jiph.2022.04.002.
8. Sharma R, Park TE, Moy S. Ceftazidime-avibactam: A novel cephalosporin/β-lactamase inhibitor combination for the treatment of resistant Gram-negative organisms. Clin Ther 2016;38(3):431–444; doi: 10.1016/j.clinthera.2016.01.018
9. De Coster W, D'Hert S, Schultz DT, et al. NanoPack: Visualizing and processing long-read sequencing data. Bioinformatics 2018;34(15):2666–2669; doi: 10.1093/bioinformatics/bty149
10. Wick RR, Judd LM, Gorrie CL, et al. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017;13(6):e1005595; doi: 10.1371/journal.pcbi.1005595
11. Wick RR, Schultz MB, Zobel J, et al. Bandage: Interactive visualization of de novo genome assemblies. Bioinformatics 2015;31(20):3350–3352; doi: 10.1093/bioinformatics/btv383
12. Schmartz GP, Hartung A, Hirsch P, et al. PLSDB: Advancing a comprehensive database of bacterial plasmids. Nucleic Acids Res 2022;50(D1):D273–D278; doi: 10.1093/nar/gkab1111
13. Lam MMC, Wick RR, Watts SC, et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nat Commun 2021;12(1):4188; doi: 10.1038/s41467-021-24448-3
14. Alcock BP, Raphenya AR, Lau TTY, et al. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 2020;48(D1):D517–D525; doi: 10.1093/nar/gkz935
15. Carattoli A, Zankari E, García-Fernández A, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 2014;58(7):3895–3903; doi: 10.1128/AAC.02412-14
16. Chen L, Zheng D, Liu B, et al. VFDB 2016: Hierarchical and refined dataset for big data analysis—10 years on. Nucleic Acids Res 2016;44(D1):D694–D697; doi: 10.1093/nar/gkv1239
17. Durrant MG, Li MM, Siranosian BA, et al. A bioinformatic analysis of integrative mobile genetic elements highlights their role in bacterial adaptation. Cell Host Microbe 2020;27(1):140–153.e9; doi: 10.1016/j.chom.2019.10.022. Erratum in: Cell Host Microbe 2020;28(5):767.
18. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 2014;30(14):2068–2069; doi: 10.1093/bioinformatics/btu153
19. Martin MJ, Corey BW, Sannio F, et al. Anatomy of an extensively drug-resistant *Klebsiella pneumoniae* outbreak in Tuscany, Italy. Proc Natl Acad Sci U S A 2021;118(48):e2110227118; doi: 10.1073/pnas.2110227118
20. Karakonstantis S, Rousaki M, Kritsotakis EI. Cefiderocol: Systematic review of mechanisms of resistance, heteroresistance and in vivo emergence of resistance. Antibiotics 2022;11(6):723; doi: 10.3390/antibiotics11060723
21. Delgado-Valverde M, Conejo MDC, Serrano L, et al. Activity of cefiderocol against high-risk clones of multidrug-resistant *Enterobacterales, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Stenotrophomonas maltophilia*. J Antimicrob Chemother 2020;75(7):1840–1849; doi: 10.1093/jac/dkaa117

Address correspondence to:
Aurora Piazza, PhD
Department of Clinical, Surgical, Diagnostic and Paediatric Sciences
University of Pavia
Via Brambilla 74
Pavia 27100
Italy

E-mail: aurora.piazza@unipv.it

Pan-pathogen deep sequencing of nosocomial bacterial pathogens in Italy in spring 2020: a prospective cohort study

Harry A Thorpe, Maiju Pesonen, Marta Corbella*, Henri Pesonen, Stefano Gaiarsa, Christine J Boinett, Gerry Tonkin-Hill, Tommi Mäklin, Anna K Pöntinen, Neil MacAlasdair, Rebecca A Gladstone, Sergio Arredondo-Alonso, Teemu Kallonen, Dorota Jamrozy, Stephanie W Lo, Chrispin Chaguza, Grace A Blackwell, Antti Honkela, Anita C Schürch, Rob J L Willems, Cristina Merla, Greta Petazzoni, Edward J Feil, Patrizia Cambieri, Nicholas R Thomson†, Stephen D Bentley†, Davide Sassera†, Jukka Corander†*

Summary

Background Nosocomial infections pose a considerable risk to patients who are susceptible, and this is particularly acute in intensive care units when hospital-associated bacteria are endemic. During the first wave of the COVID-19 pandemic, the surge of patients presented a significant obstacle to the effectiveness of infection control measures. We aimed to assess the risks and extent of nosocomial pathogen transmission under a high patient burden by designing a novel bacterial pan-pathogen deep-sequencing approach that could be integrated with standard clinical surveillance and diagnostics workflows.

Methods We did a prospective cohort study in a region of northern Italy that was severely affected by the first wave of the COVID-19 pandemic. Inpatients on both ordinary and intensive care unit (ICU) wards at the San Matteo hospital, Pavia were sampled on multiple occasions to identify bacterial pathogens from respiratory, nasal, and rectal samples. Diagnostic samples collected between April 7 and May 10, 2020 were cultured on six different selective media designed to enrich for *Acinetobacter baumannii*, *Escherichia coli*, *Enterococcus faecium*, *Enterococcus faecalis*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Streptococcus pneumoniae*, and DNA from each plate with positive growth was deep sequenced en masse. We used mSWEEP and mGEMS to bin sequencing reads by sequence cluster for each species, followed by mapping with snippy to generate high quality alignments. Antimicrobial resistance genes were detected by use of ARIBA and CARD. Estimates of hospital transmission were obtained from pairwise bacterial single nucleotide polymorphism distances, partitioned by within-patient and between-patient samples. Finally, we compared the accuracy of our binned *Acinetobacter baumannii* genomes with those obtained by single colony whole-genome sequencing of isolates from the same hospital.

Findings We recruited patients from March 1 to May 7, 2020. The pathogen population among the patients was large and diverse, with 2148 species detections overall among the 2418 sequenced samples from the 256 patients. In total, 55 sequence clusters from key pathogen species were detected at least five times. The antimicrobial resistance gene prevalence was correspondingly high, with key carbapenemase and extended spectrum ß-lactamase genes detected in at least 50 (40%) of 125 patients in ICUs. Using high-resolution mapping to infer transmission, we established that hospital transmission was likely to be a significant mode of acquisition for each of the pathogen species. Finally, comparison with single colony *Acinetobacter baumannii* genomes showed that the resolution offered by deep sequencing was equivalent to single-colony sequencing, with the additional benefit of detection of co-colonisation of highly similar strains.

Interpretation Our study shows that a culture-based deep-sequencing approach is a possible route towards improving future pathogen surveillance and infection control at hospitals. Future studies should be designed to directly compare the accuracy, cost, and feasibility of culture-based deep sequencing with single colony whole-genome sequencing on a range of bacterial species.

Funding Wellcome Trust, European Research Council, Academy of Finland Flagship program, Trond Mohn Foundation, and Research Council of Norway.

Copyright © 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Nosocomial infections caused by multidrug-resistant bacterial pathogens are a major cause of mortality and morbidity among patients in hospital and add a significant economic and management burden to health-care systems. According to estimates of infection prevalence in Europe from 2016 to 2017, 5·9% of all patients and 19·2% of patients on intensive care units (ICUs) had a nosocomial infection.[1] Another European study, which used data from 2015, estimated that 63·5% of all infections with antimicrobial resistant (AMR) bacteria were associated with health care.[2] This threat is potentially elevated under circumstances where the inflow of critically ill patients is rapidly increased, such as during the COVID-19 pandemic. Pathogen

Lancet Microbe 2024; 5: 100890

Published Online August 20, 2024
https://doi.org/10.1016/S2666-5247(24)00113-7

*Contributed equally
†Contributed equally

Department of Biostatistics, Faculty of Medicine, University of Oslo, Oslo, Norway (H A Thorpe PhD, M Pesonen PhD, H Pesonen DSc, G Tonkin-Hill PhD, A K Pöntinen PhD, N MacAlasdair PhD, R A Gladstone PhD, S Arredondo-Alonso PhD, Prof J Corander PhD); **Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy** (M Corbella MSc, S Gaiarsa PhD, C Merla PhD, G Petazzoni MSc, P Cambieri MD); **Parasites and Microbes, Wellcome Sanger Institute, Cambridge, UK** (C J Boinett PhD, G Tonkin-Hill, N MacAlasdair, D Jamrozy PhD, S W Lo PhD, C Chaguza PhD, G A Blackwell PhD, Prof N R Thomson PhD, Prof S D Bentley PhD, Prof J Corander); **Department of Computer Science, University of Helsinki, Helsinki, Finland** (T Mäklin PhD, Prof A Honkela PhD); **Institute of Biomedicine, University of Turku, Turku, Finland** (T Kallonen PhD); **Department of Medical Microbiology, Universitair Medisch Centrum Utrecht, Utrecht, Netherlands** (A C Schürch PhD, Prof R J L Willems); **Department of Medical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy** (G Petazzoni); **Milner Centre for Evolution, University of Bath, Claverton Down, Bath, UK** (Prof E J Feil PhD); **Department of Biology and Biotechnology, University of Pavia, Pavia, Italy** (D Sassera PhD); **Fondazione IRCCS Policlinico San Matteo, Pavia, Italy** (D Sassera);

Helsinki Institute for Information Technology, Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland (Prof J Corander)

Correspondence to: Dr Harry A Thorpe, Department of Biostatistics, Faculty of Medicine, University of Oslo, 0313 Oslo, Norway
harry.thorpe@medisin.uio.no

or

Prof Jukka Corander, Department of Biostatistics, Faculty of Medicine, University of Oslo, 0313 Oslo, Norway
jukka.corander@medisin.uio.no

Research in context

Evidence before this study

Soon after the emergence of SARS-CoV-2, multiple clinical studies established that mortality from severe infection was highest among older, immunocompromised individuals and patients with severe comorbidities. Several follow-up studies investigating the association between bacterial co-infections and COVID-19 mortality concluded that the effect of co-infections tends to be minimal. However, it is expected that in regions with a high burden of endemic nosocomial pathogens, there is an elevated risk of hospital transmission of virulent and multidrug-resistant bacterial strains during invasive procedures, such as ventilator treatment. We first searched PubMed from database inception to Aug 7, 2023, with the search terms "SARS-CoV-2" AND "bacterial" AND "co-infections", which resulted in 636 publications, including reviews and meta-analyses. We then searched for "SARS-CoV-2" AND "co-infection" AND "whole-genome sequencing", which returned only seven publications, five of which focused on SARS-CoV-2. This indicates a paucity of studies relying on whole-genome sequencing of bacterial pathogens that were circulating in health-care settings during the pandemic-driven disruption of infection control. Both searches were restricted to articles published in English.

Added value of this study

Our study presents the first attempt to recover bacterial strains at single nucleotide polymorphism-level resolution, while simultaneously capturing most clinically relevant major bacterial pathogens present in the gut, upper airways, or lungs of the patients. The deep-sequencing approach used in the study enabled quantification of antimicrobial resistance prevalence and estimation of hospital transmission. By sampling patients within a region severely affected by both the COVID-19 pandemic and endemic circulation of nosocomial bacterial pathogens, we were able to assess transmission in a high-risk setting.

Implications of all the available evidence

Our study serves as a proof-of-concept of how to use the power of deep sequencing to investigate multiple relevant pathogenic organisms simultaneously instead of a siloed approach based on whole-genome sequencing of single isolates of individual species. The approach we describe could facilitate the development of improved future guidelines for assessing and managing the risk of nosocomial infections for different patient populations, particularly on intensive care units.

surveillance is key to managing the burden of nosocomial infections, and in the last 15 years genomics has become a powerful tool for tracking pathogen evolution and transmission.[3,4]

Whole-genome sequencing (WGS) of DNA from a pure colony identified after culturing is considered the gold standard for pathogen surveillance.[4–6] However, this approach has a major limitation: it only provides a partial snapshot of the bacterial diversity that can be present in a clinical sample. This limitation can, in principle, be remedied by picking and sequencing multiple colonies per sample, but this is, in practice, both expensive and time-consuming. Shotgun metagenomics has the advantage of being able to capture all microbial diversity present in a sample and has been applied clinically, for example to establish the aetiology of conditions with unknown cause, but with variable success.[7–9] However, this approach is not generally suitable for genomic epidemiology and transmission analysis which require strain-level resolution, owing to the high sample biomass and sequencing depth required.

These issues have motivated the development of culture-based metagenomic profiling, whereby bacteria of interest are first enriched by culturing on selective media, followed by deep sequencing of the total DNA present.[10] This approach maintains the genetic diversity present within the taxa of interest, while eliminating the problems of low biomass samples, contamination with host DNA, or low abundance of taxa of interest. Because the selected taxa are typically closely related, potentially including several closely related strains, sequence types (STs), or sequence clusters (SCs) from the same species, standard metagenomics methods do not provide sufficient resolution to separate strains, especially from short-read sequencing data. To overcome this limitation, we developed a reference-based deconvolution algorithm (implemented in mSWEEP[11] and mGEMS[12]) that can accurately reconstruct individual bacterial genomes from metagenomic sequencing data.[13,14] This culture-based, deep-sequencing approach has been used to study the major community-acquired pathogen *Streptococcus pneumoniae*, showing that the richer data facilitated finer-scale resolution in both transmission and evolutionary analyses than by using a single colony-based approach.[14]

In this study, we used deep sequencing after culturing, combined with ST-level deconvolution with mSWEEP and mGEMS, to characterise the bacterial pathogen population within a single hospital in Pavia, northern Italy, during the first wave of the COVID-19 pandemic. Northern Italy has a high burden of AMR infections and was the first region in Europe to be severely affected by COVID-19; such conditions presented serious challenges to infection control, risking increased nosocomial transmission of bacterial pathogens. We focused on *Acinetobacter baumannii*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Enterococcus faecium*, *Enterococcus faecalis*, *Escherichia coli*, and *Staphylococcus aureus* as these species contribute the most to the overall burden of AMR and nosocomial infections according to recent estimates.[2] The aims of the study were to identify the bacterial pathogens and AMR-associated

genes present in the hospital, assess their distribution according to ward and sample type, characterise the pathogen genetic diversity, and infer transmission, and compare the deep-sequencing data with existing single-colony whole-genome sequence data.

Methods

Study design and participants

This prospective cohort study was done in an 800-bed teaching hospital serving the Lombardy region in northern Italy (Fondazione IRCCS Policlinico San Matteo, Pavia, Italy) during the first wave of the COVID-19 pandemic in the spring of 2020. Nasal, rectal, and respiratory samples collected according to the routine screening and diagnostic protocols at the hospital were used, and no exclusion was applied on the basis of hospital ward, diagnosis or patient characteristics. For each patient, we recorded the dates of admission and discharge, hospital ward, age, and sex (appendix 1 p 1).

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (2023-3.11/105). The work described herein is a prospective study done on bacterial isolates from human samples that were obtained as part of routine hospital care and hence the requirement for patient informed consent was waived.

Culturing

Samples were cultured on six different media to select for pathogen species of interest. Chocolate agar plus PolyViteX (PVX), and Columbia agar plus 5% sheep blood, (COS), were used to select for all bacteria and fungi. Columbia Colistin-Nalidixic Acid agar plus 5% sheep blood (CNA), was used to select for Gram-positive bacteria. McConkey agar (MCK), was used to select for Gram-negative bacteria. Mannitol Salt Agar (MSA), was used to select for *Staphylococcus* spp. CHROMID Candida Agar (CAN), was used to select for *Candida* spp. All agar was purchased from Biomerieux, Marcy l'Etoile, France. Nasal swabs were plated on PVX, CNA, and CAN; rectal swabs were plated on CNA, CAN, and MCK; and respiratory samples were plated on PVX, COS, MCK, MSA, and CAN.

All samples were cultured for 48 h (at 37°C, PVX at 5% CO_2), and plates were swept to collect all microorganisms, and material was dissolved in saline buffer, pelleted, and frozen (appendix 1 pp 1–2, 7).

DNA extraction, library preparation, and sequencing

DNA was extracted by use of the QIAGEN DNeasy 96 Power Soil Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. DNA sequencing library preparation was done with the QIAseq FX DNA Library Kit (QIAGEN, Hilden, Germany). WGS was done on the Illumina Novaseq 6000 platform (Illumina, San Diego, CA, USA) with 150 bp paired-end reads (appendix 1 p 2).

Reference database construction

We built two sets of reference databases for use with the mSWEEP–mGEMS[11,12] pipeline; the first contained a broad set of species, and the second consisted of detailed databases of the pathogen species of interest. To build the broad reference, we used Kraken version 1.1.1[15] to identify the species present within the sequencing read sets (appendix 1 p 3). We identified 101 species and downloaded their genomes from the National Center for Biotechnology Information and subjected them to quality control (appendix 1 pp 20–21). For each pathogen species of interest we used high-quality, detailed sets of genome assemblies from curated collections to enable SC-level identification.[16–20] Each of these reference sets was clustered into SCs by use of PopPUNK.[21] Each SC was additionally labelled by the most common ST contained within it, as these are widely recognised groups (for example *K pneumoniae* SC1_ST307). The two sets of reference genomes were combined to give a database of 10 993 genomes representing 101 species (appendix 1 pp 2–3, 21–29).

See Online for appendix 1

Species and sequence cluster detection

We used a reference-based, hierarchical approach to detect species and SCs accurately (appendix 1 p 10). Themisto version 0.2.0[22] was used to pseudoalign the sequence reads to the combined reference database, mSWEEP version 1.4.0[11] was used to estimate species abundances, and mGEMS version 1.0.0[12] was used to bin the reads into species bins. For each pathogen species of interest, the species-binned reads were used as input to a second round of the pipeline, with detailed species-specific reference databases, to estimate SC abundances and bin the reads into SC bins. To ensure that species and SC assignments were accurate, we used Mash version 2.3[23] to check that the genetic distances were within the expected range for each cluster (appendix 1 pp 3–4).

The code for our demix checking pipeline is available at https://github.com/harry-thorpe/demix_check

Antimicrobial resistance gene detection

We used ARIBA version 2.14.6[24] with the CARD version 3.2.8[25] database to identify AMR genes in the sequence read sets (appendix 1 p 4).

Mapping and single nucleotide polymorphism calling

Mapping was done with Snippy version 4.6, and Snippy-core was used to generate whole-genome and core-genome alignments. We then used snp-dists version 0.7 to calculate pairwise single nucleotide polymorphism (SNP) distances from these alignments. For the analysis of *A baumannii* SC1_ST2, we followed the same procedure, but we also incorporated the single-colony whole-genome sequence reads from Petazzoni and colleagues.[26] We did phylogenetic analysis by use of FastTree version 2.1.10 Double precision[27] (using the generalised time-reversible plus CAT model; appendix 1 pp 4–5).

Code available at https://github.com/tseemann/snippy

Code available at https://github.com/tseemann/snp-dists

Transmission analysis

We used the within-patient binned mapped genome (BMG) SNP distance distributions to identify the between-patient BMG pairs that were likely to be the result of transmission, assuming that pairs of isolates resulting from either recent transmission or repeated sampling of the same patient would be indistinguishable. For each SC, we used the 95th percentile within-patient BMG SNP distance as a threshold, and then selected the between-patient BMG pairs with SNP distance of no more than the threshold as putative transmission events (appendix 1 p 5).

See Online for appendix 3

Statistical analysis

The presence of pathogen species in ICUs versus ordinary wards was compared by use of Fisher's exact tests, whereas the presence of species versus sample types was compared by use of Chi squared tests. A two-sample *t* test was used to compare the mean number of pathogen species carried by patients between ICUs and ordinary wards. All tests were done by use of R version 4.3.3. Significant difference was considered at $p<0·05$, and p values were adjusted for multiple comparisons with Bonferroni correction where appropriate.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We applied a novel culture-based deep-sequencing approach to a set of 1130 clinical samples (497 rectal swabs, 439 nasal swabs, and 194 respiratory samples), which were collected for diagnostic purposes and processed in parallel to do this study (appendix 1 p 7). The enrolled cohort consisted of 256 patients who were admitted to the San Matteo hospital between March 1 and May 7, 2020 (169 men and 87 women; median age 63 years [IQR 51–72]), of whom 143 were on ordinary wards and 125 on ICUs (12 patients were present on both at some point). 45 (18%) patients died during the course of the study. The number of admissions per day, age distribution by sex, hospital stay duration, and sampling times per patient are shown in appendix 1 (p 8). The 1130 clinical samples were cultured on six different media types to select for pathogens of interest (appendix 1 p 7), resulting in growth on 2681 of 3786 plates (median positive samples per patient 2 [IQR 1–6]).

See Online for appendix 5

See Online for appendix 6

We sequenced the total DNA from 2418 plate sweeps to high depth (mean 30 million reads; appendix 4), and used mSWEEP and mGEMS, with a broad reference database of 10 993 genomes (appendix 1 pp 10, 20–21; appendix 2), chosen after an initial scan of the reads with Kraken (appendix 1 p 9), to identify species and bin reads into species bins. These steps resulted in 3233 high-confidence assignments to 52 different species (appendix 1 pp 11–12; appendix 4), of which 2148 (66%) were assigned to seven of the eight pathogen species of interest: *A baumannii, E coli, E faecium, E faecalis, K pneumoniae, P aeruginosa,* or *S aureus* (appendix 1 pp 11–12). For these pathogen species, we used the species-binned reads as input to a second round of the binning pipeline, with detailed species-specific reference databases (appendix 1 pp 10, 22–29; appendix 3), resulting in 2238 high-confidence assignments to 132 different SCs (figure 1; appendix 4).

See Online for appendix 4

See Online for appendix 2

The number of patients carrying each SC is shown in figure 1. All *A baumannii* assignments were identified as ST2 (global clone 2),[28] in agreement with single colony whole-genome sequences collected from the same hospital during a nosocomial outbreak that overlapped the time period covered by this study.[26] In total, 15 *E coli* SCs were identified, of which the most common (SC2_ST131) corresponded to ST131, an ST which has been frequently reported in screening of nosocomial colonisation.[29] 22 *E faecium* and 39 *E faecalis* SCs were identified, including the hospital-associated *E faecium* ST117[17] and *E faecalis* ST6 and ST28.[18] The most common of the 19 *K pneumoniae* SCs corresponded to the hospital-associated clones ST307, ST392, and ST45.[19,30] Of the 23 *P aeruginosa* SCs detected, the most common corresponded to ST253, also known as PA14, which is the most common clone in *Pseudomonas* group 2.[31] For *S aureus*, the most common of the 15 detected SCs corresponded to the hospital-associated ST22[32] and the livestock-associated ST398.[33,34] Aside from these pathogen species of main clinical interest, the most common species identified from the samples were other *Staphylococcus, Streptococcus,* and *Enterococcus* spp (appendix 1 p 13).

The number of species detected in an individual patient was significantly higher in ICUs (mean 3·8 [SD 1·8]) than in ordinary wards (mean 1·6 [SD 1·2]) (*t* test, $p<0·0001$). We then tested the association between each pathogen species and ward type, revealing that the presence of *A baumannii, K pneumoniae, P aeruginosa*, and *E faecium* was significantly associated with ICUs for both nasal and rectal samples and *E faecalis* was associated with ICUs for nasal samples, whereas *E coli* from rectal samples was associated with ordinary wards (figure 2A and appendix 5). When this analysis was done at the level of individual SCs (but without considering sample types separately), nine SCs were significantly associated with ICUs after multiple correction (appendix 6).

We screened for AMR-associated genes within the read sets using ARIBA with the CARD database, focusing on carbapenemase and extended spectrum ß-lactamase (ESBL) genes, and genes conferring vancomycin resistance (figure 2B). The most common carbapenemase genes were $bla_{\text{OXA-23}}$, $bla_{\text{KPC-2}}$, and the $bla_{\text{OXA-51}}$-like $bla_{\text{OXA-66}}$; these genes were detected in 77 (62%) of 125, 76 (61%), and 62 (50%) patients on ICUs, respectively, and in lower numbers on ordinary wards 19 (14%) of 138 patients, 31 (23%), and 12 (9%), respectively. We detected a large number of ESBL genes, of which the most common were $bla_{\text{OXA-1}}$, $bla_{\text{CTX-M-15}}$, $bla_{\text{TEM-187}}$, and $bla_{\text{CTX-M-14}}$; these were present in 81 (65%) of 125, 77 (62%), 64 (51%), and 47 (38%) patients on ICUs, respectively. We detected genes from both

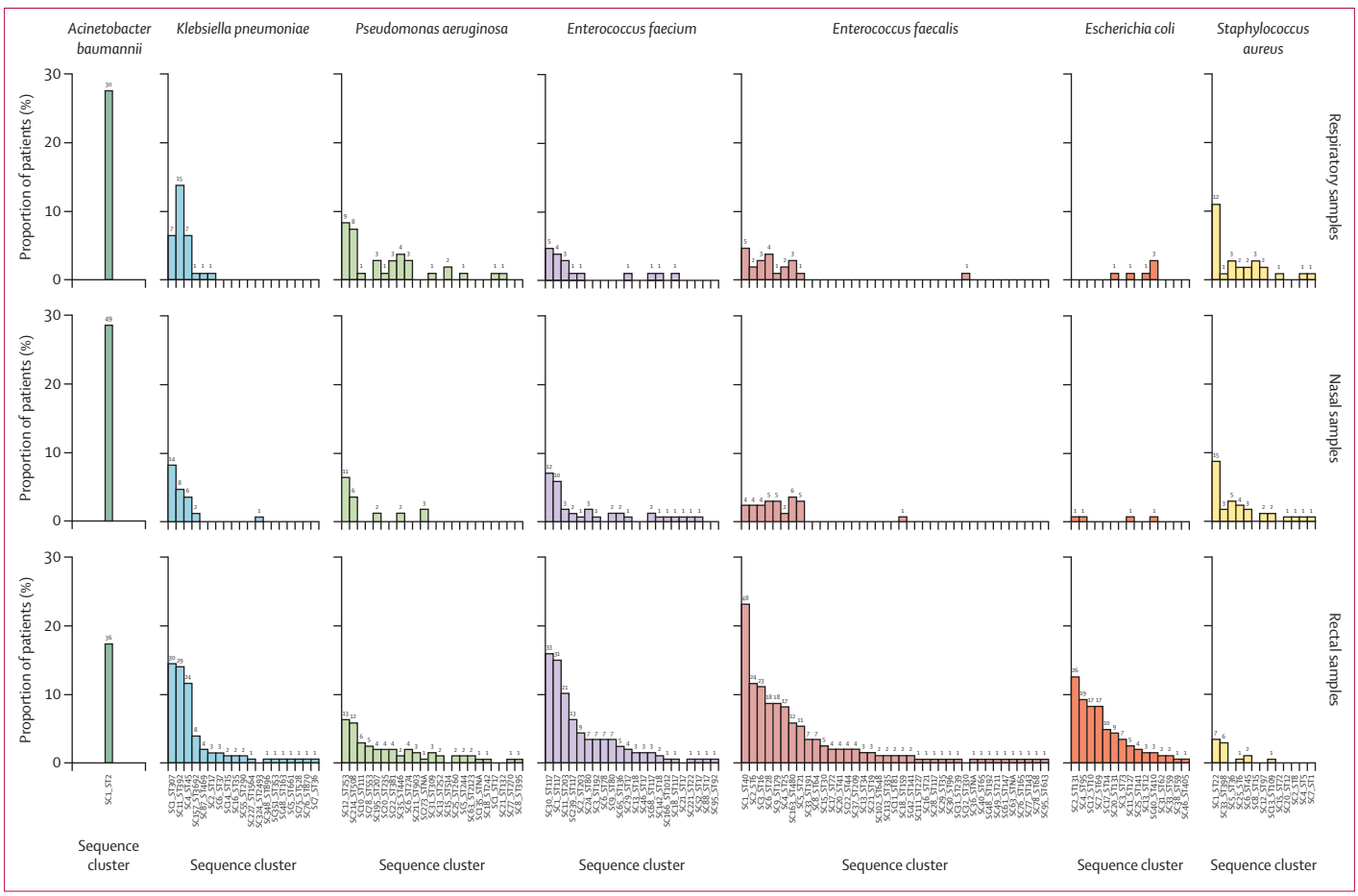


Figure 1: **The prevalence of pathogen species and SCs among the patients**
The percentage of patients who were colonised by each SC is shown, calculated separately for each sample type. The numbers above each bar indicate the corresponding numbers of patients with positive samples. Within each species, the SCs are ordered by overall prevalence. SC=sequence cluster. ST=sequence type.

the *vanA* and *vanB* vancomycin resistance operons, with the *vanA* operon being more common (50 [40%] patients on ICUs) than the *vanB* operon (nine [7%] patients on ICUs).

Deep sequencing and accurate read binning enabled us to do whole-genome mapping and SNP calling, but to differentiate our samples from traditional single colony samples we refer to ours as BMGs; mapping quality shown in appendix 1 (pp 14–15). We first investigated within-patient diversity for each SC that was detected at least five times (n=55) (figure 3; appendix 1 p 16). The core-genome SNP diversity was extremely low; 46 (84%) of 55 SCs had a median SNP distance of 0, and 49 (89%) contained no BMG pairs with a distance of ten or more SNPs (figure 3). More diversity was observed when whole-genome SNPs were considered, but this was still relatively low; 40 (73%) of 55 SCs had a median SNP distance of no more than three, and 50 (91%) had a median SNP distance of no more than eight (appendix 1 p 16). This suggests that the patients generally carried only one major lineage from each SC (along with some minor strain variation), even if they carried multiple SCs from the same species. Notable exceptions to this observation were found for *A baumannii* SC1_ST2, whereby one patient carried two sublineages of this SC, with BMGs differing by 510–13 core-genome SNPs (807–94 whole-genome SNPs), and *S aureus* SC33_ST398, whereby one patient carried two sublineages with BMGs differing by 224–26 core-genome SNPs (262–331 whole-genome SNPs).

We then investigated between-patient diversity in the same SCs (figure 3; appendix 1 p 16), which was overall much more extensive than the diversity within patients but with considerable differences between the species in their patterns of diversity. For the notorious hospital-associated species and lineages, a low amount of diversity was observed even between patients. *A baumannii* SC1_ST2 BMGs had a median pairwise distance of two core-genome SNPs (IQR 0–2) and five whole-genome SNPs (3–9), with 25% of pairs being identical according to the core-genome SNPs. There were also several much more divergent pairs, separated by 509–16 core-genome SNPs (626–954 whole-genome SNPs), and these were explained by the two separate sublineages of this SC described above.

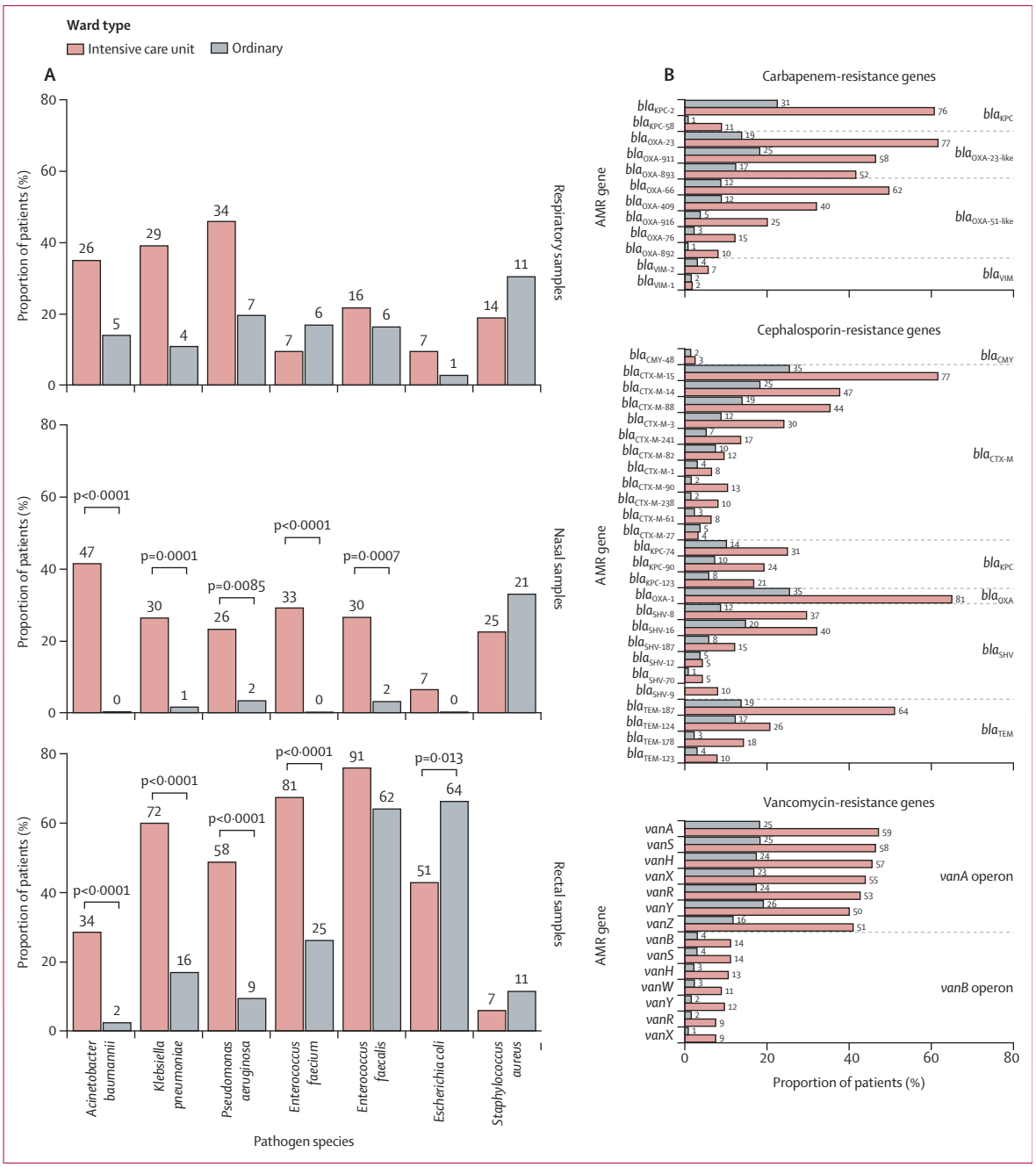


Figure 2: **Pathogen and AMR gene distribution among patients and within the hospital**

(A) Pathogen prevalence and distribution. The percentage of patients who were colonised by each pathogen species is shown, calculated separately for different ward and sample types. The numbers above each bar indicate the corresponding numbers of patients with positive samples, and the p values show the results of a Fisher's exact test of association between species prevalence and ward type. (B) AMR gene prevalence and distribution. AMR genes were detected in the raw deep sequencing reads, and the percentage of patients carrying each gene on ICU and ordinary wards is shown. The numbers to the right of each bar indicate the corresponding numbers of patients with positive samples. The three most important resistance classes among the pathogen species of interest are shown (genes conferring resistance to carbapenems, cephalosporins, and vancomycin). For carbapenem and cephalosporin resistance genes, the different *bla* types are separated by dashed lines, and ordered by prevalence within each group. For vancomycin resistance genes, the *vanA* and *vanB* operons are separated by a dashed line. AMR=antimicrobial resistance. ICU=intensive care unit.

K pneumoniae also showed minimal diversity, with four of five SCs having a median distance of 0 core-genome SNPs, corresponding to no more than ten whole-genome SNPs. For *E faecium*, seven of 12 SCs had a median pairwise distance of 0 core-genome SNPs, and three of 12 had distances between one and five SNPs. When whole-genome SNPs were considered, the distances were considerably greater, suggesting that the number of non-core SNPs was

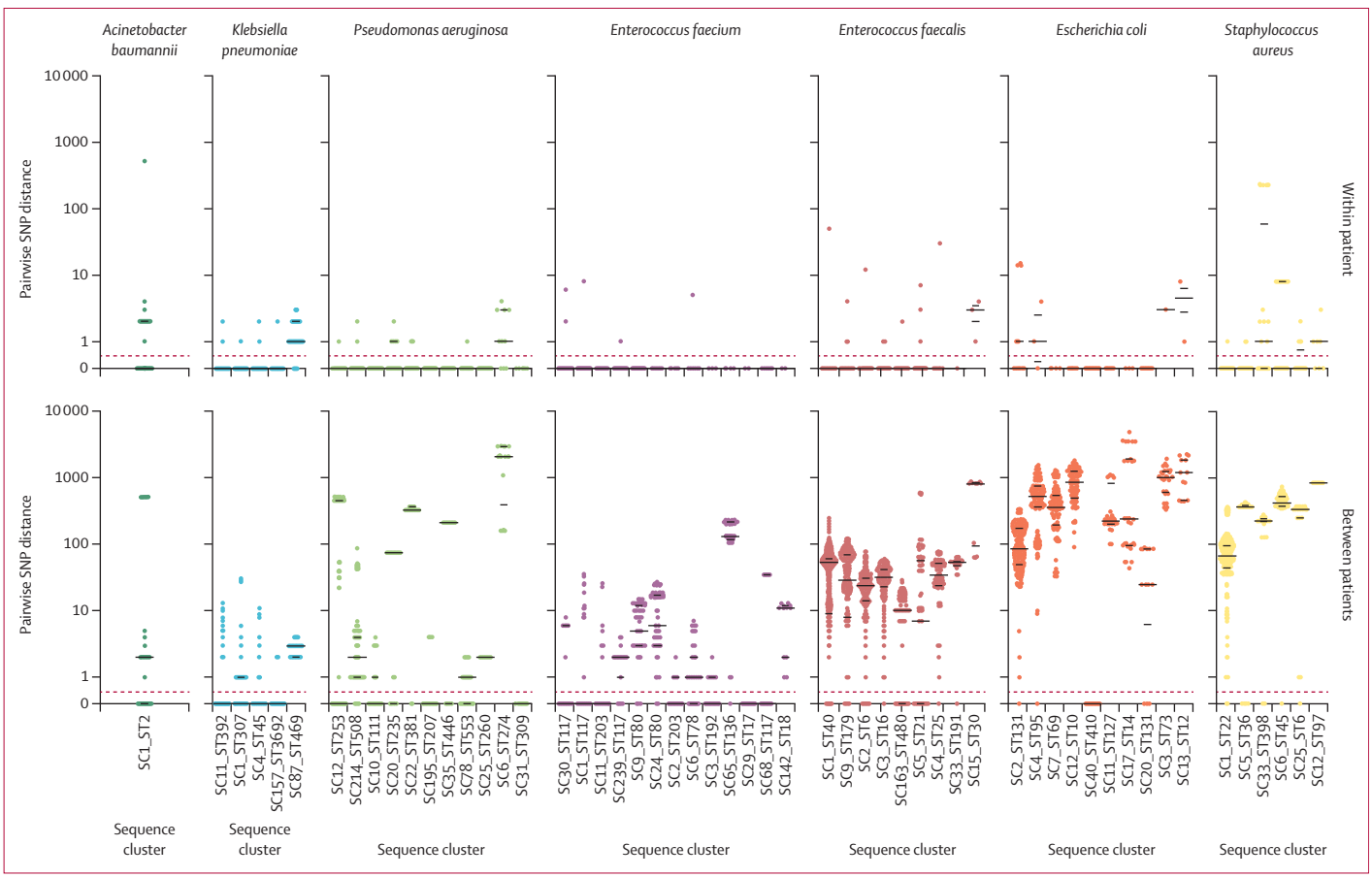


Figure 3: **Pairwise core-genome SNP distances among the SCs**

Core-genome pairwise SNP distances were calculated from mapping-based alignments for each SC. Each point shows a comparison between a pair of binned mapped genomes, and the comparisons are separated into those from the same patient (top), and different patients (bottom). The distributions are shown by violin scatter plots, with additional information given by the horizontal black bars: long bars depict medians, short bars depict 25th and 75th percentiles. SNP distances (y-axis) are displayed on a $\log_{10}$ scale; 0 has been included to ensure that identical genomes are present, with the red dashed lines indicating the break in the scale. Only SCs that were detected in at least five patients are shown. SC=sequence cluster. SNP=single nucleotide polymorphism. ST=sequence type.

significant. The SNP distances observed among *P aeruginosa* SCs were more variable, with six of 11 SCs having median distances of no more than two core-genome SNPs (≤6 whole-genome SNPs), and four of 11 having distances of more than 90 SNPs (for both core and whole genome).

For *E coli* SCs, the median SNP distance was more than 90 SNPs for eight of ten SCs in both core and whole genome, with the majority of SCs having SNP distances in the hundreds, indicating extensive between-patient diversity (figure 3; appendix 1 p 16). The exceptions were identical isolates, according to core-genome SNPs, of SC2_ST131 (carried by four patients), SC20_ST131 (two patients), SC40_ST410 (three patients), and SC4_ST95 (three patients). ST131 is a known health-care-associated clone,[35–37] and ST410 has previously been described as an emerging problem in hospitals.[38] *S aureus* also showed high diversity, with median distances of more than 100 SNPs for five of six SCs (both core-genome and whole-genome SNPs). However, there were a small number of identical BMGs (core-genome SNPs) shared between 24 patients in these SCs. *E faecalis* similarly showed considerable, albeit lower, diversity, with median distances of more than 32 core-genome SNPs for seven of nine SCs (>108 whole-genome SNPs). Identical BMGs were present in eight of nine SCs from this species according to both core-genome and whole-genome SNPs, and in two SCs these made up 25% of the pairs (core-genome SNPs).

The very low BMG SNP distances observed among the pathogen species strongly indicated hospital transmission, so we used the sequence data to estimate this burden for each pathogen. For each SC, we identified patients who were linked by at least one pair of BMGs, and then partitioned these occurrences into those that were putatively transmission-linked and those that were not. Putative transmission was highest within and between ICUs, followed by ICU–ordinary ward links (appendix 1 pp 17–18). The species with the highest percentage of transmission-linked occurrences were *A baumannii, K pneumoniae, P aeruginosa,* and *E faecium,* with more than 75% occurrences being putatively transmission-linked when calculated from core-genome SNPs (>81% from whole-genome SNPs; figure 4;

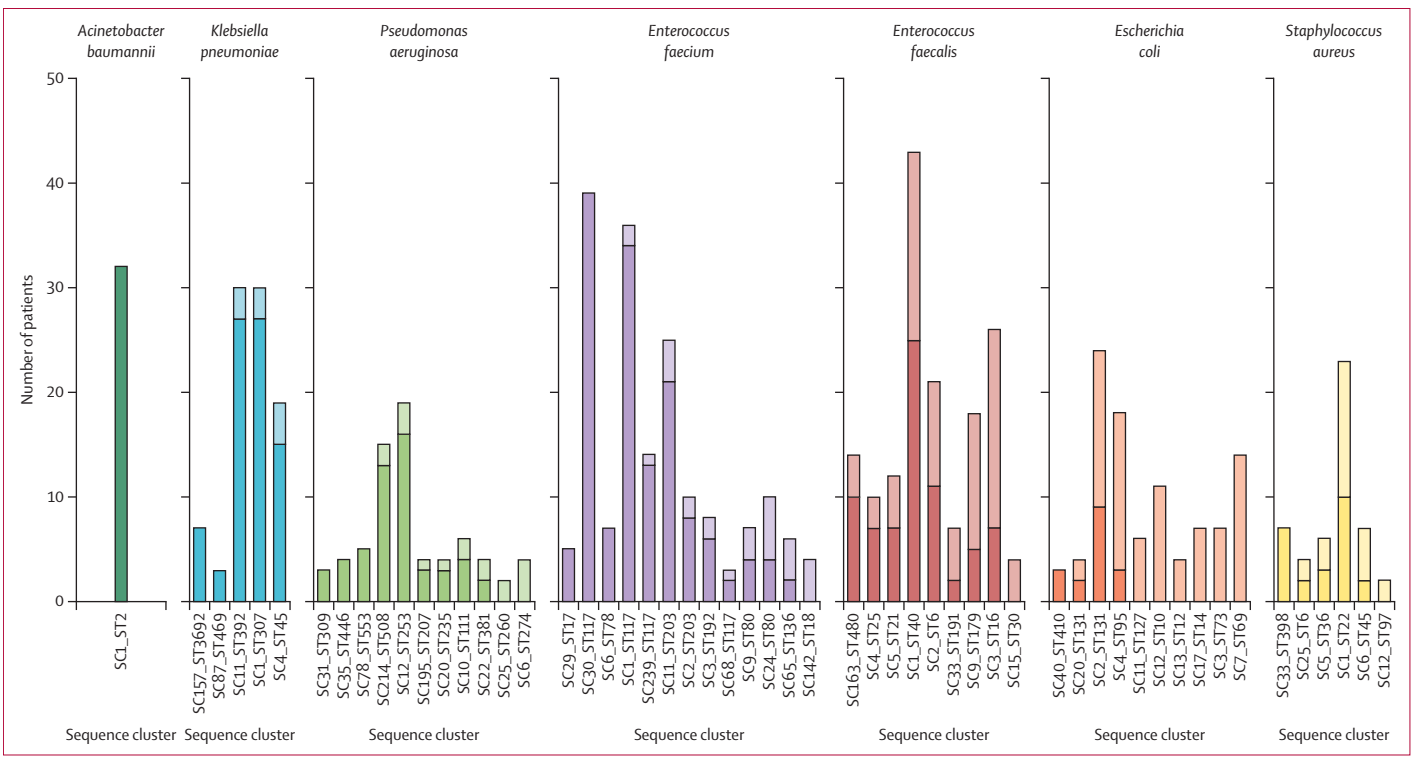


Figure 4: **The burden of hospital transmission based on estimates from core-genome SNP distances**
Core-genome SNP distances were used to infer networks of patients who carried binned mapped genomes that were closely related enough to be linked by transmission, according to SC-specific thresholds obtained from within-patient samples. For each SC, the occurrences of the binned mapped genomes among patients were partitioned into those that were transmission-linked (eg, those that were closely related enough to another binned mapped genome from a different patient; dark colour) and those that were not (light colour). As the height of the bar indicates the total number of patients from whom high-quality binned mapped genomes were obtained, the dark portion of the bar shows the proportion of these occurrences that were probably the result of hospital transmission. Only SCs that were detected in at least five patients are shown. SNP=single nucleotide polymorphism. ST=sequence type.

appendix 1 p 19). Within these species, the dominant SCs were *E faecium* SC1_ST117 and SC30_ST117, *A baumannii* SC1_ST2, and *K pneumoniae* SC1_ST307 and SC11_ST392, each with at least 25 linked patients. Approximately half of occurrences in *E faecalis* and *S aureus* were putatively transmission-linked (figure 4; appendix 1 p 19). *E coli* was least associated with transmission, with a few scattered examples of linked patients contributing to 17 (17%) of 98 occurrences (both core-genome and whole-genome SNPs). The results were highly concordant when calculated from core-genome and whole-genome SNPs, indicating that both measures captured similar information.

Motivated by the pattern of within-patient and between-patient diversity of *A baumannii* SC1_ST2, we compared all the *A baumannii* BMGs identified by mSWEEP-mGEMS with the reads from single-colony *A baumannii* isolates obtained previously from the same hospital and period[26] by combining them in a core-genome SNP phylogeny (figure 5A; microreact project). This showed two clearly separated clades of nearly identical genomes, belonging to the two SC1_ST2 subclones previously identified by Petazzoni and colleagues,[26] and both clades included samples from the two approaches (single colony and deep sequencing).

For more on the **microreact project** see https://microreact.org/project/7kCEtVo1po9Mb2vTxfX4uM-abau-demixed-single-colony-reads-core-sites-fasttree

Additionally, deep sequencing yielded reads from both subclones in samples from a single patient, whereas single colony whole-genome sequencing only identified one of the two, although colonies from multiple clinical samples were sequenced from this same patient. We therefore did a third round of the binning pipeline with a reference database consisting of the 91 single colony genomes belonging to SC1_ST2, labelled by subclone (SC1_1, n=68; SC1_2, n=23; figure 5B). This confirmed that both subclones were carried by the patient.

Discussion

Applying our deep-sequencing after culturing method to a cohort of hospital patients during the first COVID-19 wave in spring 2020, we discovered and characterised a substantial pathogen burden in the patient population, with 55 of 132 SCs being detected at least five times. The pathogen burden was highest in the ICUs, with *A baumannii*, *K pneumoniae*, *P aeruginosa*, and *E faecium* being significantly associated with ICUs. These pathogens were found to be colonising multiple body sites in a typical ICU patient; this is probably due to the high hospital prevalence of the pathogens and severely ill state of the patients, but it might also have been a consequence of the difficult situation regarding infection control during the

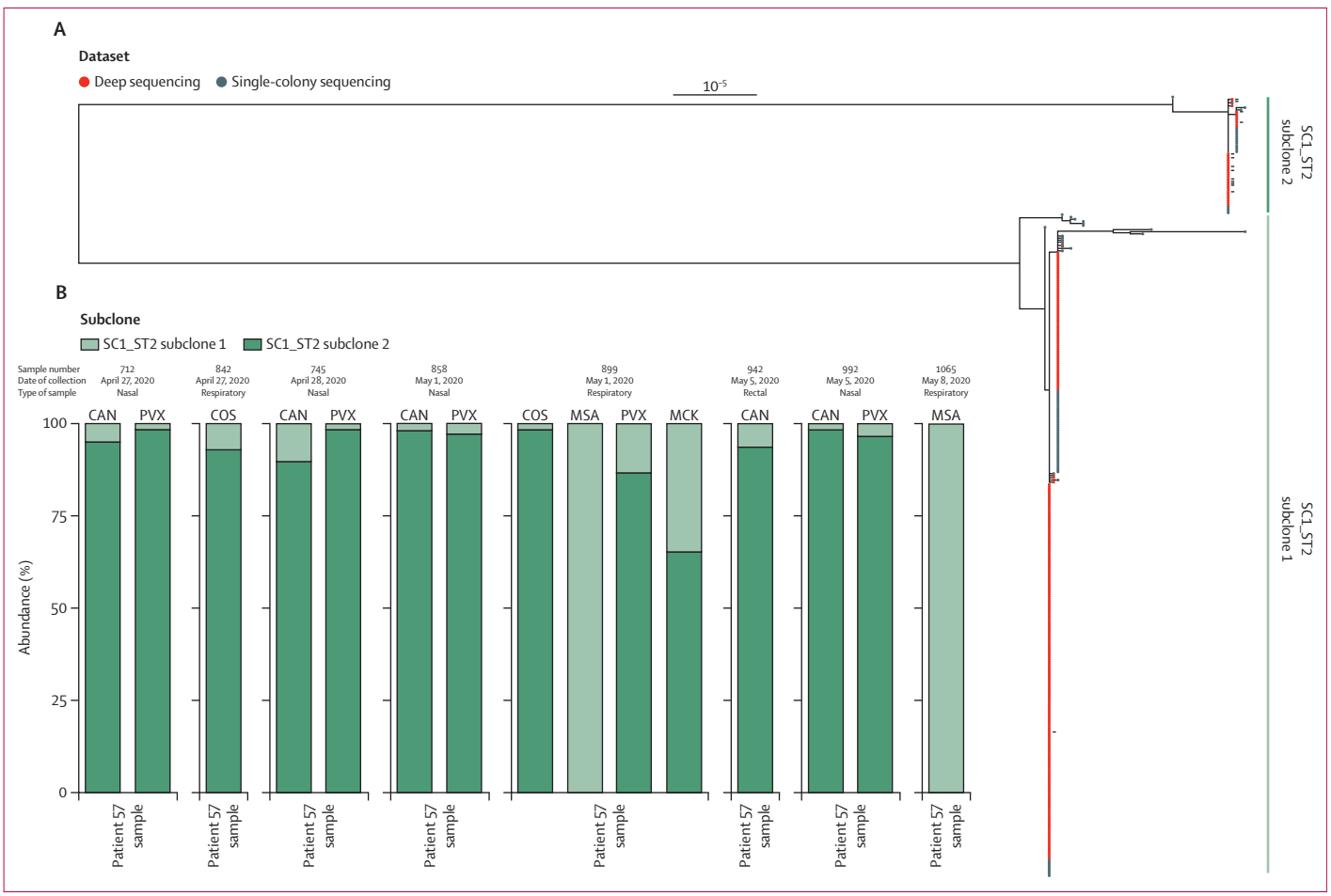


Figure 5: **Fine scale analysis of the *Acinetobacter baumannii* SC1_ST2 subclones**

(A) Phylogenetic tree of the two SC1_ST2 subclones. Single-colony whole-genome sequence data corresponding to *Acinetobater baumannii* from the same hospital as the current study (grey) was combined with the demixed deep sequencing reads from the current study (red). Reads were mapped to a common reference (SAMN07258611) by use of Snippy, and FastTree was used to generate a phylogenetic tree from the core-genome alignment (generalised time-reversible plus CAT model). The two SC1_ST2 subclades are labelled, and the scale bar indicates SNPs per site. (B) Reanalysis of the *A baumannii* SC1_ST2 subclones from a single patient who was found to carry both subclones. Each bar shows a single sequencing sample, and the bars are coloured by the subclone detected. The samples are grouped by the original swab sample (as one swab sample resulted in multiple sets of sequencing reads) and ordered by date of sample collection. This information, along with the sample type and culture media is shown above the bars. CAN=CHROMID Candida Agar. COS=Columbia agar + 5% sheep blood. MCK=McConkey agar. MSA=Mannitol Salt Agar. PVX=Chocolate agar + PolyViteX. SC=sequence cluster. ST=sequence type.

early pandemic. AMR gene prevalence was also very high, with several carbapenemase and ESBL genes being present in over 40% of patients on ICUs. Previous research from the same hospital with single colony whole-genome sequencing identified bla_{KPC-2} and bla_{OXA-23} as the dominant carbapenemase genes carried by *K pneumoniae* and *A baumannii*, respectively.[19,26] Such systemic presence of AMR bacteria highlights their ability to sustain themselves successfully in various niches in the hospital, as has been shown by other studies where sampling of the hospital environment was done.[4] We observed considerable differences in the genetic diversity among the pathogen species: *A baumannii* and *K pneumoniae* were dominated by very closely related BMGs; *E coli, E faecalis,* and *S aureus* were dominated by distantly related BMGs; and *P aeruginosa* and *E faecium* exhibited a mixture of both. Using these distances to infer transmission, we estimated that hospital transmission was likely to be a significant mode of acquisition for each of the pathogen species. For *A baumannii*, *K pneumoniae*, *P aeruginosa*, and *E faecium*, this was particularly significant with more than 75% of occurrences being potentially linked by transmission. These results emphasise the need for more comprehensive approaches to tracking the dissemination and evolution of these pathogens.

By comparing our results for *A baumannii* with the single colony-based whole-genome sequencing surveillance done during the same period at the study hospital,[26] we showed that our approach is able to detect and distinguish highly related strains with the sensitivity of the single-colony approach, and has the important added benefit of being able to delineate co-infections or co-colonisation within single patients.

Our study has some important limitations. Sequencing was done only on samples available through standard

diagnostic procedures, thus providing uneven sampling. Approximately one-third of the patients had already been admitted to the hospital when the sampling started, patients were not always sampled on admission, and we did not have access to samples taken before the start of the pandemic. Additionally, we did not perform power calculations before starting the sampling due to uncertainties about sample and personnel availability as a result of the COVID-19 pandemic. Another limitation is that although we detected AMR genes, we did not attempt to link them to their host species or strains. This was because many of these genes are known to be carried on plasmids and are present in multiple species, for example the $bla_{CTX\text{-}M\text{-}15}$ ESBL gene is common in both *K pneumoniae* and *E coli*. It would be attractive to extend the methodology to work with both long-read and short-read data in the future, as this would enable mobile genomic elements such as plasmids to be tracked. It is also possible that the thresholds used to validate the SC-level assignments by mSWEEP were too conservative, leading to missed assignments. However, we maintain that this step was key to ensure that incorrect assignments were not included.

The wealth of data generated by deep sequencing is an increasingly accessible way to explore bacterial variation in several ecological niches. Our study shows that this approach is also feasible for scrutiny of clinically relevant nosocomial bacteria, serving as a first step towards its further development and use as a research tool in clinical microbiology.

Contributors

JC, DS, SDB, and NRT conceptualised the study and obtained funding. JC, DS, and CJB administered the project. JC, DS, HAT, MC, SG, CJB, GT-H, TM, AKP, NM, RAG, SA-A, TK, DJ, SWL, CC, GAB, AH, ACS, RJLW, CM, GP, EJF, PC, NRT, and SDB obtained resources. HAT, MP, TM, DS, and JC developed the methodology. MC, SG, CM, GP, PC, and DS did laboratory investigations. HAT, MP, HP, GT-H, TM, SG, and JC did computational investigations. HAT, MC, SG, CJB, GT-H, TM, AKP, NM, RAG, SA-A, TK, DJ, SWL, CC, GAB, CM, GP, PC, DS, and JC curated the data. MP did the statistical analysis. HAT and JC did the bioinformatic analysis. HAT visualised the data. JC and GT-H accessed and verified the underlying data. HAT, JC, and DS wrote the manuscript. All authors reviewed and edited the manuscript, had full access to all the data in the study, and accept responsibility for the decision to submit for publication.

Declaration of interests

SWL received Robert Austrian Research Award sponsored by Pfizer, outside of the scope of this study. All other authors declare no competing interests.

Data sharing

Deep sequence data generated for this study are available at the ENA under project PRJEB39567. Sequence data of *E faecium* clade B (*Enterococcus lactis*) isolates are available at the ENA under project PRJEB28495.

Acknowledgments

This research was funded in part by the Wellcome Trust Grant number 206194. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Additional funding was obtained from the European Research Council (grant 742158, JC, HAT, HP), Academy of Finland Flagship programme (JC, AH), Trond Mohn Foundation (BATTALION grant, JC, AKP, RAG, NM), Research Council of Norway (grant 2999131 JC, GT-H). The authors thank Pasquale Piemontese, Debora De Vitis, Chiara Rebuffa, Vincenzo Brunco, Marco Ardizzone, and Alessia Girello, the laboratory technicians of the Microbiology and Virology Unit at Fondazione IRCCS Policlinico San Matteo (the study hospital) for helping with the microbiology work.

References

1 Suetens C, Latour K, Kärki T, et al. Prevalence of healthcare-associated infections, estimated incidence and composite antimicrobial resistance index in acute care hospitals and long-term care facilities: results from two European point prevalence surveys, 2016 to 2017. *Euro Surveill* 2018; **23:** 1800516.

2 Cassini A, Högberg LD, Plachouras D, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect Dis* 2019; **19:** 56–66.

3 Harris SR, Feil EJ, Holden MTG, et al. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 2010; **327:** 469–74.

4 Doughty EL, Liu H, Moran RA, et al. Endemicity and diversification of carbapenem-resistant *Acinetobacter baumannii* in an intensive care unit. *Lancet Reg Health West Pac* 2023; **37:** 100780.

5 Köser CU, Holden MTG, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012; **366:** 2267–75.

6 Tong SYC, Holden MTG, Nickerson EK, et al. Genome sequencing defines phylogeny and spread of methicillin-resistant *Staphylococcus aureus* in a high transmission setting. *Genome Res* 2015; **25:** 111–18.

7 Qi C, Hountras P, Pickens CO, et al. Detection of respiratory pathogens in clinical samples using metagenomic shotgun sequencing. *J Med Microbiol* 2019; **68:** 996–1002.

8 Rodino KG, Toledano M, Norgan AP, et al. Retrospective review of clinical utility of shotgun metagenomic sequencing testing of cerebrospinal fluid from a US tertiary care medical center. *J Clin Microbiol* 2020; **58:** e01729-20.

9 Oechslin CP, Lenz N, Liechti N, et al. Limited correlation of shotgun metagenomics following host depletion and routine diagnostics for viruses and bacteria in low concentrated surrogate and clinical samples. *Front Cell Infect Microbiol* 2018; **8:** 375.

10 Whelan FJ, Waddell B, Syed SA, et al. Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. *Nat Microbiol* 2020; **5:** 379–90.

11 Mäklin T, Kallonen T, David S, et al. High-resolution sweep metagenomics using fast probabilistic inference. *Wellcome Open Res* 2021; **5:** 14.

12 Mäklin T, Kallonen T, Alanko J, et al. Bacterial genomic epidemiology with mixed samples. *Microb Genom* 2021; **7:** 000691.

13 Mäklin T, Thorpe HA, Pöntinen AK, et al. Strong pathogen competition in neonatal gut colonisation. *Nat Commun* 2022; **13:** 7417.

14 Tonkin-Hill G, Ling C, Chaguza C, et al. Pneumococcal within-host diversity during colonization, transmission and treatment. *Nat Microbiol* 2022; **7:** 1791–804.

15 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014; **15:** R46.

16 Horesh G, Blackwell GA, Tonkin-Hill G, Corander J, Heinz E, Thomson NR. A comprehensive and high-quality collection of *Escherichia coli* genomes and their genes. *Microb Genom* 2021; **7:** 000499.

17 Arredondo-Alonso S, Top J, McNally A, et al. Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. *MBio* 2020; **11:** e03284-19.

18 Pöntinen AK, Top J, Arredondo-Alonso S, et al. Apparent nosocomial adaptation of *Enterococcus faecalis* predates the modern hospital era. *Nat Commun* 2021; **12:** 1523.

19 Thorpe HA, Booton R, Kallonen T, et al. A large-scale genomic snapshot of *Klebsiella* spp isolates in northern Italy reveals limited transmission between clinical and non-clinical settings. *Nat Microbiol* 2022; **7:** 2054–67.

20 Gladstone RA, Lo SW, Lees JA, et al. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine* 2019; **43:** 338–46.

21 Lees JA, Harris SR, Tonkin-Hill G, et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res* 2019; **29:** 304–16.

22 Alanko JN, Vuohtoniemi J, Mäklin T, Puglisi SJ. Themisto: a scalable colored *k*-mer index for sensitive pseudoalignment against hundreds of thousands of bacterial genomes. *Bioinformatics* 2023; **39:** i260–69.

23 Ondov BD, Treangen TJ, Melsted P, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016; **17:** 132.

24 Hunt M, Mather AE, Sánchez-Busó L, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genom* 2017; **3:** e000131.

25 Alcock BP, Huynh W, Chalil R, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 2023; **51:** D690–99.

26 Petazzoni G, Bellinzona G, Merla C, et al. The COVID-19 Pandemic sparked off a large-scale outbreak of carbapenem-resistant *Acinetobacter baumannii* from the endemic strains at an Italian hospital. *Microbiol Spectr* 2023; **11:** e0450522.

27 Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; **5:** e9490.

28 Hamidian M, Nigro SJ. Emergence, molecular mechanisms and global spread of carbapenem-resistant *Acinetobacter baumannii*. *Microb Genom* 2019; **5:** e000306.

29 Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev* 2014; **27:** 543–74.

30 David S, Reuter S, Harris SR, et al. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat Microbiol* 2019; **4:** 1919–29.

31 Wiehlmann L, Cramer N, Tümmler B. Habitat-associated skew of clone abundance in the *Pseudomonas aeruginosa* population. *Environ Microbiol Rep* 2015; **7:** 955–60.

32 Hsu L-Y, Harris SR, Chlebowicz MA, et al. Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol* 2015; **16:** 81.

33 Kinross P, Petersen A, Skov R, et al. Livestock-associated meticillin-resistant *Staphylococcus aureus* (MRSA) among human MRSA isolates, European Union/European Economic Area countries, 2013. *Euro Surveill* 2017; **22:** 16-00696.

34 Merla C, Kuka A, Petazzoni G, et al. Livestock-associated methicillin-resistant *Staphylococcus aureus* in inpatients: a snapshot from an Italian hospital. *J Glob Antimicrob Resist* 2022; **30:** 10–15.

35 Brodrick HJ, Raven KE, Kallonen T, et al. Longitudinal genomic surveillance of multidrug-resistant *Escherichia coli* carriage in a long-term care facility in the United Kingdom. *Genome Med* 2017; **9:** 70.

36 Mills EG, Martin MJ, Luo TL, et al. A one-year genomic investigation of *Escherichia coli* epidemiology and nosocomial spread at a large US healthcare network. *Genome Med* 2022; **14:** 147.

37 Roberts LW, Hoi LT, Khokhar FA, et al. Genomic characterisation of multidrug-resistant *Escherichia coli, Klebsiella pneumoniae*, and *Acinetobacter baumannii* in two intensive care units in Hanoi, Viet Nam: a prospective observational cohort study. *Lancet Microbe* 2022; **3:** e857–66.

38 Roer L, Overballe-Petersen S, Hansen F, et al. *Escherichia coli* sequence type 410 is causing new international high-risk clones. *MSphere* 2018; **3:** e00337-18.

Travel Medicine and Infectious Disease 59 (2024) 102698

Contents lists available at ScienceDirect

Travel Medicine and Infectious Disease

journal homepage: www.elsevier.com/locate/tmaid

Dynamics of viral DNA shedding and culture viral DNA positivity in different clinical samples collected during the 2022 mpox outbreak in Lombardy, Italy

Antonio Piralla [a,1], Davide Mileto [b,1], Alberto Rizzo [b], Guglielmo Ferrari [a], Federica Giardina [c], Stefano Gaiarsa [a], Greta Petazzoni [a], Micol Bianchi [b], Federica Salari [b], Fiorenza Bracchitta [b], Josè Camilla Sammartino [c], Alessandro Ferrari [a], Gloria Gagliardi [b], Alessandro Mancon [b], Claudio Fenizia [d], Mara Biasin [d], Francesca Rovida [c,a], Stefania Paolucci [a], Elena Percivalle [a], Alessandra Lombardi [b], Valeria Micheli [b], Silvia Nozza [e], Antonella Castagna [e], Davide Moschese [f], Spinello Antinori [g], Andrea Gori [h], Paolo Bonfanti [i], Roberto Rossotti [j], Antonella D'Arminio Monforte [k], Federica Attanasi [l], Marcello Tirani [m], Danilo Cereda [m], Fausto Baldanti [c,a,*], Maria Rita Gismondo [b,d], MPXV Lombardy Network

[a] *Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy*
[b] *Clinical Microbiology, Virology and Bioemergency Diagnostics, ASST Fatebenefratelli Sacco, Luigi Sacco Hospital, Milan, Italy*
[c] *Department of Clinical-Surgical, Diagnostic and Pediatric Sciences, Università degli Studi di Pavia, Pavia, Italy*
[d] *Department of Biomedical and Clinical Sciences "L.Sacco", University of Milan, Italy*
[e] *Clinic of Infectious Diseases, Vita-Salute University, San Raffaele Scientific Institute, Milan, Italy*
[f] *I Division of Infectious Diseases, Luigi Sacco Hospital, ASST Fatebenefratelli Sacco, Milan, Italy*
[g] *Department of Biomedical and Clinical Science, Università degli Studi di Milano, III Division of Infectious Diseases, ASST Fatebenefratelli Sacco, Milan, Italy*
[h] *Department of Infectious Diseases Ospedale "Luigi Sacco", Centre for Multidisciplinary Research in Health Science (MACH), University of Milan, Italy*
[i] *Department of Infectious Diseases, San Gerardo Hospital – University of Milano Bicocca, Monza, Italy*
[j] *Department of Infectious Diseases, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy*
[k] *Clinic of Infectious Diseases, Department of Health Sciences, ASST Santi Paolo E Carlo, University of Milan, Milan, Italy*
[l] *Agenzia per la Tutela della Salute della Brianza, Monza, Italy*
[m] *Directorate General for Health, Lombardy Region, Milan, Italy*

ARTICLE INFO

Keywords:
Mpox virus
Molecular epidemiology
Next generation sequencing
Re-Emerging virus
Multiple samples

ABSTRACT

Background: Mpox virus (MPXV) has recently spread outside of sub-Saharan Africa. This large multicentre study was conducted in Lombardy, the most densely populated Italian region accounting for more than 40% of Italian cases. The present study aims to: i) evaluate the presence and the shedding duration of MPXV DNA in different body compartments correlating the MPXV viability with the time to onset of symptoms; ii) provide evidence of MPXV persistence in different body compartment as a source of infection and iii) characterize the MPXV evolution by whole genome sequencing (WGS) during the outbreak occurred in Italy.

Material and methods: The study included 353 patients with a laboratory-confirmed diagnosis of MPXV infection screened in several clinical specimens in the period May 24th - September 1st, 2022. Viral isolation was attempted from different biological matrices and complete genome sequencing was performed for 61 MPXV strains.

Results: MPXV DNA detection was more frequent in the skin (94.4%) with the longest median time of viral clearance (16 days). The actively-replicating virus in cell culture was obtained for 123/377 (32.6%) samples with a significant higher viral quantity on isolation positive samples (20 vs 31, $p < 0.001$). The phylogenetic analysis highlighted the high genetic identity of the MPXV strains collected, both globally and within the Lombardy region.

* Corresponding author. Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, 27100, Pavia, Italy.
E-mail address: f.baldanti@smatteo.pv.it (F. Baldanti).
[1] These authors have equally contributed.

https://doi.org/10.1016/j.tmaid.2024.102698
Received 13 December 2023; Received in revised form 15 February 2024; Accepted 17 February 2024
Available online 29 March 2024
1477-8939/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Conclusion: Skin lesion is gold standard material and the high viral load and the actively-replicating virus observed in genital sites confirms that sexual contact plays a key role in the viral transmission.

1. Introduction

Since May 2022, *Mpox virus* (MPXV) has been responsible for a global outbreak, the first related to *Orthopoxvirus* (OPXV) after *Smallpox virus* eradication, with more than 86,746 cases and 112 fatalities worldwide, the majority of which in countries where MPXV was never reported before [1]. Until April 12th, 2023, European Union (EU) accounted for 25,874 cases, with a hospitalization rate of 6% [2]. The 2022 MPXV outbreak mainly involves men who have sex with men (MSM), displaying characteristics of a sexually transmitted infection (STI) [3–7]. Mpox virus is a linear double-stranded DNA zoonotic virus belonging to the *Orthopoxvirus* genus *Poxviridae* family, to which the WHO has attributed an epidemic, or even a pandemic potential. Two distinct genetic clades were described: the central African (Congo Basin, Clade I), which causes a more severe disease, and the west African clade (Clade II) causing a milder disease and associated with the 2022 outbreak (named Clade IIb) [3,8,9]. Human-to-human transmission occurs through direct contact with skin lesions and infectious body fluids or from indirect contact with contaminated materials, such as clothing or linens. Furthermore, prolonged face-to-face contact is necessary for human-to-human transmission through large respiratory droplets, sneezing, or coughing [10]. Incubation period lasts 3–17 days, followed by a two-stage disease of 2–4 weeks. In detail, the invasive stage (0–5 days) is characterized by fever, lymphadenopathy, headache, myalgia, and fatigue, while in the second stage, distinctive skin lesions occur, and their number is descriptive of disease severity [10,11]. Severe outcomes are usually rare [12], and the overall case fatality has been assessed up to 3–6%, depending on health conditions, and access to therapies [2,12]. Immune deficiencies or coinfection may lead to a faster progression to a worse clinical picture [13,14]. The recent emergency has stepped up the research on MPXV pathogenesis, transmission, diagnosis, and management, which, according to recent bibliometric analyses, is still too limited [13,14]. New data on route of transmission, the extent of asymptomatic infection, and correspondence between PCR cycle threshold value and, infectivity are recently reported by several studies in France, Spain and Australia [15–17]. Lombardy, the most densely populated Italian region (10 million inhabitants) accounted for more than 40% of national cases [18].

In this epidemiological scenario, the present study described clinical and virological data collected during the monitoring of 353 MPXV-positive subjects. Clinical specimens from different anatomical sites (oropharyngeal, anal, urethral, and skin lesion swabs, together with plasma, urine, and semen) were collected, and analysed to i) evaluate the presence of MPXV DNA in different body compartments; ii) measure the shedding duration of MPXV DNA; iii) correlate the MPXV viability with the time to onset of symptoms iv) provide evidence of MPXV persistence in different body compartment as a source of infection and v) finally, characterize the MPXV evolution by whole genome sequencing (WGS) during the outbreak occurred in Italy.

2. Material and Methods

2.1. Study population

A total of 793 subjects were tested during the Regional Surveillance Program using molecular assays targeting *Orthopoxvirus* and MPXV-specific real-time PCR as confirmation in the period May 24th - September 1st, 2022. The study included 353 individuals with a laboratory-confirmed diagnosis of MPXV infection. Infection was defined as the detection of viral DNA in at least one of tested biological specimens. Demographical, clinical, and epidemiological data were collected at first access and during the follow-up period and reported in a Regional shared database. Follow-up samples were available only for a series of patients included in the present study and described below.

All diagnostic and experimental procedures were performed in the two Regional Mpox virus References Centres: i) Laboratory of Clinical Microbiology, Virology, and Bioemergencies, "L. Sacco" University Hospital (Milan) and ii) Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo (Pavia). This study complies with the Declaration of Helsinki. The institutional review board approved the study and the use of residual clinical specimens for complete genome characterization of both institutions (no.2022/ST/124 and no.44007/2022).

2.2. Molecular diagnosis

The following biological matrices were collected and analysed for diagnostic and follow-up monitoring purpose: swabs from the oropharynx, skin lesion, anus, and urethra in Universal Transport Medium swabs (UTM-RT®; COPAN Diagnostics, Italy), blood, urine, and seminal fluid samples. DNA was extracted with QIAsymphony® DSP Virus/Pathogen Kit on QIAsymphony® SP automated platform (QIAGEN, Germany). OPXV screening was performed by means of RealStar® *Orthopoxvirus* PCR Kit 1.0 (Altona DIAGNOSTICS) real-time PCR, targeting variola virus and non-variola OPXV species (*Cowpox virus*, *Mpox virus*, *Raccoonpox virus*, *Camelpox virus*, *Vaccinia virus*). The presence of MPXV DNA was confirmed using a specific homemade real-time PCR protocol as previously described [19,20]. Results were given as cycle of quantification (Cq) values that inversely reflects viral load. A Cq $\geq$ 40 was set as negative cut-off.

2.3. Viral isolation

A total of 377 MPXV-positive samples were used to attempt viral isolation. An aliquot of 200 μL of each transport medium was plated in duplicate in 24-well plates containing 80–90% confluent Vero E6 cells, adding 800 μL of Dulbecco's Modified Eagle Medium with L-glutamine (Gibco ThermoFisher Scientific) supplemented with 2% of heat-inactivated fetal bovine serum (Gibco ThermoFisher Scientific) and 1% penicillin-streptomycin (5000 U/mL; Pen-Strep, Gibco ThermoFisher Scientific). Plates were incubated at 37 °C and at 5% CO2 atmospheric pressure and checked every 24 h. Wells were monitored daily for virus-induced cytopathic effect (CPE) showing typical monolayer separation and cell rounding and CPE readings were recorded by two independent readers for each sample. In the majority of samples with viable virus, extensive CPE was observed between 2 and 6 days.

2.4. Viral sequencings

In a series of cultured-positive samples, (nearly 10% of total) WGS was performers. In detail, a Nextera XT paired-end library (Illumina) was prepared using 1 ng of DNA extracted from culture supernatants, using a QIAamp DNA mini kit (Qiagen). The library was sequenced on the MiSeq platform using paired-end sequencing, with a read length of 150 nucleotides. Sequencing reads were mapped to a collection of high-quality human Mpox virus genomes. Mapping reads were assembled using Spades [21]. The sequence of all 179 viral genes (reference: NC_063383) was retrieved from the 61 genomes obtained in this study and from a collection of high-quality genomes obtained from GISAID (https://gisaid.org/, N = 523). Genes were aligned, concatenated, and

used to infer a maximum-likelihood phylogeny with the software IQ-TREE [22]. Gene alignments were further analysed to obtain all genic mutations and their nucleotide-substitution pattern. Lastly, the association between all mutations and the topology of the phylogenetic tree was tested using Cramer's V index.

2.5. Genome assembly

All Sequencing reads pairs (N = 61) were quality-filtered using Trimmomatic1, which was also used to trim sequencing primers and low-quality bases at both ends of all reads. High-quality reads were mapped to a collection of 121 MPXV (available in the NCBI Assembly database on May 31, 2022) using Bowtie22. Mapping reads were assembled using SPAdes3 in "careful" mode.

2.6. Phylogenetic analyses

All high-coverage complete genomes with a full collection date available and relative to viral samples collected after January 1, 2018 were retrieved from the GISAID database on October 17, 2022 (gisaid.org; N = 618). The Coding Sequences (CDSs) of genome NC_063383 (N = 179) were retrieved from the NCBI repository and blast-searched on both the 61 genomes of the study and the database genomes. Database genomes from which it was not possible to retrieve all CDSs were excluded from the study. The final dataset included the CDSs of 585 genomes (including 61 study genomes, 523 database genomes, and the reference). All CDSs were aligned using MAFFT4 and concatenated. The resulting alignment of 166090 bp was used as input IQ-TREE5 to infer the phylogeny, using the TN + F + I + I + R6 substitution model (chosen according to BIC within the IQ-TREE internal pipeline).

2.7. Mutation analysis

All CDS alignments were scanned triplet by triplet for mutations from the reference using an in-house python script. Each mutation was classified as synonymous or non-synonymous and the nucleotide substitution pattern was extracted. Lastly, the occurrences of each mutation were counted in the genomes inside and outside the phylogenetic cluster including the 2022 outbreak. The Association of each mutation with the 2022 outbreak cluster was assessed using the Cramer's V index, after the removal of all occurrences of non-standard nucleotides (e.g. "N″s).

2.8. Statistical analysis

Comparisons were calculated using $\chi 2$ or Fisher exact for categorical variables and Kruskal-Wallis tests for continuous variables, since they were not normally distributed. Non-parametric survival analysis presenting the Kaplan-Meier curve was performed to assess the persistence of MPXV-DNA in different clinical samples and curves were compared using the Long Rank test. Spearman's correlation coefficient was used to evaluate the association with Cq and time of onset symptoms. Differences were considered statistically significant at $p < 0.05$ for all tests. All statistical analyses were performed using GraphPad software version 8.3.0 (Prism).

3. Results

3.1. Study population and clinical characteristics

A total of 793 patients were screened for MPXV infection and 353 (44.5%) were confirmed by real-time PCR. A total of 4018 clinical specimens were analysed for diagnosis as well as during follow-up. Of these, 1191 (29.6%) were oropharyngeal swabs, 1166 (29.0%) were vesicular or pustular swab samples, 639 (15.9%) were anogenital swabs,

Table 1
Demographic, clinical characteristics and disease severity of confirmed mpox cases in Lombardy (n = 353).

Categories		overall		vaccinated		unvaccinated		unknown		p value[a]
		No.	%	No.	%	No.	%	No.	%	
Total cases		353	100.0	30	8.5	231	65.4	92	26.1	
Median Age yrs (range)		37 (15–67)		50 (30–58)		37 (15–6)		37 (19–67)		**<0.001**
Age group	0–19	3	0.8	0	0.0	2	0.9	1	1.1	**<0.001**
	20–29	55	15.5	0	0.0	43	18.6	12	13.0	
	30–39	164	46.6	3	10.0	115	49.8	46	50.0	
	40–49	97	27.4	9	30.0	64	27.7	24	26.1	
	50–59	29	8.2	18	60.0	6	2.6	5	5.4	
	>60	5	1.4	0	0.0	1	0.4	4	4.3	
Gender	Male	350	99.2	29	96.7	229	99.1	92	100.0	>0.99
	Female	3	0.8	1	3.3	2	0.9	0	0.0	
HIV status	Positive	37	10.5	3	10.0	22	9.5	12	13.0	>0.99
	Negative	18	5.1	1	3.3	14	6.1	3	3.3	
	Unknown	300	84.7	26	86.7	195	84.4	77	83.7	
Other STIs	Yes	7	2.0	3	10.0	6	2.6	0	0.0	0.08
	No	32	9.0	1	3.3	18	7.8	2	2.2	
	Unknown	315	89.0	26	86.7	207	89.6	90	97.8	
Rash	Yes	331	93.5	27	90.0	218	94.4	85	92.4	0.14
	No	14	4.0	3	10.0	9	3.9	2	2.2	
	Unknown	9	2.5	0	0.0	4	1.7	5	5.4	
Lymphadenopathy	Yes	154	43.5	10	33.3	109	47.2	35	38.0	0.23
	No	190	53.7	19	63.3	118	51.1	52	56.5	
	Unknown	10	2.8	1	3.3	4	1.7	5	5.4	
Fever	Yes	193	54.5	11	36.7	135	58.4	47	51.1	**0.02**
	No	149	42.1	18	60.0	91	39.4	39	42.4	
	Unknown	12	3.4	1	3.3	5	2.2	6	6.5	
Hospitalization	Yes	11	3.1	0	0.0	7	3.0	4	4.3	>0.99
	No	342	96.6	30	100.0	224	97.0	87	94.6	
	Unknown	1	0.3	0	0.0	0	0.0	1	1.1	
Local transmission case	Yes	247	69.8	21	70.0	160	69.3	60	65.2	>0.99
	No	86	24.3	8	26.7	60	26.0	20	21.7	
	Unknown	21	5.9	1	3.3	11	4.8	12	13.0	

[a]p value is referred to comparison between vaccinated and unvaccinated MPXV cases.

431 (10.7%) were blood samples, 251 (6.2%) were urine, 176 (4.4%) were urethral swabs, and 164 (4.1%) were semen. As summarized in Table 1, the median age of MPXV-confirmed cases was 37.0 years (IQR, 32–43 years: range 15–67 years) and the great majority were males (350/353, 99.2%). A minority of patients (11/353; 3.1%) required hospitalization, while the vast majority (342, 96.6%) were managed as outpatients. For one patient no information was available. Clinical reports were available for 345 (97.5%) confirmed cases with cutaneous rash and lymphadenopathy present in 331 (93.5%) and 154 (43.5%) subjects, respectively while 193 (54.5%) complained of fever (Table 1). A total of 329 (93.2%) exposure histories were available and 244 (74.1%) of them were autochthonous cases while transmission likely occurred abroad in 85 (25.8%) cases (mainly in Spain, France Germany, and Great Britain). Smallpox vaccination status was reported in 261 cases, showing that more than half of MPXV-positive patients were unvaccinated (231, 65.4%). Considering that smallpox vaccination was waived in 1980, it was expected that immunized subjects were older than those unvaccinated (median age 50 *vs* 36 years, $p < 0.01$). A great majority of vaccinated patients belonged to 50–59 age group (18/30; 60.0%), while in unvaccinated patients the most represented age group was 30–39 (115/231; 49.8%) (Table 1).

3.2. MPXV load and persistence in clinical specimens

Overall, 1285 samples were collected at the time of diagnosis, with a median time from the onset of symptoms of 6 days (IQR 3–9). MPXV detection was more frequent from the skin (289 of 306, 94.4%), anogenital (155 of 188, 82.4%), oropharyngeal (266 of 345, 78.0%) and plasma/blood (133 of 186, 71.5%) samples, than from urethral (45 of 69, 65.2%), semen (37 of 77, 48.1%) and urine (23 of 114, 20.2%) samples. The MPXV load, inversely reflected by Cq values, was significantly higher from skin lesions (median Cq 20, IQR 17–26) than from anogenital samples (median Cq 24, IQR 19–33), and oropharynx (median Cq 28, IQR 25–23) (Fig. 1A). In the remaining urethral (median Cq 31, IQR 25.5–35), urine (median Cq 32, IQR 25–34), semen (median Cq 34, IQR 30.5–36), and blood (median Cq 34, IQR 32–35) samples, the median viral load was higher than Cq = 30 and significantly lower than detected from the skin, anogenital and oropharyngeal samples (Fig. 1A, $p < 0.001$).

In 302 patients, skin and oropharyngeal samples were simultaneously collected at the median of 6 days from the onset of symptoms (IQR 3–9 days). Among these, in 229 paired samples (both positive), a significant difference in the median MPXV load was observed (median Cq 19, IQR 17–24 *vs* median Cq 28, IQR 24–32; $p < 0.001$). In 56 and 13 paired samples, only skin (median Cq 23 IQR 18.3–32) and oropharyngeal (median Cq 34 IQR 30–35.5) samples were positive, respectively. Finally, in four paired samples MPXV DNA was detected in neither skin nor oropharynx samples and the diagnosis was performed on anogenital samples.

Follow-up samples were collected only in a subset of patients and the MPXV DNA clearance was investigated in the different body compartments using the Kaplan-Meier method (Fig. 1B and C). The median clearance time of MPXV DNA detection was 16 days in the skin (n = 110 patients), 14 days in the oropharynx (n = 167), 13 days in anogenital (n = 116) and urethral samples (n = 30), 9 days in urine (n = 14), 8 days in blood (n = 80) and finally 7 days in semen (n = 24) (Fig. 1B). Persistent shedding defined as duration >21 days was observed in 10.0% (11/110) of skin samples, 6.1% (10/165) of oropharyngeal swabs, and 3.4% (4/116) of anogenital samples. The most prolonged MPXV DNA shedding was observed in one oropharyngeal sample still positive at 56 days. Kaplan-Meier curves of sample type with a more prolonged viral shedding (skin and oropharyngeal samples) when compared with the log-rank test resulted significantly different (median time, 16 *vs* 14 days, HR, 0.56; 95% CI, 0.41 to 0.77; $p < 0.001$, Fig. 1C).

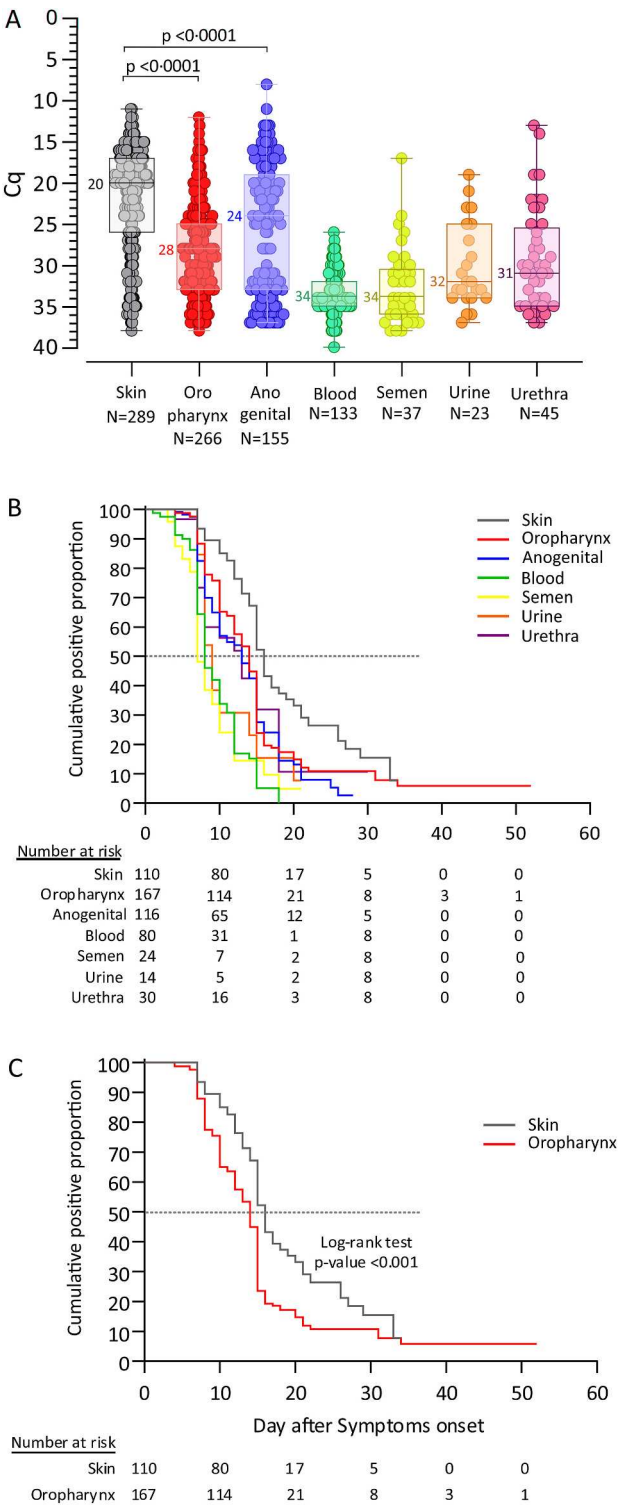


Fig. 1. The cycle of quantification values and viral shedding duration. (A) Cycle of quantification in different clinical specimens (B) Cumulative positive proportion in different clinical specimens. (C) Cumulative positive proportion in the clinical specimens (skin and oropharynx) with the high median MPXV DNA persistence (16 *vs* 14 days). Curves comparison was performed using log rank test analysis.

3.3. Isolation vs cq

MPXV isolation was attempted for 377 samples as summarized in Table 2. An actively-replicating virus was demonstrated for a total of

Table 2
Results of viral isolation assay performed on different MPXV DNA positive biological samples.

Clinical sample	Total (n = 377)		Positive (n = 123)		Negative (n = 248)		Cross-contaminated (n = 6)	
	no.	%[a]	no.	%[b]	no.	%[b]	no.	%
skin	120	31.8	39	32.5	81	67.5	0	0.0
nasopharynx	117	31.0	12	10.3	105	89.7	0	0.0
anogenital	71	18.8	46	64.8	19	26.8	6	8.5
plasma/blood	25	6.6	0	0.0	25	100.0	0	0.0
semen	3	0.8	3	100.0	0	0.0	0	0.0
urine	4	1.1	3	75.0	1	25.0	0	0.0
urethra	37	9.8	20	54.1	17	45.9	0	0.0

[a] calculated based on total samples.
[b] calculated based on total for each sample categories'.

123/377 (32.6%) samples, while it was unsuccessful in 248 and the remaining 6 samples were excluded due to bacterial or fungal contamination. Stratifying isolation results according to sample type, the rate of isolation was as follows: 100% for semen (3/3), 64.8% (46/71) for anogenital swabs, 10.3% (12/117) for oropharynx samples and 0% (0/25) for blood samples (Table 2).

In 359 samples the time between the date of sample collection and the symptoms onset was available and thus results of MPXV isolation were stratified into five categories: 0–7, 8–14, 15–21, 22–28, ≥29 days (Fig. 2A). Higher isolation rate was observed in samples collected between 0 and 7 days (65/158, 41.1%) from the onset of symptoms. The proportion decreased to 31.5% (34/108) for samples collected between 8 and 14 days and below 20% in the other categories (Fig. 2A). Analysing the rate of isolation among each sample category, in skin samples, the rate was over 40.0% for 0-7- and 8-14-day samples and strongly decreased in the 15–21, 22–28, and ≥29 days categories (Fig. 2B). Among oropharyngeal samples, the rate of isolation was less than 20 % in all categories with a maximum rate in samples collected at 0–7 days

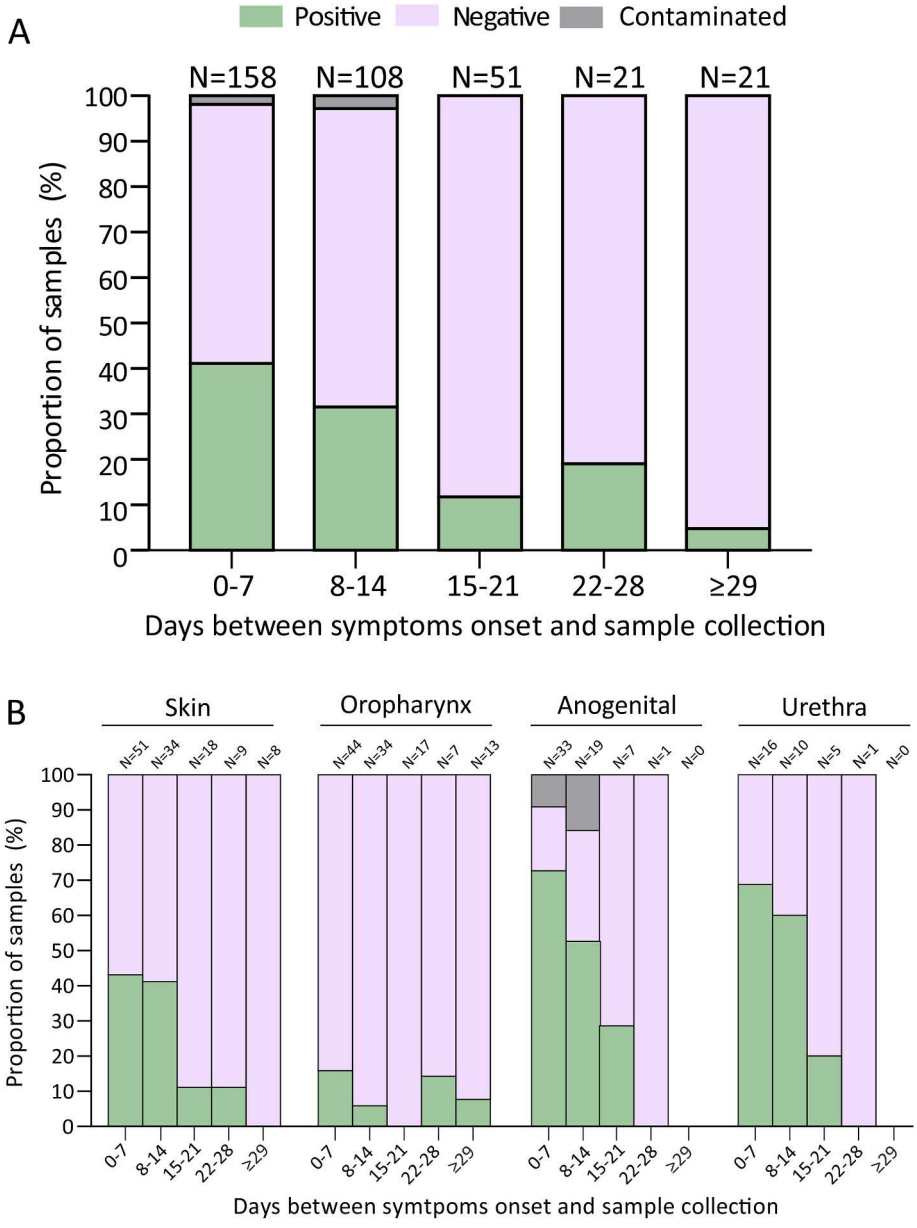


Fig. 2. Proportion of viral culture isolation. (A) Overall proportion of MPXV isolation results stratified by days between symptoms onset and samples collection. (B) Proportion of MPXV isolation results stratified by days between symptoms onset and samples collection according to the clinical sample with at least 30 samples.

from the symptoms onset. Among anogenital samples, the rate of isolation was 72.7% in the 0-7-days, 52.6% in the 8-14-days, and declined to 28.6% in the 15–21-days sample category. Similar findings were observed in the few urethral samples analysed. Overall, the isolation rate was higher than 60.0% for the 0-7- and 8-14-days categories with rates of 68.8% and 60.0%, respectively. For almost all samples (360/377, 95.5%) in which isolation was attempted, Cq values were available. Thus, a correlation between Cq value and isolation success, as well as timing from symptoms onset was assessed (Fig. 3). The median Cq values (Fig. 3A) between isolated and non-isolated samples were significantly different in the 0-7-days (19 vs 30; $p < 0.001$) as well as in 8–14-days (22 vs 29; $p < 0.001$) categories. As for the other categories, including samples collected ≥15 days after the onset of symptoms. Overall, the median Cq of samples with MPXV isolation was lower than other ones (20 vs 31, $p < 0.001$, Fig. 3B). Comparison of Cq values according to isolation success/failure showed a significant difference in the following biological specimens: skin (17 vs 28; $p < 0.001$), oropharyngeal (21 vs 31; $p = 0.0016$), anogenital (20 vs 35, $p < 0.001$) and urethral (28.5 vs 34.0, $p < 0.001$) samples. No comparison on Cq values could be done on plasma/blood, semen and urine samples.

3.4. Genomics and phylogenetic characterization

Sixty-one samples were cultured for virus isolation and their full genome was obtained by MPXV read selection and subsequent de novo assembly. The mean read-depth of the 61 genomes obtained was 118.2, while the mean N50 was 153074 and the mean contig number was 8.1. The 61 novel genomes were included in a Maximum-Likelihood phylogeny, together with 523 genomes available in public database (https://gisaid.org/, Fig. 4). The tree highlights the presence of a large monophylum including all the 2022 outbreak isolates. In addition, bootstrap values do not support the base nodes topology of the outbreak cluster, impeding the identification of outbreak sub-lineages. This result highlights the high genetic identity of the MPXV strains collected, both globally and within the Lombardy region. Nevertheless, it is possible to identify small high-confidence clusters in the most recent nodes of the tree, which are supported by a high bootstrap value and could be used to hypothesize small-scale epidemiological links. The 61 genomes

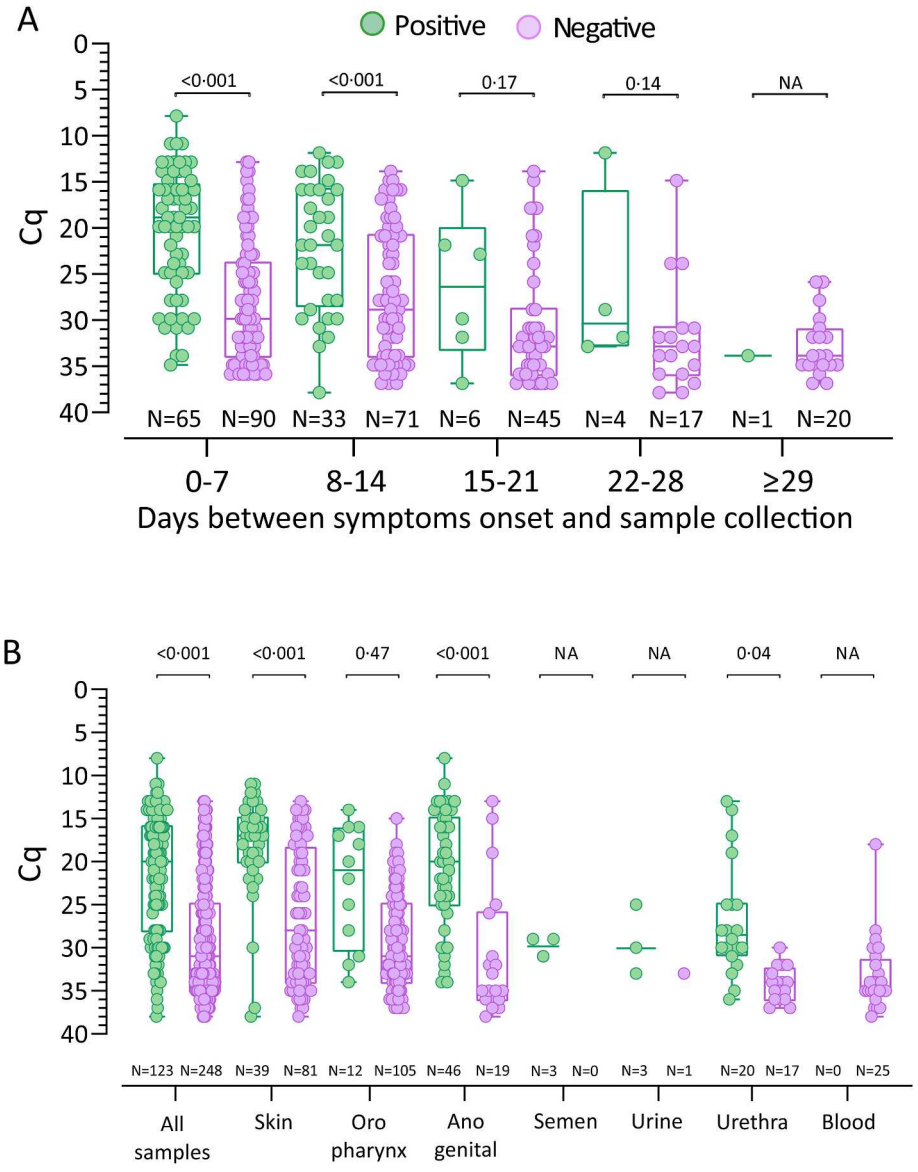


Fig. 3. The cycle of quantification values and viral culture isolation. (A) Comparison of Cq in virus culture positive and negative samples stratified by days between symptoms onset and samples collection. (B) Comparison of Cq in virus culture positive and negative samples stratified by different clinical specimens collected. NA, not applicable.

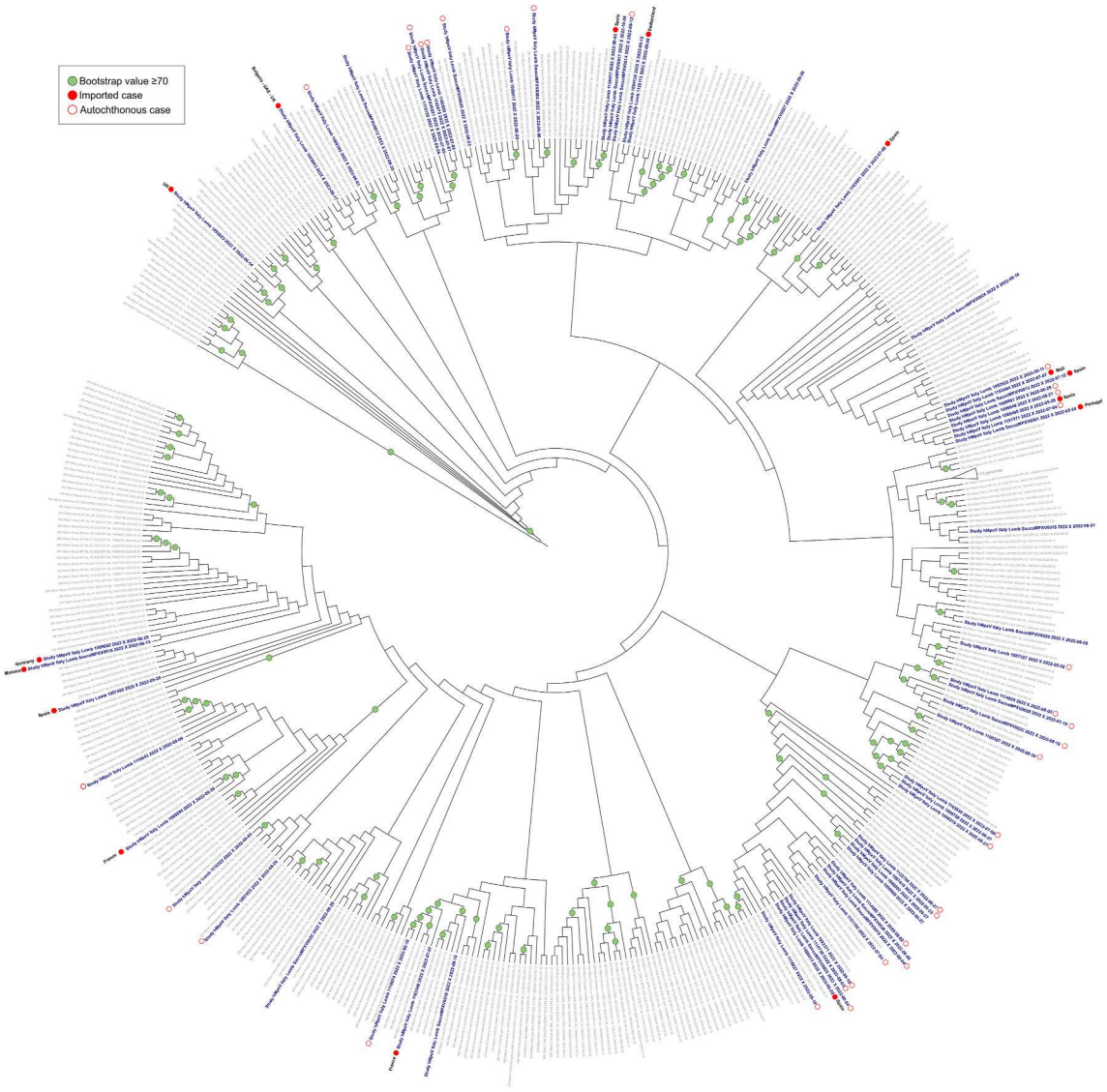


Fig. 4. Phylogenetic three of complete coding sequences of MPXV strains originated in this study (n = 61) and references (n = 523). Nodes of tree with a bootstrap value ≥ 70 were highlighted with a green dot. In order to differentiate the MPXV cases, autochthonous cases are highlighted with a white circle (with red border) while imported cases are highlighted with a red circle. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

described in this study are distributed in 37 local subclusters of the outbreak clade. Two clusters contain genomes relative to both imported infections and to autochthonous ones. E.g., sample 1096013 is associated with a patient that was infected in Spain and clusters with three other MPXV genomes associated with local infections. Both the tree topology and the sampling date suggest that the virus was imported by the patient who had travelled, and it was spread to the other three subjects.

After running association tests, 44 genic mutations were found linked to the 2022 outbreak clade (Table S1) of which 24/44 were non-synonymous. Interestingly, 42 characterizing mutations were also identified by Isidro et al. and by Wang et al., 40 of which were common to both studies [8,23]. Moreover, the two aforementioned works also identified intergenic mutations, which were not the focus of the present study. Notably, the vast majority of the single nucleotide mutations detected (617/757; 81.5%) followed the G > A or the C > T mutation pattern, which is known to be associated to the APOBEC enzymes' function. The prevalence of these substitutions is even higher in these outbreak-associated mutations (43/44; 97.7%).

4. Discussion

The recent MPXV global outbreak represented a major public health concern, requiring a strong effort in terms of individuals' management, spread containment, and diagnostic resources. However, it also gave the opportunity to elucidate the clinical and virological aspects of a neglected tropical disease. Our results confirm that viral DNA detected by qPCR on swabs from lesions is the most appropriate standard tool to make a prompt MPXV laboratory diagnosis [24]. Taken alongside data from previous studies [4,15–17,25–27], swabs from skin lesions were PCR positive at the time of diagnosis in almost all samples, showed the highest viral load (low cq value), the longer viral shedding (median 16 days after symptoms onset) and an isolation rate of nearly 40%. Suner et al. reported a median time of 25 days for DNA detection by qPCR in skin lesions but a shorter period of DNA detection for other body fluids as also observed in our study [26]. On the contrary, the isolation rate from the oropharyngeal swabs (n = 117), which proved to be an optimal biological sample at diagnosis (median shedding 14 days), was below 10%. This finding suggests the persistence of MPXV DNA in the upper respiratory tract that was not associated with a viable virus. Our results,

are slightly different from those reported by Hernaez in a Spanish cross-sectional study involving respiratory samples from 44 patients describing viable virus in 66% of qPCR-positive saliva samples [25]. The reduced infectivity in aerosol samples in our study may reflect the efficiency of virus replication our culture system used, but conversely, our results on other body fluids (e.g. skin lesion, anogenital) are in keeping with other reports [16]. Higher isolation rates (more than 50%) were observed in anogenital and urethral samples with higher isolation rate in the first 14 days after symptoms onset, supporting the possibility of a sexual transmission of the disease. Conversely, MPXV was not isolated from any of the 25 positive plasma samples. The associated high Cq values (low viral load) assessed in these samples likely justifies the lack of viral isolation.

In our study, vaccinated subjects showed milder symptoms (no skin lesions), faster viral clearance and lower Cq values in multiple biological specimens. Indeed, most of the recruited patients were below 50 years of age and therefore unvaccinated. However, based on our data no major conclusions could be draw on the implication of MPXv vaccine discontinuation in the origin and spread of 2022 MPXV outbreak. Moreover, independently from vaccination, none of the individuals included in this study showed severe complications and hospitalization was necessary only in 3.1% of cases, mostly at the beginning of the outbreak due to a worrying symptomatology, pain management and the need for antiviral treatment. Indeed, our hospitalization rate was slightly lower than those observed in a recent systemic review reporting however a high level of heterogeneity worldwide [28].

Since the beginning of this outbreak, most cases in Europe have been registered among MSM [2], and among those who have multiple sex partners [29]. This epidemiologic feature was also observed in the population included in our study. Indeed, sexual activity entails close physical contacts, favouring the chance of transmission, irrespective of sexual orientation and route of infection, as furtherly supported by the three MPXV-positive females, who reported sexual intercourse with a confirmed positive partner. Yet, whether MPXV is transmitted through sexual secretions and/or oral, genital, or anal mucosa to date remains under investigation. So far, viral DNA in seminal fluid has been detected in four and 22 patients in Italy [30,31], two cases in Germany [32], and in 29/32 (91%) patients belonging to a larger case series [3]. In addition, viable virus has been recently documented few other reports [33,34] also summarized in a recent systematic review by Reda et al. [35]. In our series, replication-competent virus was isolated in 100% (3/3) seminal specimens supporting the evidence of infectiousness of MPXV in semen. Virus isolation confirms semen as a potential source of infection, but additional analyses are required to assess whether the virus could be associated with seminal cell infection, whether it stems from passive diffusion from urethra, or genital lesions, and whether viral replication occurs in the genital tract.

How the infective virus is conveyed from the entry site to the several infected anatomical sites remains to be elucidated. Based on this observation, it emerged that the first two weeks after symptoms onset likely represent the most important phase in terms of virus infectiousness. Moreover, our data suggest that, on average, during the fourth week from symptoms onset the infectious potential might be considered drastically reduced.

The whole sequencing data consistently detected the IIb clade (previously named "West African"), mirroring the other reports on the 2022 outbreak. Genomic diversity among the outbreak samples is low and to date there is no evidence of emerging variants of concern suggestive of immune escape [8]. Phylogeny of previously published and novel genomes highlighted the scarce genomic variability of the outbreak samples, while allowing the identification of small high-confidence local epidemiological clusters (Fig. 4). Single mutation analysis led us to identify 24 non-synonymous mutations that characterize the 2022 outbreak clade (Table S1). Of note, three of these mutations (L108F in the DNA polymerase, S30L and D88 N in the Late transcription factor VLTF-1) were indicated among the putative causes of the 2022 outbreak by Kannan et al. [36]. In this study, the authors observe that the mutations contributing to the enhanced viral spread are related to the replication process. Indeed, a large part of the non-synonymous mutations that we have identified as outbreak-related affect proteins that regulate the genome replication (e.g. gene_124 DNA Helicale, gene_50 DNA Polymerase). Moreover, two other mutations affect proteins that regulate the interaction with the host immune system: S105L in chemokine binding protein [37] and S54F in Crm-B TNF-alpha-receptor-like protein [38] (both are in the inverted repeat regions and are thus listed twice in Table S1). The two aminoacidic changes might have improved the protein affinity with the human TNF and chemokines. The vast majority of all the mutations detected in our dataset follow the substitution pattern associated with the APOBEC deaminases. As pointed out by Isidro et al. [8] and Gigante et al. [39] as well such enzymes can be considered the evolutionary driving force of MPXV, which led to the development of the 2022 epidemic. In fact, 43 of the 44 outbreak-associated mutations follow these patterns.

Our study has some limitations. First, follow-up samples could not be collected from all patients and not for all clinical samples, thus viral shedding duration has been assessed only in subset of patients. The rapid evolving of outbreak has reduced the capacity to collected information regarding on skin healing (e.g. fresh pustules or desquamation of crust), thus the onset of symptoms was used as starting point to calculate the starting point of infection.

5. Conclusions

Our study describes clinical and epidemiological overview of the 2022 MPXV outbreak, including a clinical track record to real-time PCR performed on specimens from different anatomical sites (oropharyngeal, anal, urethral, and skin lesion swabs, together with plasma, urine and sperm) as well as MPXV isolation in an "*in vitro*" model. Our study will contribute to a better understanding of the MPXV dissemination, with a particular focus on the different route of transmissions. The provided insight on MPXV 2022 outbreak has highlighted the need to refine the clinical management and diagnosis, and for defining appropriate public health guidelines and preventive strategies, suited to the most affected communities.

Funding statement

This research was supported by EU funding within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT).

Data sharing

All sequenxing data included in this study are available under BioProject ID: PRJNA909472 (http://www.ncbi.nlm.nih.gov/bioproject/909472). De-identified participant data collected for the study, including individual participant data will be made available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Antonio Piralla: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Davide Mileto:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Alberto Rizzo:** Methodology, Formal analysis, Data curation. **Guglielmo Ferrari:** Methodology, Formal analysis, Data curation. **Federica Giardina:** Methodology, Formal analysis, Data curation. **Stefano Gaiarsa:** Writing – original draft, Visualization, Software, Data curation. **Greta Petazzoni:** Writing – original draft,

Visualization, Software, Data curation. **Micol Bianchi:** Investigation, Formal analysis, Data curation. **Federica Salari:** Methodology, Formal analysis, Data curation. **Fiorenza Bracchitta:** Methodology, Formal analysis, Data curation. **Josè Camilla Sammartino:** Methodology, Formal analysis, Data curation. **Alessandro Ferrari:** Methodology, Formal analysis, Data curation. **Gloria Gagliardi:** Methodology, Formal analysis, Data curation. **Alessandro Mancon:** Methodology, Formal analysis, Data curation. **Claudio Fenizia:** Methodology, Formal analysis, Data curation. **Mara Biasin:** Methodology, Formal analysis, Data curation. **Francesca Rovida:** Formal analysis, Data curation. **Stefania Paolucci:** Formal analysis, Data curation. **Elena Percivalle:** Formal analysis, Data curation. **Alessandra Lombardi:** Formal analysis, Data curation. **Valeria Micheli:** Formal analysis, Data curation. **Silvia Nozza:** Data curation. **Antonella Castagna:** Data curation. **Davide Moschese:** Data curation. **Spinello Antinori:** Data curation. **Andrea Gori:** Data curation. **Paolo Bonfanti:** Data curation. **Roberto Rossotti:** Data curation. **Antonella D'Arminio Monforte:** Data curation. **Federica Attanasi:** Data curation. **Marcello Tirani:** Data curation. **Danilo Cereda:** Data curation. **Fausto Baldanti:** Writing – review & editing, Supervision, Funding acquisition. **Maria Rita Gismondo:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We thank Daniela Sartori for manuscript editing. We thank all health care workers involved in the individual's management in inpatients as well as outpatient's department.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tmaid.2024.102698.

References

[1] CDC. Monkeypox in the U.S. Centers for Disease Control and Prevention; 2022. published online Aug 31, https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html. [Accessed 12 April 2023].
[2] Monkeypox situation update, as of 1 March 2023. European Centre for Disease Prevention and Control; 2023. published online Mar 01, https://monkeypoxreport.ecdc.europa.eu/. [Accessed 12 March 2023].
[3] Thornhill JP, Barkati S, Walmsley S, Rockstroh J, Antinori A, Harrison LB, et al. Monkeypox virus infection in humans across 16 countries - april-june 2022. N Engl J Med 2022;387:679–91. https://doi.org/10.1056/NEJMoa2207323.
[4] Adler H, Gould S, Hine P, Snell LB, Wong W, Houlihan CF, et al. Clinical features and management of human monkeypox: a retrospective observational study in the UK. Lancet Infect Dis 2022;22:1153–62. https://doi.org/10.1016/S1473-3099(22)00228-6.
[5] Eisenstadt R, Liszewski WJ, Nguyen CV. Recognizing minimal cutaneous involvement or systemic symptoms in monkeypox. JAMA Dermatol 2022;158:1457–8. https://doi.org/10.1001/jamadermatol.2022.4652.
[6] Likos AM, Sammons SA, Olson VA, Frace AM, Li Y, Olsen-Rasmussen M, et al. A tale of two clades: monkeypox viruses. J Gen Virol 2005;86:2661–72. https://doi.org/10.1099/vir.0.81215-0.
[7] Català A, Clavo-Escribano P, Riera-Monroig J, Martín-Ezquerra G, Fernandez-Gonzalez P, Revelles-Peñas, et al. Monkeypox outbreak in Spain: clinical and epidemiological findings in a prospective cross-sectional study of 185 cases. Br J Dermatol 2022;187:765–72. https://doi.org/10.1111/bjd.21790.
[8] Isidro J, Borges V, Pinto M, Sobral D, Santos JD, Nunes A, et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. Nat Med 2022;28:1569–72. https://doi.org/10.1038/s41591-022-01907-y.
[9] Mileto D, Riva A, Cutrera M, Moschese D, Mancon A, Meroni L, et al. New challenges in human monkeypox outside Africa: a review and case report from Italy. Trav Med Infect Dis 2022;49:102386. https://doi.org/10.1016/j.tmaid.2022.102386.
[10] Vaughan A, Aarons E, Astbury J, Brooks T, Chand M, Flegg P, et al. Human-to-Human transmission of monkeypox virus, United Kingdom, october 2018. Emerg Infect Dis 2020;26:782–5. https://doi.org/10.3201/eid2604.191164.
[11] Altindis M, Puca E, Shapo L. Diagnosis of monkeypox virus - an overview. Trav Med Infect Dis 2022 Nov-Dec;50:102459. https://doi.org/10.1016/j.tmaid.2022.102459. Epub 2022 Sep 13. PMID: 36109000; PMCID: PMC9534096.
[12] WHO. Monkeypox outbreak: global trends. https://worldhealthorg.shinyapps.io/mpx_global/. [Accessed 12 April 2023].
[13] Billioux BJ, Mbaya OT, Sejvar J, Nath A. Potential complications of monkeypox. Lancet Neurol 2022;21:872.
[14] Huhn GD, Bauer AM, Yorita K, Graham MB, Sejvar J, Likos A, et al. Clinical characteristics of human monkeypox, and risk factors for severe disease. Clin Infect Dis 2005;41:1742–51. https://doi.org/10.1086/498115.
[15] Peiró-Mestres A, Fuertes I, Camprubí-Ferrer D, Marcos MÁ, Vilella A, Navarro M, et al. Frequent detection of monkeypox virus DNA in saliva, semen, and other clinical samples from 12 patients, Barcelona, Spain, May to June 2022. Euro Surveill 2022;27:2200503. https://doi.org/10.2807/1560-7917.ES.2022.27.28.2200503.
[16] Lim CK, McKenzie C, Deerain J, Chow EPF, Towns J, Chen MY, et al. Correlation between monkeypox viral load and infectious virus in clinical specimens. J Clin Virol 2023;161:105421. https://doi.org/10.1016/j.jcv.2023.105421.
[17] Palich R, Burrel S, Monsel G, Nouchi A, Bleibtreu A, Seang S, et al. Viral loads in clinical samples of men with monkeypox virus infection: a French case series. Lancet Infect Dis 2023;23:74–80. https://doi.org/10.1016/S1473-3099(22)00586-2.
[18] Number of confirmed monkeypox cases in Italy as of March 17. by region, https://www.statista.com/statistics/1323947/monkeypox-confirmed-cases-by-region-italy/. [Accessed 12 April 2023].
[19] Li Y, Zhao H, Wilkins K, Hughes C, Damon IK. Real-time PCR assays for the specific detection of monkeypox virus West African and Congo Basin strain DNA. J Virol Methods 2010;169:223–7. https://doi.org/10.1016/j.jviromet.2010.07.012.
[20] Maksyutov RA, Gavrilova EV, Shchelkunov SN. Species-specific differentiation of variola, monkeypox, and varicella-zoster viruses by multiplex real-time PCR assay. J Virol Methods 2016;236:215–20. https://doi.org/10.1016/j.jviromet.2016.07.024.
[21] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77. https://doi.org/10.1089/cmb.2012.0021.
[22] Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32:268–74. https://doi.org/10.1093/molbev/msu300.
[23] Wang L, Shang J, Weng S, Aliyari SR, Ji C, Cheng G, et al. Genomic annotation and molecular evolution of monkeypox virus outbreak in 2022. J Med Virol 2023;95:e28036. https://doi.org/10.1002/jmv.28036.
[24] WHO. Monkeypox posted on 19 may 2022. https://www.who.int/news-room/fact-sheets/detail/monkeypox. [Accessed 13 April 2023].
[25] Hernaez B, Muñoz-Gómez A, Sanchiz A, Orviz E, Valls-Carbo A, Sagastagoitia I, et al. Monitoring monkeypox virus in saliva and air samples in Spain: a cross-sectional study. Lancet Microbe 2023;4:e21–8. https://doi.org/10.1016/S2666-5247(22)00291-9.
[26] Suñer C, Ubals M, Tarín-Vicente EJ, Mendoza A, Alemany A, Hernández-Rodríguez Á, et al. Viral dynamics in patients with monkeypox infection: a prospective cohort study in Spain. Lancet Infect Dis 2023;23:445–53. https://doi.org/10.1016/S1473-3099(22)00794-0.
[27] Tarín-Vicente EJ, Alemany A, Agud-Dios M, Ubals M, Suñer C, Antón A, et al. Clinical presentation and virological assessment of confirmed human monkeypox virus cases in Spain: a prospective observational cohort study. Lancet 2022;400:661–9. https://doi.org/10.1016/S0140-6736(22)01436-2.
[28] DeWitt ME, Polk C, Williamson J, Shetty AK, Passaretti CL, McNeil CJ, et al. Global monkeypox case hospitalisation rates: a rapid systematic review and meta-analysis. EClinicalMedicine 2022;54:101710. https://doi.org/10.1016/j.eclinm.2022.101710.
[29] Brown K, Leggat PA. Human monkeypox: current state of knowledge and implications for the future. Trav Med Infect Dis 2016;1:8. https://doi.org/10.3390/tropicalmed1010008.
[30] Antinori A, Mazzotta V, Vita S, Carletti F, Tacconi D, Lapini LE, et al. Epidemiological, clinical and virological characteristics of four cases of monkeypox support transmission through sexual contact, Italy, May 2022. Euro Surveill 2022;27:2200421. https://doi.org/10.2807/1560-7917.ES.2022.27.22.2200421.
[31] Raccagni AR, Candela C, Mileto D, Canetti D, Bruzzesi E, Rizzo A, et al. Monkeypox infection among men who have sex with men: PCR testing on seminal fluids. J Infect 2022;85:573–607. https://doi.org/10.1016/j.jinf.2022.07.022.
[32] Noe S, Zange S, Seilmaier M, Antwerpen MH, Fenzl T, Schneider J, et al. Clinical and virological features of first human monkeypox cases in Germany. Infection 2023;51:265–70. https://doi.org/10.1007/s15010-022-01874-z.
[33] Lapa D, Carletti F, Mazzotta V, Matusali G, Pinnetti C, Meschi S, et al. Monkeypox virus isolation from a semen sample collected in the early phase of infection in a patient with prolonged seminal viral shedding. Lancet Infect Dis 2022;22:1267–9. https://doi.org/10.1016/S1473-3099(22)00513-8.
[34] Colavita F, Mazzotta V, Rozera G, Abbate I, Carletti F, Pinnetti C, et al. Kinetics of viral DNA in body fluids and antibody response in patients with acute Monkeypox virus infection. iScience 2023;26:106102. https://doi.org/10.1016/j.isci.2023.106102.
[35] Reda A, Abdelaal A, Brakat AM, Lashin BI, Abouelkheir M, Abdelazeem B, et al. Monkeypox viral detection in semen specimens of confirmed cases: a systematic

review and meta-analysis. J Med Virol 2023;95:e28250. https://doi.org/10.1002/jmv.28250.
[36] Kannan SR, Sachdev S, Reddy AS, Kandasamy SL, Byrareddy SN, Lorson CL, et al. Mutations in the monkeypox virus replication complex: potential contributing factors to the 2022 outbreak. J Autoimmun 2022;133:102928. https://doi.org/10.1016/j.jaut.2022.102928.
[37] Smith CA, Smith TD, Smolak PJ, Friend D, Hagen H, Gerhart M, et al. Poxvirus genomes encode a secreted, soluble protein that preferentially inhibits beta chemokine activity yet lacks sequence homology to known chemokine receptors. Virology 1997;236:316–27. https://doi.org/10.1006/viro.1997.8730.
[38] Cunnion KM. Tumor necrosis factor receptors encoded by poxviruses. Mol Genet Metabol 1999;67:278–82. https://doi.org/10.1006/mgme.1999.2878.
[39] Gigante CM, Korber B, Seabolt MH, Wilkins K, Davidson W, Rao AK, et al. Multiple lineages of monkeypox virus detected in the United States, 2021-2022. Science 2022;378:560–5. https://doi.org/10.1126/science.add4153.

Rapid communication

Clinical isolates of ST131 *bla*OXA-244-positive *Escherichia coli*, Italy, December 2022 to July 2023

Aurora Piazza[1,2] , Marta Corbella[3] , Vittoria Mattioni Marchetti[1,4] , Cristina Merla[3] , Irene Mileto[3,4] , Angela Kuka[3,4] , Greta Petazzoni[1,3] , Stefano Gaiarsa[3] , Roberta Migliavacca[1,2] , Fausto Baldanti[1,2,3] , Patrizia Cambieri[3]

1. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia, Pavia, Italy
2. IRCCS Fondazione Policlinico San Matteo, Pavia, Italy
3. Microbiology and Virology Unit, IRCCS Fondazione Policlinico San Matteo, Pavia, Italy
4. Specialization School of Microbiology and Virology, University of Pavia, Italy

Correspondence: Vittoria Mattioni Marchetti (vittoria.mattionimarchetti@unipv.it)

Citation style for this article:
Piazza Aurora, Corbella Marta, Mattioni Marchetti Vittoria, Merla Cristina, Mileto Irene, Kuka Angela, Petazzoni Greta, Gaiarsa Stefano, Migliavacca Roberta, Baldanti Fausto, Cambieri Patrizia. Clinical isolates of ST131 blaOXA-244-positive Escherichia coli, Italy, December 2022 to July 2023. Euro Surveill. 2024;29(8):pii=2400073. https://doi.org/10.2807/1560-7917.ES.2024.29.8.2400073

Article submitted on 02 Feb 2024 / accepted on 22 Feb 2024 / published on 22 Feb 2024

The dissemination of carbapenemase-producing *Escherichia coli*, although still at low level, should be continuously monitored. OXA-244 is emerging in Europe, mainly in *E. coli*. In Italy, this carbapenemase was reported from an environmental river sample in 2019. We report clinical isolates of OXA-244-producing ST131 *E. coli* in four patients admitted to an acute care hospital in Pavia, Italy. The association of this difficult-to-detect determinant with a globally circulating high-risk clone, ST131 *E. coli*, is of clinical relevance.

OXA-244, a single-point mutation variant (Arg214Gly) of OXA-48, is a rarely reported carbapenemase with reduced activity against carbapenems, first described in 2011 in a *Klebsiella pneumoniae* clinical isolate in Spain [1,2]. To date, the OXA-244 variant has mainly been described in *Escherichia coli* and found in several countries across the European Union, Lebanon, Russia and the United Kingdom [3-6]. Although the origin of this variant is not yet certain, there are reports of not healthcare-related community and environmental spread.

The aim of the present report was to investigate the genetic diversity of the ST131 *E. coli* isolated at the Fondazione IRCCS Policlinico San Matteo (OSM) in Pavia, a 1,000-bed teaching hospital, using whole genome sequencing (WGS), with a focus on the possible presence of OXA-244-producing strains.

Case description

During the 10-month period from November 2022 to September 2023, we collected and sequenced 30 carbapenem-resistant *E. coli* (CR-Ec) isolates. Ten of them, collected from 10 patients in the period December 2022 to July 2023, belonged to the high-risk clone ST131. Bacteria of this ST were isolated at OSM throughout most of the study period. The patients had a mean age of 58.3 years, ranging from 35 to 90 years, and eight of the 10 were male. Two patients were admitted to the pneumology unit, two to the infectious diseases unit, three to intensive care units (ICU), two to surgery wards, and one to general medicine. Among the 10 ST131 CR-Ec isolates, six were isolated from rectal swabs, two from nasopharyngeal aspirate, one from a urine sample and one from a blood culture.

Antimicrobial profiles

Species identification and determination of antibiotic susceptibility profiles were carried out by MALDI-TOF MS (Bruker Daltonics, Bremen, Germany) using the software Bruker Biotyper 3.1 and Phoenix M50 BD (Becton Dickinson, Franklin Lakes, United States (US)) and interpreted according to the 2023 European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints [9]. The presence of carbapenem resistance mechanism was confirmed with the NG-test Carba 5 immunochromatographic assay (NG Biotech Laboratories). All 10 isolates were resistant to ertapenem, cephalosporins, amoxicillin/clavulanate, ciprofloxacin and levofloxacin. Seven of the 10 isolates were resistant to gentamicin, and only two were resistant to trimethoprim/sulfamethoxazole. All 10 *E. coli* ST131 strains were susceptible to meropenem, and nine were resistant to imipenem and amikacin (Table 1).

Genetic characterisation of OXA-244-producing ST131

We performed WGS on the Illumina MiSeq platform. The reads obtained were de novo assembled with Shovill [8]. We reconstructed the resistome, plasmidome and virulome of the isolates using ResFinder, PlasmidFinder and the Virulence Factors Database (VFDB) via ABRicate (github.com/tseemann/ABRicate), and determined the multilocus sequence typing (MLST) profiles according to the Achtman scheme on Enterobase [9].

Table 1

Antibiotic susceptibility profiles of ST131 *Escherichia coli* strains, Italy, December 2022–July 2023 (n = 10)

Strain	Isolation date	Specimen	Ward	AMC S≤0,25 R›0,5	P/T S≤8 R›8	FEP S≤1 R›4	CTX S≤2 R›2	CAZ S≤1 R›4	CZA S≤8 R›8	C/T S≤2 R›2	CFX S≤8 R›8	CRX S≤1 R›2	CIP S≤0.25 ›0.5	LVX S≤0.5 R›1	ERT S≤0.5 R›0.5	IMP S≤2 R›4	MEM S≤2 R›8	AK S≤8 R›8	CN S≤2 R›2	SXT S≤2 R›4
32829	Dec 2022	Blood	Intensive care unit	›32/2 (R)	›32/4 (R)	›16 (R)	›4 (R)	›16 (R)	0,5/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	2 (R)	1 (R)	2 (S)	0,25 (S)	≤4 (S)	≤1 (S)	›8/152 (R)
7662	Dec 2022	Rectal swab	Infectious diseases	›32/2 (R)	›32/4 (R)	›16 (R)	›4 (R)	›16 (R)	2/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	›8 (R)	›2 (R)	4 (I)	2 (S)	≤4 (S)	≤1 (S)	≤1/19 (S)
7730	Jan 2023	Urine	Infectious diseases	›32/2 (R)	›16/4 (R)	›8 (R)	›4 (R)	›8 (R)	0,125 (S)	32/4 (R)	›8 (R)	›4 (R)	›1 (R)	›2 (R)	›1 (R)	4 (I)	4 (I)	16 (R)	›4 (R)	≤1/19 (S)
7944	Feb 2023	Nasopharyngeal aspirate	Pneumology	›32/2 (R)	›16/4 (R)	8 (R)	›4 (R)	8 (R)	0,064 (S)	4/4 (R)	›8 (R)	›4 (R)	›1 (R)	›2 (R)	›1 (R)	0,5 (S)	0,25 (S)	8 (S)	›4 (R)	≤1/19 (S)
8136	Mar 2023	Rectal swab	Intensive care unit	›32/2 (R)	›32/4 (R)	›16 (R)	›4 (R)	›16 (R)	0,5/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	›8 (R)	›2 (R)	≤0,25 (S)	2 (S)	≤4 (S)	≤1 (S)	›8/152 (R)
8171	Apr 2023	Rectal swab	General medicine	›32/2 (R)	›32/4 (R)	›16 (R)	›4 (R)	›16 (R)	1/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	›8 (R)	›2 (R)	1 (S)	2 (S)	≤4 (S)	2 (S)	≤1/19 (S)
8270	May 2023	Urine	Intensive care unit	›32/2 (R)	›32/4 (R)	›16 (R)	›4 (R)	›16 (R)	4/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	›8 (R)	›2 (R)	1 (S)	2 (S)	≤4 (S)	≤1 (S)	≤1/19 (S)
8345	May 2023	Rectal swab	Surgery	›32/2 (R)	›32/4 (R)	8 (R)	›4 (R)	›16 (R)	0,5/4 (S)	›2/4 (R)	›16 (R)	›4 (R)	›1 (R)	›8 (R)	›2 (R)	8 (R)	2 (S)	≤4 (S)	≤1 (S)	≤1/19 (S)
8638	Jul 2023	Nasopharyngeal aspirate	Pneumology	›32/2 (R)	›64/4 (R)	8 (R)	›4 (R)	8 (R)	0,5/4 (S)	›4/4 (R)	›8 (R)	›4 (R)	›1 (R)	›1 (R)	1 (R)	≤0,25 (S)	0,25 (S)	8 (S)	›4 (R)	≤1/19 (S)
8718	Jul 2023	Rectal swab	Surgery	›32/2 (R)	64/4 (R)	›8 (R)	›4 (R)	›16 (R)	0,5/4 (S)	2/4 (S)	›8 (R)	›4 (R)	›1 (R)	›1 (R)	›1 (R)	≤0,25 (S)	≤0,125 (S)	8 (S)	2 (S)	≤1/19 (S)

AK: amikacin; AMC: amoxicillin/clavulanate; CAZ: ceftazidime; CFX: cefuroxime; CFZ: cefazolin; CIP: ciprofloxacin; CN: gentamycin; CRX: ceftriaxone; C/T: ceftolozane/tazobactam; CTX: cefotaxime; CZA: ceftazidime/avibactam; ETP: ertapenem; FEP: cefepime; IMP: imipenem; LVX: levofloxacin; MEM: meropenem; P/T: piperacillin/tazobactam; SXT: trimethoprim/sulfamethoxazole.

Interpretation according to EUCAST breakpoints [7] are displayed in brackets.

Table 2

Metadata and genomic results of ST131 *Escherichia coli* isolates, Italy, December 2022–July 2023 (n = 10)

Strain	Isolation date	Specimen	Serotype	FimH	Clade	Resistome	Plasmidome
32829	Dec 2022	Blood	O16:H5	H41	A	*aadA5*, *blaCTX*-M-15, ***blaOXA-244***, *dfrA17*, *mphA*, *sul1*	Col156, ColRNAI, IncFIA, IncFIB, IncFII
7662	Dec 2022	Rectal swab	O25b:H4	H30	C1-M27	*aph(3")-Ib*, *aph(6)-Id*, *bla*CTX-M-27, *bla*KPC-3, sul2, *tetA*	Col(MG828), Col156, IncFIA, IncFIB, IncFII
7730	Jan 2023	Urine	O25b:H4	H30	C2	*aac(3)-IIe*, *aac(6')-Ib*, *bla*CTX-M-15, *bla*OXA-1, ***blaOXA-244***	IncFIA, IncFIB, IncFIC
7944	Feb 2023	Nasopharyngeal aspirate	O25b:H4	H30	C2	*aac(3)-IIe*, *aac(6')-Ib*, *bla*CTX-M-15, *blaOXA*-1, ***blaOXA-244***, tetA	Col156, IncFIA, IncFIB, IncFIC
8136	Mar 2023	Rectal swab	O25b:H4	H30	C1-M27	*aadA5*, *aph(3")-Ib*, *aph(6)-Id*, *bla*CTX-M-27, *dfrA17*, *mphA*, *sul1*, *sul2*, *tetA*	Col(MG828), Col156, Col8282, ColRNAI, IncFIA, IncFIB, IncFII
8171	Apr 2023	Rectal swab	O25b:H4	H30	C2	*aph(3)-Ib*, *bla*CTX-M-15, *bla*TEM-235, *sul2*	Col(MG828), IncB/O/K/Z, IncFIB, IncFII
8270	May 2023	Urine	O25b:H4	H30	C1-M27	*bla*CTX-M-27	Col(BS512), Col(MG828), Col156, IncFIA, IncFIB
8345	May 2023	Rectal swab	O25b:H4	H30	C	*bla*KPC-3, *bla*SHV-158, *bla*TEM-1, *tetA*	Col(MG828), Col156, Col8282, IncFIA, IncFIB, IncFII, IncN, IncX3
8638	Jul 2023	Nasopharyngeal aspirate	O25b:H4	H30	C2	*aac(3)-IIe*, *aac(6')-Ib*, *bla*CTX-M-15, *bla*OXA-1, ***blaOXA-244***, *tetA*	IncFIA, IncFIB, IncFIC
8718	Jul 2023	Rectal swab	O25b:H4	H30	C2	*aac(6')-Ib*, *aph(3')-VI*, *bla*CTX-M-15, *bla*NDM-1, *bla*OXA-1, *dfrA5*, *mphA*, *mphE*, *msrE*, *qnrS1*, *sul1*, *sul2*	Col156, ColRNAI, IncFIB, IncFIB(pNDM-Mar), IncFII, IncHI1B, IncI1

Nine of the 10 strains harboured a *bla*CTX-M gene, of which six were the *bla*CTX-M-15 and three the *bla*CTX-M-27 variant. Among carbapenemases, the *bla*KPC-3 was detected in two strains from different wards, *bla*NDM-1 in one isolate, and four isolates from three wards were positive for *bla*OXA-244 (OXA-244-Ec). Table 2 adds information on genomic resistance determinants and plasmids.

Nine of the 10 ST131 strains showed the O25b:H4 serotype with a H30 fimbrial type, while the remaining, *bla*OXA-244-harbouring strain showed the O16:H5 with a H41 fimbrial type. Concerning the ST131 sublineages, three isolates belonged to C1-M27 (positive for *bla*CTX-M-27), five belonged to C2 and produced CTX-M-15, one belonged to clade C and encoded the carbapenemase KPC-3, and one to clade A with the H41 fimbrial type (Table 2).

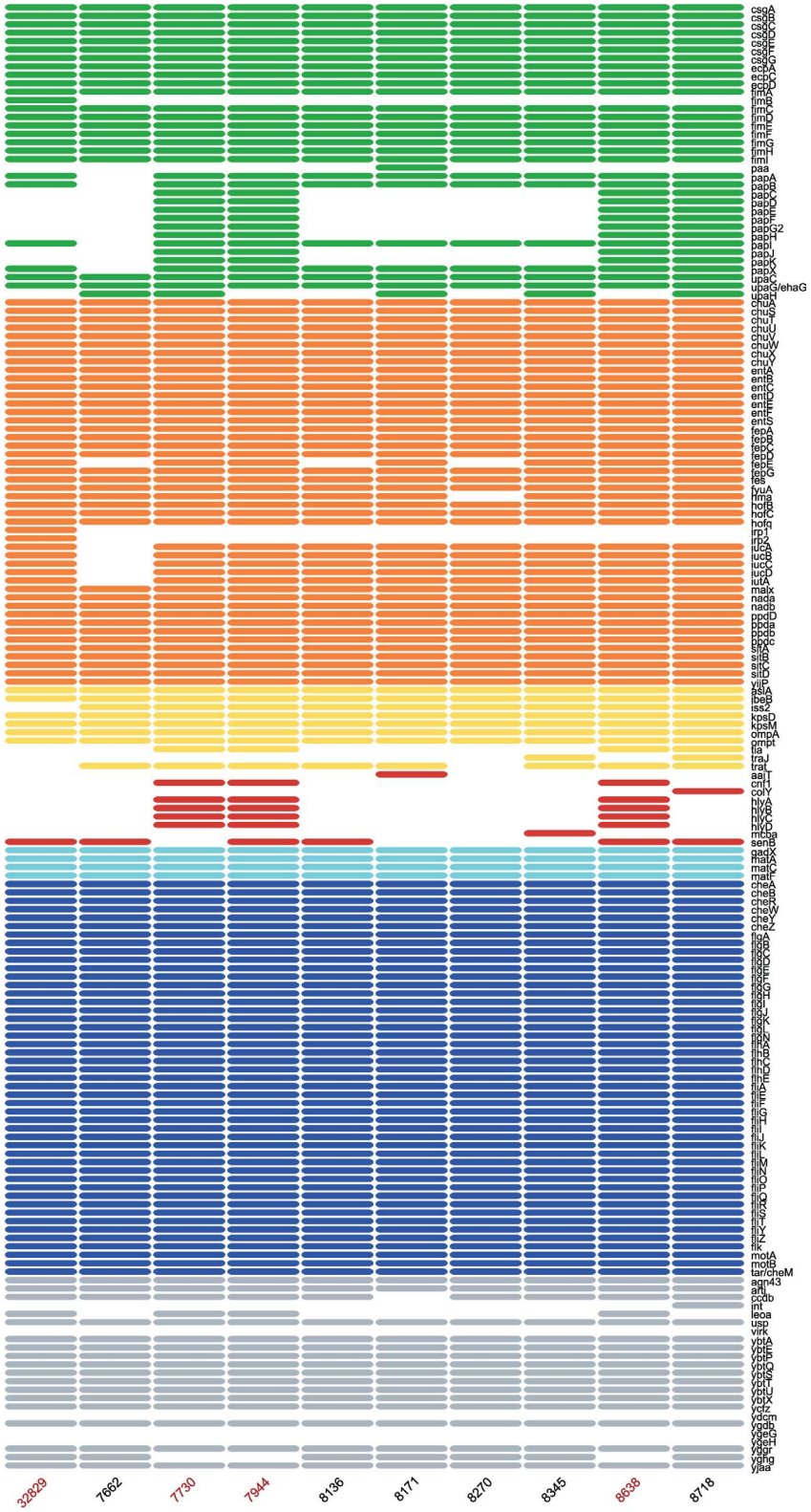
All strains showed a conserved virulence pattern, associated with ST131 isolates, including genes (operons/gene groups) for iron uptake systems (*sitABCD, chuA, fyuA and malX*), for adhesion factors (*ecpACD* and *fimH*) and for invasion (*kpsM II* and *ibeB*). The operon *iucABCD* and the determinant *iutA* were identified in all but one strain. Among genes encoding for toxins, all 10 strains presented the *usp* gene, encoding a colicin-like toxin, six strains contained *senB,* the gene for enterotoxin TieB. The three OXA-244-positive O25b:H4 isolates carried in addition the cytotoxic necrotising factor 1 gene *cnf1,* and the hemolysin gene *hlyA* (Figure 1).

The plasmid content showed that the IncF family with IncFI replicons was present in all 10 isolates, followed by IncFII which was found in six isolates. Among plasmids of Col-type, the Col156 plasmid was found in seven strains, co-existing in four cases with other specific Col-like replicons Col(MG828) and Col8282. Less frequent plasmid families such as IncX3, IncN and IncB/O/K/Z were also identified (Table 2).

The *bla*OXA-244 gene was found to be part of the Tn*1999.2* transposon in all four strains; the association of *bla*OXA-244 with Tn1999.2 has already been described in the literature, usually with a chromosomal location [4,10]. Single nucleotide polymorphism (SNP)-based maximum-likelihood phylogeny was inferred using P-DOR [11] on the four OXA-244-Ec and the 46 most similar high-quality genomes belonging to ST131 available from the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) database (https://www.bv-brc.org) (Figure 2). Based on the SNPs matrix, our four OXA-244-Ec belonged to two different clusters. The isolate 32829 was closely related to a small cluster of OXA-244-Ec collected from a rectal swab in Germany in 2019 [3]. Moreover, isolate 32829 revealed a relevant identity (SNP range: 194–197) with two other OXA-244-Ec collected from a water basin in Italy in 2019, in the same geographical area as OSM [10]. In contrast, the remaining three strains showed a monophyletic

Figure 1

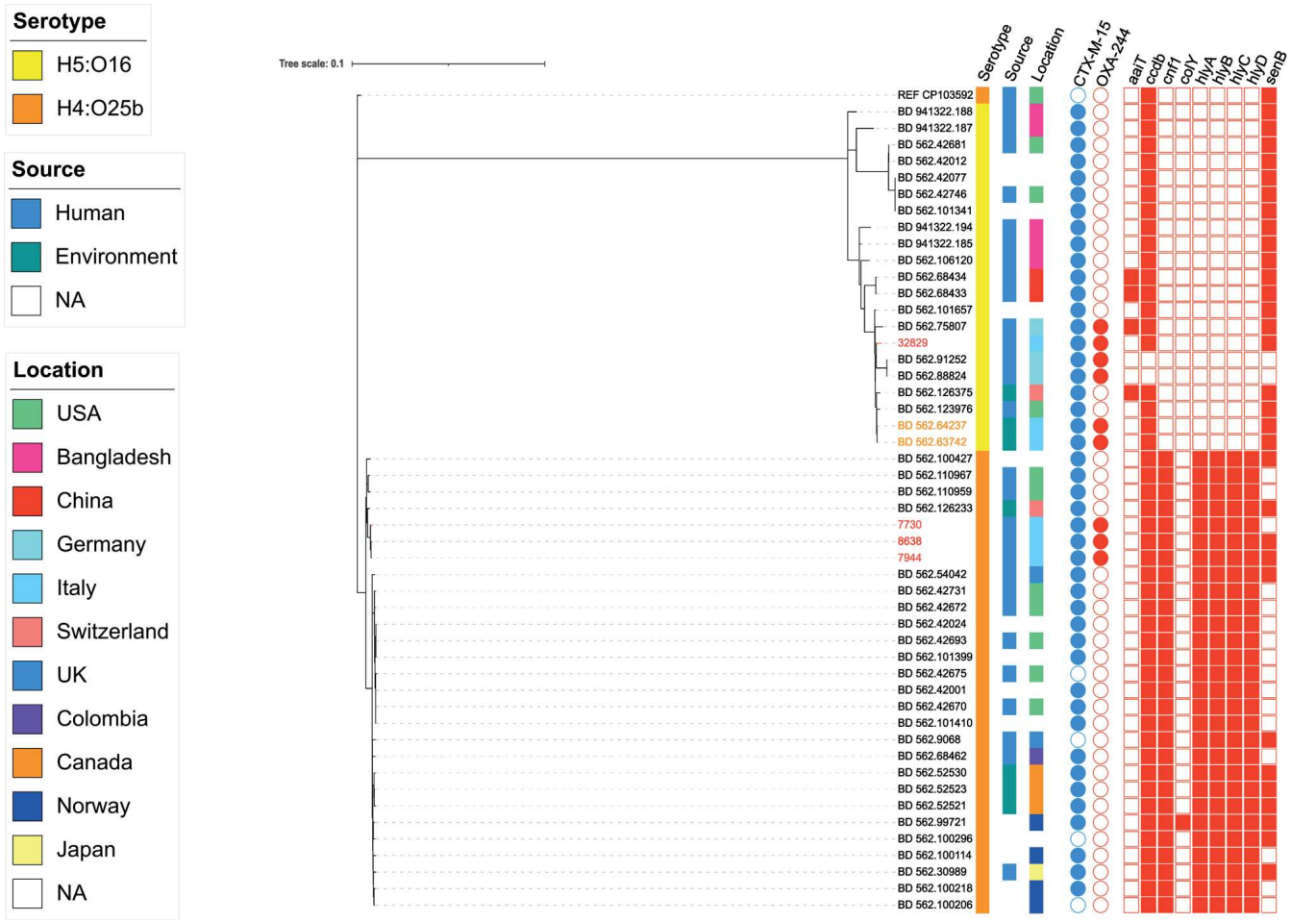
Heatmap representation of the virulence genes of *Escherichia coli* ST131 isolates, Italy, December 2022–July 2023 (n = 10)



Green: adhesion; orange: metabolism; yellow: invasion; red: toxins; light blue: regulation; blue: chemotaxis; grey: others. The four OXA-244-producing *E. coli* ST131 are labelled in red.

Figure 2

Phylogenetic analysis of OXA-244-producing *Escherichia coli* ST131, Italy, December 2022–July 2023 (n = 4)



NA: not available.

Maximum-likelihood phylogeny of 52 *E. coli* ST131 strains inferred on coreSNPs visualised using iTOL v6, including the four Italian OXA-244-producing *E. coli* ST131 (red labels) and 48 from the BV-BRC database (grey labels). Labelled in orange are the two Italian OXA-244-producing *E. coli* ST131 collected from water. The presence/absence matrix shows the pattern of resistance determinants (blue and red dots), toxin determinants (red square) and resistance determinants (red).

relation (SNP range: 52–95) and clustered together with clinical and environmental strains of CTX-M-15-producing *E. coli* ST131 collected in the US and in Switzerland, respectively. These data suggested a separate importation of OXA-244-Ec to the OSM.

Epidemiological investigation

The four *bla*OXA-244-positive *E. coli* were collected from four patients admitted to three different OSM wards at distinct times. The first isolate was collected in December 2022 from an ICU, the second in January 2023 on the infectious diseases ward, and the remaining two in the pneumology unit in February and July 2023. The two patients admitted to the pneumology ward did not overlap, with a gap of 5 months between their hospitalisations. The epidemiological investigation into possible contact among the patients revealed that three of them had been admitted to the ICU in February, October and November 2022, with no overlap. Environmental samples were not taken on the above wards, but no other similar strains have since been identified in any patient samples. The four patients had been transferred to OSM from other hospitals but were negative when screened for carbapenem-resistant *Enterobacterales* at admission to OSM. In 2019, an OXA-244-producing ST131 H5:O16 *E. coli* had been detected in canal water in the same geographical area [10], but we could not clearly assess the role of the environment or the community setting as a reservoir for such a strain.

Discussion

Escherichia coli belonging to sequence type (ST) 131 is well known as a successful high-risk multidrug-resistant clone. Typically, isolates belonging to ST131 clade C (the most prevalent), are resistant to fluoroquinolones, due to mutations in the quinolone resistance-determining region, and to third-generation cephalosporins, due to the association with *bla*CTX-M-type determinants. However, they mostly remain

susceptible to carbapenems [12]. In July 2021, a risk assessment published by European Centre for Disease Prevention and Control (ECDC) pointed out an increase in the prevalence of OXA-244-producing *E. coli* in Europe and evidence of healthcare-associated transmission [13,14], in contrast to what had been pointed out in the literature since early 2021. OXA-244 is known to be a difficult-to-detect carbapenemase, due to its weak hydrolytic activity towards carbapenems and temocillin and, thus, its difficulty to grow on selective and screening media [6]. To date, OXA-244-producing *E. coli* is predominantly found in ST38 but reported also in several other clonal groups such as ST10, ST69, ST167, ST361 and ST3268 [5,15]. The ability of OXA-244-producing *E. coli* ST131 to occur in clinical settings has already been reported in Germany and France, while in Italy, the sole report had been from water environment [3,6,10,15]. Here we describe the first evidence of such strains in a clinical setting in Italy. The lack of environmental sampling on the involved wards represents a limitation; indeed, the entry timeline of OXA-244-Ec in the OSM is hard to track. Nevertheless, the association of the difficult-to-detect determinant OXA-244 with a high-risk clone such as ST131 represents a concern for global health.

Conclusion

The presence of OXA-244 in the healthcare setting poses an additional challenge for microbiological diagnosis and surveillance. The source and route of transmission of OXA-244-producing *E. coli* in Europe is still unclear, and there is a need for further investigations in order to implement evidence-based control measures. It is important to not underestimate the sole ertapenem resistance, and to use rapid tests for the detection of carbapenemases in patients on wards such as ICUs, followed by molecular-based characterisation of the specific resistance genes, in order to rapidly initiate targeted and specific antibiotic therapy when required.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105). The work described herein is a retrospective study performed on bacterial isolates from human samples that were obtained as part of routine hospital care. An informed consent was signed by all the patients as part of the hospital routine activity, and according to 'Comitato Etico Territoriale Lombardia 6'.

Funding statement

This research was supported by EU funding within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. P Eoooooo07, INF-ACT) and by Ricerca Corrente n 08022218 "Real-Time bioinformatic surveillance of nosocomial infections and multidrug resistant bacteria to improve infection control".

Conflict of interest

None declared.

Authors' contributions

Sample collection: AK, IM. Conceptualisation: MC, CM, PC. Formal analysis: VMM, AP, GP, SG. Supervision: RM, MC, PC, FB. Writing—original draft: AP, VMM. Writing—review and editing: VMM, AP, RM, CM, MC. All authors have read and agreed to the published version of the manuscript.

References

1. Potron A, Poirel L, Dortet L, Nordmann P. Characterisation of OXA-244, a chromosomally-encoded OXA-48-like β-lactamase from Escherichia coli. Int J Antimicrob Agents. 2016;47(1):102-3. https://doi.org/10.1016/j.ijantimicag.2015.10.015 PMID: 26655033
2. Oteo J, Hernández JM, Espasa M, Fleites A, Sáez D, Bautista V, et al. Emergence of OXA-48-producing Klebsiella pneumoniae and the novel carbapenemases OXA-244 and OXA-245 in Spain. J Antimicrob Chemother. 2013;68(2):317-21. https://doi.org/10.1093/jac/dks383 PMID: 23034714
3. Welker S, Boutin S, Miethke T, Heeg K, Nurjadi D. Emergence of carbapenem-resistant ST131 Escherichia coli carrying blaOXA-244 in Germany, 2019 to 2020. Euro Surveill. 2020;25(46):2001815. https://doi.org/10.2807/1560-7917.ES.2020.25.46.2001815 PMID: 33213685
4. Davidova-Gerzova L, Lausova J, Sukkar I, Nesporova K, Nechutna L, Vlkova K, et al. Hospital and community wastewater as a source of multidrug-resistant ESBL-producing Escherichia coli. Front Cell Infect Microbiol. 2023;13:1184081. https://doi.org/10.3389/fcimb.2023.1184081 PMID: 37256105
5. Hammerum AM, Porsbo LJ, Hansen F, Roer L, Kaya H, Henius A, et al. Surveillance of OXA-244-producing Escherichia coli and epidemiologic investigation of cases, Denmark, January 2016 to August 2019. Euro Surveill. 2020;25(18):1900742. https://doi.org/10.2807/1560-7917.ES.2020.25.18.1900742 PMID: 32400363
6. Emeraud C, Biez L, Girlich D, Jousset AB, Naas T, Bonnin RA, et al. Screening of OXA-244 producers, a difficult-to-detect and emerging OXA-48 variant? J Antimicrob Chemother. 2020;75(8):2120-3. https://doi.org/10.1093/jac/dkaa155 PMID: 32363407
7. The European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters Version 13.0, valid from 2023-01-01. Växjö: EUCAST; 2023. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_13.0_BreakBreak_Tables.pdf
8. Seeman T, Edwards R, Goncalves da Silva A, Kiil K. Shovill. Assemble bacterial isolate genomes from Illumina paired-end reads. San Francisco; GitHub; 2020. Available from: https://github.com/tseemann/shovill
9. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Agama Study Group, Achtman M. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res. 2020;30(1):138-52. https://doi.org/10.1101/gr.251678.119 PMID: 31809257
10. AbuAlshaar A, Piazza A, Mercato A, Marchesini F, Mattioni Marchetti V, Bitar I, et al. OXA-244-producing ST131 Escherichia coli from surface and groundwaters of Pavia urban area (Po plain, Northern Italy). Front Microbiol. 2022;13:920319. https://doi.org/10.3389/fmicb.2022.920319 PMID: 35756019
11. Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics. 2023;39(9):btad571. https://doi.org/10.1093/bioinformatics/btad571 PMID: 37701995
12. Kocsis B, Gulyás D, Szabó D. Emergence and dissemination of extraintestinal pathogenic high-risk international clones of Escherichia coli. Life (Basel). 2022;12(12):2077. https://doi.org/10.3390/life12122077 PMID: 36556442
13. European Centre for Disease Prevention and Control (ECDC). OXA-244-producing Escherichia coli in the European Union/European Economic Area and the UK since 2013, first update – 20 July 2021. Stockholm: ECDC; 2021. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/OXA-244-producing-E-coli-in-EU-EEA-since-2013-first-update.pdf

14. Lindemann PC, Pedersen T, Oma DH, Janice J, Grøvan F, Chedid GM, et al. Intraregional hospital outbreak of OXA-244-producing Escherichia coli ST38 in Norway, 2020. Euro Surveill. 2023;28(27):2200773. https://doi.org/10.2807/1560-7917.ES.2023.28.27.2200773 PMID: 37410380

15. Falgenhauer L, Nordmann P, Imirzalioglu C, Yao Y, Falgenhauer J, Hauri AM, et al. Cross-border emergence of clonal lineages of ST38 Escherichia coli producing the OXA-48-like carbapenemase OXA-244 in Germany and Switzerland. Int J Antimicrob Agents. 2020;56(6):106157. https://doi.org/10.1016/j.ijantimicag.2020.106157 PMID: 32919009

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2024.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Spectrum

Bacteriology | Research Article

One-year surveillance for hypervirulent *Klebsiella pneumoniae* detected carbapenem-resistant superbugs

C. Merla,[1] A. Kuka,[1] I. Mileto,[1,2] G. Petazzoni,[1,3] S. Gaiarsa,[1] D. De Vitis,[1] M. Ardizzone,[1] M. Corbella,[1] F. Baldanti,[1,3] P. Cambieri[1]

AUTHOR AFFILIATIONS See affiliation list on p. 9.

ABSTRACT Hypervirulent *Klebsiella pneumoniae* (hvKp) can cause infections in clinically healthy people, such as young and immunocompetent patients. Genes involved in the capsule synthesis or those encoding the siderophores have been adopted as predictors of hvKp. Certain sequence types, such as ST23 and ST86, have been associated with hvKp strains, too. The aim of this study was to investigate the presence of hvKp among 354 *K. pneumoniae* strains isolated from clinical samples of patients admitted to an Italian 900-bed hospital between 21 May 2021 and April 2022. All the isolates were screened by PCR for the amplification of virulence loci. Whole genome sequencing was performed in strains tested positive for at least one target gene. Thirteen out of 354 (3.7%) were hvKp. Five were wild type and belonged to the hypervirulent clones ST23, ST86, ST5, and ST375 and to the new clone ST6310. Six strains carried the *blaKPC* gene: three belonged to ST101, two to ST512, and one to ST395. Two isolates were ST147 and carried the *blaNDM* gene. Although hvKp isolation is not frequent, their presence should be systematically investigated to avoid the spreading of both virulent strains and strains with combined increase in virulence and resistance to antibiotics. PCR-based protocols are essential for surveillance of these strains, which do not always show a recognizable phenotype. Moreover, hvKp strains were isolated also from patients without history of recent foreign travels, indicating an increased spreading of these strains as well as an underestimated of their circulation so far.

IMPORTANCE *Klebsiella pneumoniae* is a healthcare-associated pathogen frequently resistant to antibiotics. Hypervirulent strains of pneumoniae (hvKp) can spread from the primary site of infection to multiple sites causing life-threatening infections also in young otherwise healthy individuals. This study described the isolation of 13 isolates of *K. pneumoniae* with increased virulence in a large tertiary hospital over a 1-year period. Among them, eight strains were multidrug resistant and hypervirulent. Although these hypervirulent strains are still rare in Italy, their presence is particularly concerning since they can cause difficult-to-treat life-threatening infections. Moreover, not all the hypervirulent isolates were positive by the string test, so hvKp isolates were not always phenotypically detectable. Molecular biology techniques such as PCR amplification and next generation sequencing are therefore necessary for the detection of hvKp isolates, and surveillance programs exploiting molecular techniques are highly desirable.

KEYWORDS hypervirulent *Klebsiella pneumoniae*, NDM, multidrug resistance, KPC, ST23, ST86, ST147, ST6310

Klebsiella pneumoniae has been recognized by the World Health Organization as a critical priority healthcare-associated pathogen due to antibiotic resistance. Strains of *K. pneumoniae* with increased virulence have spread from Asia, where the first strain with this characteristic was first isolated in the 1980s (1) to Europe and the United States (2–5).

Editor Florence Claude Doucet-Populaire, University Paris-Saclay, AP-HP Hôpital Antoine Béclère, Service de Microbiologie, Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, France

Address correspondence to C. Merla, c.merla@smatteo.pv.it.

C. Merla and A. Kuka contributed equally to this article. The order of names between the two first authors was decided based on the contribution in writing the manuscript.

The authors declare no conflict of interest.

Received 6 September 2023
Accepted 25 December 2023
Published 30 January 2024

[This article was published on 30 January 2024 with inaccurate information in the legend of Fig. 1. The legend was updated in the current version, posted on 22 February 2024.]

Copyright © 2024 Merla et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

Hypervirulent strains of *K. pneumoniae* (hvKp) can cause infections in young otherwise healthy individuals, differently from nosocomial multidrug resistant (MDR) *K. pneumoniae*. hvKp can disseminate from the primary site of infection to multiple sites causing life-threatening infections, such as hepatic abscesses, pneumonia, necrotizing fasciitis, endophthalmitis, meningitis, and sepsis (4).

Currently, the diagnosis of hvKp infections is based on clinical criteria, which includes unusual metastatic spread of *K. pneumoniae* infections, clusters of *K. pneumoniae* infections with increased severity and mortality. On the other hand, no single microbiological characteristic can define hvKp (6). Hypermucoviscosity is one of the main characteristics of hvKp strains, as most of these strains overexpress genes for capsular polysaccharides such as *rmpA* and *magA*, resulting in hypermucoid phenotype and higher resistance to phagocytosis and intercellular killing by neutrophils. Hypermucoviscosity is detected by string test, in which the colony is stretched with a loop. The formation of a viscous string longer than 5 mm is considered a positive string test (7). However, recent studies (8, 9) have shown that hypermucosity is not an exclusive feature of hvKp, while conversely, not all hvKp isolates exhibit a hypermucoviscous phenotype. Most of hvKp isolates belong to sequence types ST23, ST25, ST65, ST86, and ST375 (6).

The presence of iron acquisition systems (e.g., aerobactin, salmochelin, yersiniabactin) involved in bacterial growth and survival is indicative of hvKp strains (4, 6). Detection of genes such as *iutA* and *iroN*, which regulate iron acquisition systems, by PCR has been used in several studies (5, 10).

Recently, the *peg-344* gene has been included as a reliable, sensitive, and specific marker for the detection of hvKp such as aerobactin, salmochelin, *rmpA*, and *rmpA2* (8, 11). Indeed, PEG-344, whose function is still unclear, was proved to be required for maximal virulence in a pneumonia model (12).

Most hvKp strains remain susceptible to several commonly used antibiotics. However, carbapenem-resistant *K. pneumoniae* strains belonging to ST11 became hypervirulent by acquiring a pLVPK-like virulence plasmid encoding aerobactin, salmochelin, and RmpA (13). On the other hand, some hypervirulent strains of ST25 and ST65 exhibit carbapenem resistance (14, 15).

The presence of both increased virulence and resistance in carbapenem-resistant hvKp limits the number of effective antimicrobial agents and therefore poses a serious challenge to treatment, infection control, and public health.

The aim of this study is to investigate the presence of kvKp isolates among *K. pneumoniae* strains isolated over a period of 1 year from patients at Fondazione IRCCS Policlinico San Matteo in Pavia (Italy).

MATERIALS AND METHODS

Samples

Fondazione IRCCS Policlinico San Matteo is a 900-bed hospital, located in Lombardy region of Italy, where *Klebsiella pneumoniae* carbapenemase (KPC)- producing *K. pneumoniae* has been endemic since 2013 (16). In the study period, 1,822 *K. pneumoniae* isolates were collected during the routines of the bacteriology laboratory from a total of 627 patients. In this study, we searched for hvKp strains in 354 non-repetitive clinical *K. pneumoniae* isolates (one per patient) collected from patients at Fondazione IRCCS Policlinico San Matteo between May 2021 and April 2022. The selected 354 strains included strains with characteristics associated with virulence and strains from invasive infections such as bloodstream infections. In addition, we wanted to evaluate the presence of strains with both virulence and resistance characteristics that had never been reported in our hospital. Therefore, extended spectrum beta-lactamase (ESBL)-producing and carbapenem-resistant strains, hereafter referred to as MDR, were also considered. In detail, the 354 strains included:

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

- 21 *K. pneumoniae* strains positive by string test and isolated from any clinical specimen;
- 225 *K. pneumoniae* MDR strains negative by string test and isolated from any clinical specimen except for blood cultures; and
- 108 *K. pneumoniae* wild-type isolates obtained from blood cultures.

The type of specimens and the type of patient from which the 354 *K. pneumoniae* isolates have been isolated, as well as the resistance profile of the isolates, are shown in Table 1.

MALDI-TOF (Bruker Daltonics GmbH, Bremen, Germany) equipped with Bruker Biotyper 3.1 database was used for species identification. The antibiotic resistance profile was determined by the BD-Phoenix instrument (Becton Dickinson, USA), using the NMIC-505 panel.

PCR amplification, sequencing, and genome analysis

Genomic DNA was extracted with a Blood and Tissue kit (QIAGEN, Düsseldorf, Germany) following the manufacturer's instructions. PCR reactions were performed to investigate the presence of *iutA*, *iroN* loci (10), *rmpA* (17), and *peg-344* (18).

The genomes of the isolates positive for at least one virulence gene were sequenced. Short reads were obtained by Illumina MiSeq with a 2 × 150 paired-end run, after Nextera XT library preparation (Illumina Inc., San Diego, USA). Genomes were assembled using Shovill 1.1. Kleborate (19) was used to assess the sequence type and to search for virulence and resistance genes. The presence of plasmids was evaluated with PlasmidFinder (20) and manually curated.

A coreSNP-based phylogeny was performed for isolates belonging to ST147, since an outbreak caused by *K. pneumoniae* ST147 started in Tuscany region of Italy in 2018 and continued until 2021 (21). The data set consisted of 181 strains of *K. pneumoniae* ST147 from the outbreak (21, 22) and the two isolates from this study. A genomic background was constructed by retrieving the 20 most similar high-quality genomes

TABLE 1 Type of specimens and the type of patient from which the 354 *K. pneumoniae* isolates have been isolated[a]

Criterion of selection	Number of samples for the criterion	Inpatients	Outpatients	Type of specimen	Number of samples	Wild type	ESBL	KPC	NDM	VIM	Other
String test positive	21	13	8	Urine	13	7	5	1	NA[b]	NA	NA
				Blood culture	3	3	NA	NA	NA	NA	NA
				Respiratory samples	2	2	NA	NA	NA	NA	NA
				Liver biopsy	1	1	NA	NA	NA	NA	NA
				Rectal swab	1	NA	NA	1	NA	NA	NA
				Urethral swab	1	NA	NA	1	NA	NA	NA
String test negative and MDR	225	221	4	Rectal swab	160	NA	8	144	4	1	3 (2 KPC + VIM; 1 KPC + NDM)
				Respiratory samples	30	NA	6	24	NA	NA	NA
				Urine	25	NA	2	23	NA	NA	NA
				Wound/biopsy/dreinage samples	6	NA	NA	5	1	NA	NA
				Other	4	NA	NA	3	1	NA	NA
Blood cultures	108	80	28	Blood culture	108	57	19	30	1	NA	1 KPC + VIM

[a]The resistance profile of the isolates is also shown.
[b] NA was used to indicate that zero strains have that characteristic.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

TABLE 2 Characteristics of the 20 patients from whom *Klebsiella pneumoniae* with increased virulence was isolated[a]

Strain number	Isolation date	Sample	Sex	Age	Hospital ward	Outcome	String test	Resistance mechanisms	PCR *rmpA*	PCR *iutA*	PCR *iroN*	PCR *peg-344*
5980	1May 2021	Rectal swab	F	74	General medicine	NA	–	KPC	–	+	–	+
5985	1 May 2021	Bronchial aspirate	F	63	Intensive care unit	Recovered	–	KPC	–	+	–	+
5999	6 May 2021	Bronchoalveolar lavage	M	74	Intensive care unit	Deceased	–	KPC	–	+	–	+
6002	8 May 2021	Rectal swab	M	52	Intensive care unit	NA	–	KPC	–	+	–	+
6007	12 May 2021	Rectal swab	M	56	Hematology	NA	–	KPC	–	+	+	–
6008	13 May 2021	Rectal swab	M	59	General medicine	NA	–	KPC	–	+	–	–
6082	8 June 2021	Liver biopsy	F	65	Intensive care unit	Deceased	+	Wild type	+	+	+	+
6114	22 June 2021	Urine	M	81	Infectious diseases (outpatients)	NA	+	KPC	–	+	–	–
hvKp_2	3 Augurst 2021	Bronchoalveolar lavage	M	44	Intensive care unit	Recovered	+	Wild type	+	+	+	+
hvKp_4	29 August 2021	Blood culture	F	71	Obstetrics and gynecology	Recovered	+	Wild type	+	+	+	+
6311	25 September 2021	Rectal swab	M	77	Intensive care unit	NA	–	KPC	–	+	+	–
6337	27 September 2021	Dreinage	F	49	Obstetrics and gynecology	Recovered	–	NDM	–	+	+	+
6452	21 October 2021	Blood culture	M	53	Intensive care unit	Recovered	+	Wild type	+	+	+	+
6399	27 October 2021	Rectal swab	F	75	Pneumology	NA	+	KPC	–	+	+	–
6403	27 October 2021	Rectal swab	F	76	Emergency room	NA	–	KPC	–	+	+	–
6445	11 November 2021	Rectal swab	M	37	Otolaryngology	NA	–	NDM	+	+	+	+
6465	12 November 2021	Urethral swab	F	76	Pneumology	Recovered	+	KPC	–	+	–	–
30715	24 November 2021	Blood culture	M	71	Intensive care unit	Deceased	+	Wild type	–	+	–	+
hvKp_19	24 November 2021	Urine	F	80	Outpatient sample collection center	NA	+	Wild type	+	+	+	+
6843	14 April 2022	Rectal swab	F	54	Pneumology	NA	–	KPC	–	+	+	–

[a]Outcome for the patients colonized by *K. pneumoniae* with increased virulence is indicated with NA (Not Applicable).

(*N* = 94) to those in the data set from the BV-BRC database (https://www.bv-brc.org/) according to k-mer content similarity (MASH). The P-DOR pipeline (https://github.com/SteMIDIfactory/P-DOR) was used to align all genomes in the data set to a reference (573.4026 in BV-BRC) and extract coreSNPs. Maximum likelihood phylogeny was inferred with RAxML (using 100 bootstrap resamples) on the resulting coreSNP alignment using the general time reversible model, as suggested by ModelTest-NG, with ascertainment bias correction (23).

RESULTS

K. pneumoniae was isolated from 627 patients during the study period. A total of 354 *K. pneumoniae* isolates were considered for this study, including 21 isolates with positive string test, 225 MDR isolates with negative string test, and 108 wild-type isolates from blood cultures with negative string test.

Twenty out of 354 isolates (5.6%) had at least one target gene of *rmpA*, *iutA*, *iroN*, and *peg-344* gene amplified, and six isolates had all four genes amplified. The characteristics

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

TABLE 3 Characteristics of the genomes' virulence score are attributed by Kleborate as follows: 0 = no yersinabactin, colibactin, or aerobactin; 1 = yersiniabactin only; 2 = yersiniabactin and colibactin (or colibactin only); 3 = aerobactin without yersiniabactin or colibactin; 4 = aerobactin with yersiniabactin (no colibactin); 5 = yersiniabactin, colibactin, and aerobactin[a]

Strain number	Sample	contig_count	N50	largest_contig	total_size	ST	virulence_score	resistance_score
5980	Rectal swab	163	203938	449385	5704928	ST512	0	2
5985	Bronchial aspirate	161	161807	544125	5712389	ST512	0	2
5999	Bronchoalveolar lavage	148	203931	449816	5708277	ST512	0	2
6002	Rectal swab	184	109995	334526	5847712	ST395	3	2
6007	Rectal swab	275	149647	331468	5702544	ST512	0	2
6008	Rectal swab	118	157687	405925	5560881	ST512	1	3
6082	Liver biopsy	100	195392	498479	5445960	ST375	3	0
6114	Urine	140	158510	449810	5601209	ST512	0	2
hvKp_2	Bronchoalveolar lavage	115	116031	303700	5590050	ST6310	3	0
hvKp_4	Blood culture	85	316230	436906	5473727	ST86	4	0
6311	Rectal swab	131	202114	426621	5757831	ST101	4	3
6337	Dreinage	179	219779	823629	5748271	ST147	3	2
6452	Blood culture	116	246519	788521	5655516	ST23	5	0
6399	Rectal swab	202	136210	489606	5939332	ST512	3	2
6403	Rectal swab	164	205081	380663	5818972	ST101	4	3
6445	Rectal swab	228	173045	386555	5930012	ST147	4	2
6465	Urethral swab	155	201170	523991	5941613	ST512	3	2
30715	Blood culture	109	244625	449964	5425355	ST727	1	0
hvKp_19	Urine	61	323151	680484	5437774	ST5	4	0
6843	Rectal swab	172	205082	520390	5820568	ST101	4	3

[a]Resistance score is attributed by Kleborate as follows: 1 = ESBL; 2 = carbapenemase; 3 = carbapenemase plus colistin resistance; 0 no resistance mechanism to beta-lactams.

of the 20 patients from whom the 20 isolates were obtained are shown in Table 2. The median age of the 20 patients was 64.35 years (range: 37–81), and 11 were female.

Nine strains (2.54%) were isolated from rectal swabs and were resistant to carbapenems, with eight strains bearing *blaKPC* and one bearing *blaNDM* encoding for the New Delhi metallo-beta-lactamase (NDM). Only one isolate from rectal swabs was string test positive. The remaining 11 strains were isolated mainly from blood cultures (n = 3; 0.8%), bronchoalveolar lavage (BAL) (n = 2; 0.56%), and urine (n = 2; 0.56%). Four of 11 strains had KPC carbapenemase and one strain had NDM. The characteristics of all the 20 genomes are shown in Table 3.

Kleborate analyses showed that 13 (3.64%) strains had virulence score above 3 and seven of them were string test positive (Table 4). One strain isolated from a blood culture belonged to ST23 and had the highest virulence score. This strain is the only one in this study displaying aerobactin, salmochelin, yersiniabactin, and colibactin.

Eight strains meet the definition of convergent antimicrobial resistant and virulent strains as proposed by Lam et al. (19), having virulence score ≥3 and resistance score ≥1 (Table 3). Three strains belonging to ST101 isolated from rectal swabs and one strain ST147 isolated from a rectal swab had a virulence score of 4 and had aerobactin, salmochelin, and yersiniabactin. The three strains ST101 were KPC producing and the strain ST147 was NDM producing. The other four convergent antimicrobial resistant and virulent strains had a virulence score of 3 and belonged to ST147, ST395, and ST512.

A phylogeny that included the two strains ST147 isolated in this study and the 181 ST147 strains from the outbreak that occurred in Tuscany (21, 22) is shown in Fig. 1. One of the two strains of this study (SanMatteo_6445) clustered in a monophyletic clade with three other strains from the study by Martin et al. (21) and the common branch is highly supported (100/100 bootstraps). Evaluation of clonal relatedness by core-genome SNPs reveals that clones were overall highly related to the cluster (SNP range: 19–20). As for other strains in the same cluster, the FIB(pQil)-type plasmid (pQil-NDM-147Tu; 56,064 bp; GenBank accession number CP071030) was identified. This plasmid carried the bla_{NDM-1} gene, along with other resistant determinants such as *aacA4*-cr, bla_{OXA-1}, and $bla_{CTX-M-15}$. In addition, the HIB-FIB(Mar) plasmid (pVIR-147Tu; 341,914 pb; GenBank

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

TABLE 4 Characteristics of the 20 strains that were sequenced[a]

Strain	String test	Resistance mechanisms	ST	Virulence score	Yersiniabactin	Colibactin	Aerobactin	Salmochelin	RmpADC	rmpA2
5980	–	KPC	ST512	0	–	–	–	–	–	–
5985	–	KPC	ST512	0	–	–	–	–	–	–
5999	–	KPC	ST512	0	–	–	–	–	–	–
6002	–	KPC	ST395	3	–	–	iuc 1	–	–	–
6007	–	KPC	ST512	0	–	–	–	–	–	–
6008	–	KPC	ST512	1	ybt 9; ICEKp3	–	–	–	–	–
6082	+	Wild Type	ST375	3	–	–	iuc 1	iro 1	rmp 1; KpVP-1	rmpA2_3-47%
6114	+	KPC	ST512	0	–	–	–	–	–	–
hvKp_2	+	Wild Type	ST6310	3	–	–	iuc 1	iro 1	rmp 1; KpVP-1 (truncated)	–
hvKp_4	+	Wild Type	ST86	4	ybt 9; ICEKp3	–	iuc 1	iro 1	rmp 1; KpVP-1	rmpA2_9
6311	–	KPC	ST101	4	ybt 9; ICEKp3	–	iuc 1	–	–	–
6337	–	NDM	ST147	3	–	–	iuc 1	–	rmp 1; KpVP-1	rmpA2_6*–47%
6452	+	Wild Type	ST23	5	ybt 1; ICEKp10	clb 2	iuc 1	iro 1	rmp 1; KpVP-1	rmpA2_5-54%
6399	+	KPC	ST512	3	–	–	iuc 1	–	–	–
6403	–	KPC	ST101	4	ybt 9; ICEKp3	–	iuc 1	–	–	–
6445	–	NDM	ST147	4	ybt 9; ICEKp3	–	iuc 1	–	rmp 1; KpVP-1	rmpA2_6*–47%
6465	+	KPC	ST512	3	–	–	iuc 1	–	–	–
30715	+	Wild Type	ST727	1	ybt 4; plasmid (incomplete)	–	–	iro 1	rmp unknown	–
hvKp_19	+	Wild Type	ST5	4	ybt 2; ICEKp1	–	iuc 3 (truncated)	iro 3 (truncated)	rmp 3; ICEKp1 (truncated)	–
6843	–	KPC	ST101	4	ybt 9; ICEKp3	–	iuc 1	–	–	–

[a]Resistance mechanism, sequence type, and the presence of virulence loci are displayed. A negative string test is indicated with "–", while a positive one is indicated with "+".

accession number CP071028) was found. This one carried several resistant determinants such as *armA*, *sul1* and *sul2*, $bla_{CTX\text{-}M\text{-}15}$, and virulence ones such as *rmpADC*, *iucABCD-iutA*, and *peg-344*. On the contrary, the other strain (SanMatteo_6337) clustered with high branch support (94/100 bootstraps) with four strains from the BV-BRC DB isolated in Russia between 2018 and 2019 (SNP range: 41–104). The SanMatteo_6337 strain did not carry the FIB(pQil)-type plasmid and carried the HIB-FIB(Mar)-type plasmid, instead.

DISCUSSION

This study reports the isolation of 13 isolates of hypervirulent *K. pneumoniae,* based on sequence type and the presence of virulence genes such as aerobactin and salmochelin. Six out of 13 strains were isolated from rectal swab (colonization), while seven strains caused infections.

The 13 strains had different characteristics. Two strains belonged to ST23 and hvKp ST86, respectively, and caused bloodstream infections in two patients who had not recently traveled abroad. hvKp ST23 was reported for the first time in Italy in 2014 (24), and in our hospital, it was isolated from urine and BAL samples between 2017 and 2018 (25). However, this sequence type is very uncommon in Italy and only five genomes of *K. pneumoniae* ST23 were found in the BV-BRC database as isolated in Italy, including the 6453 isolate of this study and the isolate from a BAL sample of the same patient (last accessed on 30 March 2023). ST23 strain isolated in this study was wild type but hvKp

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

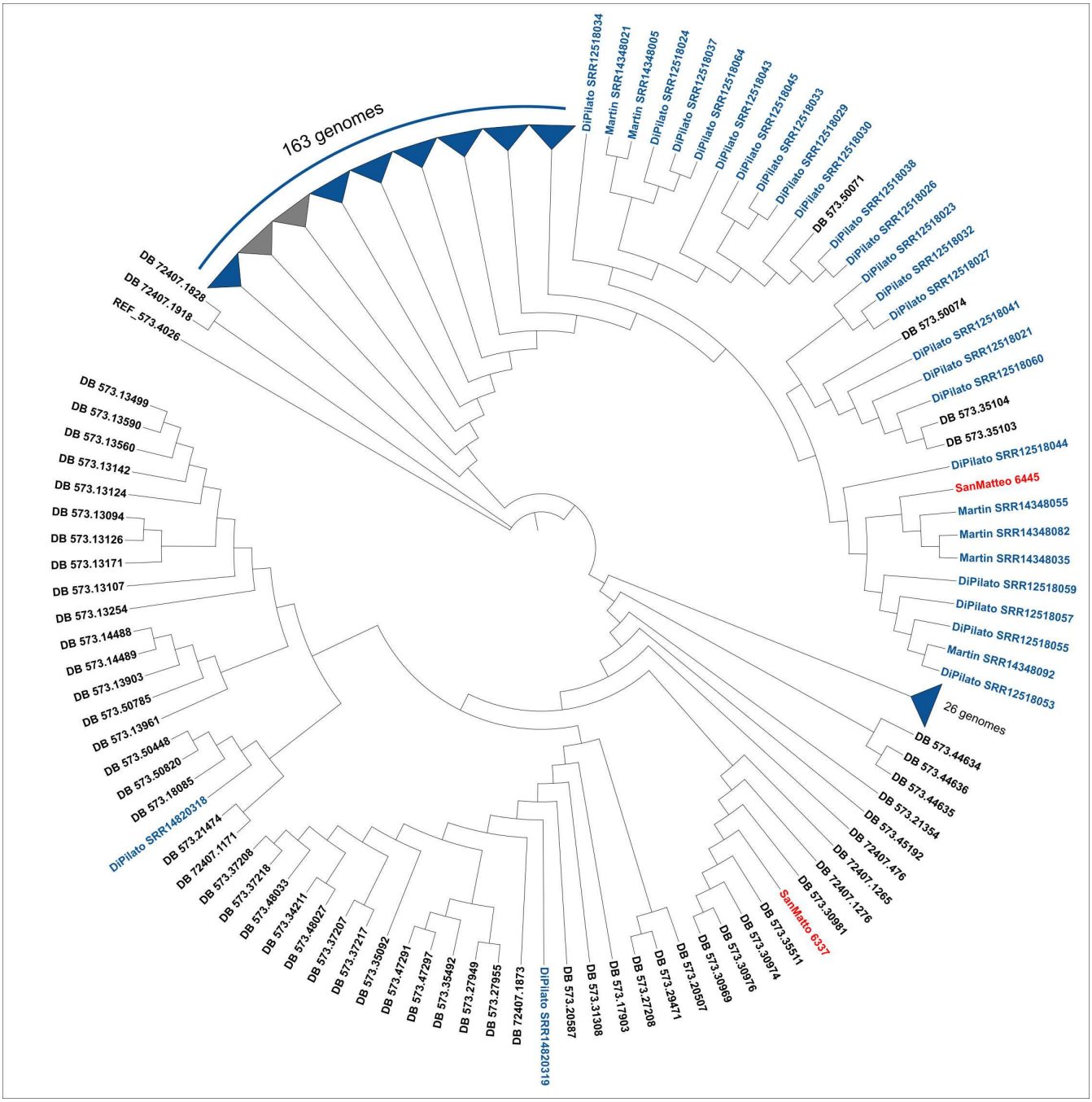


FIG 1 Maximum likelihood phylogeny of *Klebsiella pneumoniae* ST147 based on SNP cores obtained using the P-DOR pipeline (https://github.com/SteMIDIfactory/P-DOR). The two strains of this study are indicated in red, while 181 genomes ST147 from the outbreak that occurred in Tuscany reported by Martin et al. (21) and Di Pilato et al. (22) are in blue.

ST23 bearing carbapenemases were described in several European countries also from blood and respiratory samples (3, 24).

Equally rare is the ST86, which presence is documented in Italy so far only by the presence in BV-BRC databases of three genomes of *K. pneumoniae* ST86, including strain hvKp_4 of this study (reported in the BV_BRC database as 10028588). In Europe, ST86 hvKp strains have already been reported in Spain where three isolates out of 878 were found to be hypervirulent ST86 in a 7-year study (2) and in France where three out

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

of 59 *K. pneumoniae* isolates were hypervirulent ST86 in a 5-year study (3). In France, carbapenem-resistant hvKp bearing OXA-48 have been isolated as well (26).

Other sequence types frequently associated with hypervirulence, such as ST5 and ST375, were found in this study. ST5 was previously isolated in Italy in a multicentric study (27). This strain, as well as hvKp_19 of this study, bears virulence factors such as ICEKp1, aerobactin, salmochelin, and rmpADC hypermucoidy locus. Only four genomes belonging to ST375 were found in the BV-BRC database as isolated in Europe, none in Italy. Hypervirulent ST375 was previously isolated in Japan (28) from liver abscesses of two patients. The ST375 isolate of this study was isolated from a liver biopsy, too, and as the isolates from Japan, it carried aerobactin and *rmpA* genes.

The new sequence type ST6310, a single locus variant of ST25 for *mdh* gene, was described for the first time in this study. Strain hvKp_2, which belonged to ST6310, had aerobactin, salmochelin, and *rmpA* loci and it was isolated from a patient who had lived in China.

hvKp isolates belonging to well-known MDR high-risk clones such as ST101, ST147, ST395, and ST512 were also found in this study. hvKp are usually susceptible to most antibiotics, but the ability of *Klebsiella* to acquire genetic elements leads to the emergence of MDR-hvKp clones (4, 13). Most MDR strains of this study did not have multiple virulence genes, except for three strains belonging to ST101 and one from ST147 that bore a wider range of virulence genes. The latter ST147 strain resulted closely related to the strains that caused a large hospital outbreak in the Tuscany region of Italy. Indeed, this isolate along with the strains described in the study of Di Pilato et al. are part of subclade of ST147, called ST147-vir, characterized by the presence of the *ybt*-encoding element ICEKpn3 and the pQil-NDM-147Tu-like and the pVIR-147Tu-like plasmids (22). The sequence type 147 is a high-risk clone associated with hospital-acquired infections worldwide and mediating the global spread of NDM-like metallo-beta-lactamase genes (29). Therefore, it is crucial to track the dissemination of these potential successful clones in hospital settings and understand their evolution to avoid the spread of resistant and virulent determinants, especially those determinants carried on plasmids, such as bla_{NDM-1}.

In conclusion, this study described the isolation of 13 isolates of *K. pneumoniae* with increased virulence in a large tertiary hospital over a 1-year period. Although the isolation of these strains is not frequent, their presence is particularly concerning, as some of these isolates are not only hypervirulent but also MDR, effectively limiting the therapeutic options in case of infections.

Not all the hypervirulent isolates were positive by the string test, so hvKp are not always phenotypically detectable. Molecular biology techniques such as PCR amplification and next generation sequencing are therefore necessary for the detection of hvKp isolates. Surveillance programs exploiting PCR for the detection of hvKp isolates are desirable also because virulence genes are often carried on mobile genetic elements, which can be transferred within and between different bacterial species. In addition, the presence of hvKp should be systematically investigated in both wild-type and MDR strains, to avoid the spreading of strains with both increased virulence and resistance to antibiotics. Finally, hvKp strains were also isolated from patients with no history of recent travels abroad, suggesting an increased prevalence or an undetected circulation of hvKp isolates.

ACKNOWLEDGMENTS

Conceptualization: C.M. and M.C.; formal analysis: G.P. and S.G.; investigation: A.K., I.M., D.D.V., and M.A.; supervision: M.C. and P.C.; writing—original draft: C.M. and S.G; writing—review and editing: F.B. and P.C.

All authors have read and agreed to the published version of the manuscript.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

AUTHOR AFFILIATIONS

[1]Struttura Complessa di Microbiologia e Virologia, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[2]Scuola di Specializzazione in Microbiologia e Virologia, Università degli Studi di Pavia, Pavia, Italy
[3]Dipartimento di Scienze Clinico-Chirurgiche, Diagnostiche e Pediatriche, Università degli Studi di Pavia, Pavia, Italy

AUTHOR ORCIDs

C. Merla http://orcid.org/0000-0003-2986-6644
S. Gaiarsa http://orcid.org/0000-0002-5847-6568

AUTHOR CONTRIBUTIONS

C. Merla, Conceptualization, Writing – original draft | A. Kuka, Investigation | I. Mileto, Investigation | G. Petazzoni, Formal analysis | S. Gaiarsa, Formal analysis, Writing – original draft | D. De Vitis, Investigation | M. Ardizzone, Investigation | M. Corbella, Conceptualization, Supervision | F. Baldanti, Writing – review and editing | P. Cambieri, Supervision, Writing – review and editing

DATA AVAILABILITY

Genome assembly data are available at NCBI under BioProject ID PRJNA814467.

ETHICS APPROVAL

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105).

REFERENCES

1. Liu YC, Cheng DL, Lin CL. 1986. *Klebsiella pneumoniae* liver abscess associated with septic endophthalmitis. Arch Intern Med 146:1913–1916. https://doi.org/10.1001/archinte.1986.00360220057011
2. Cubero M, Grau I, Tubau F, Pallarés R, Dominguez MA, Liñares J, Ardanuy C. 2016. Hypervirulent *Klebsiella pneumoniae* clones causing bacteraemia in adults in a teaching hospital in Barcelona, Spain (2007-2013). Clin Microbiol Infect 22:154–160. https://doi.org/10.1016/j.cmi.2015.09.025
3. Rafat C, Messika J, Barnaud G, Dufour N, Magdoud F, Billard-Pomarès T, Gaudry S, Dreyfuss D, Branger C, Decré D, et al. 2018. Hypervirulent *Klebsiella pneumoniae*, a 5-year study in a French ICU. J Med Microbiol 67:1083–1089. https://doi.org/10.1099/jmm.0.000788
4. Russo TA, Marr CM. 2019. Hypervirulent *Klebsiella pneumoniae*. Clin Microbiol Rev 32:e00001-19. https://doi.org/10.1128/CMR.00001-19
5. Parrott AM, Shi J, Aaron J, Green DA, Whittier S, Wu F. 2021. Detection of multiple hypervirulent *Klebsiella pneumoniae* strains in a New York city hospital through screening of virulence genes. Clin Microbiol Infect 27:583–589. https://doi.org/10.1016/j.cmi.2020.05.012
6. Harada S, Doi Y. 2018. Hypervirulent *Klebsiella pneumoniae*: a call for consensus definition and international collaboration. J Clin Microbiol 56:e00959-18. https://doi.org/10.1128/JCM.00959-18
7. Kumabe A, Kenzaka T. 2014. String test of hypervirulent *Klebsiella pneumonia*. QJM 107:1053. https://doi.org/10.1093/qjmed/hcu124
8. Russo TA, Olson R, Fang C-T, Stoesser N, Miller M, MacDonald U, Hutson A, Barker JH, La Hoz RM, Johnson JR. 2018. Identification of biomarkers for differentiation of hypervirulent *Klebsiella pneumoniae* from classical *K. pneumoniae*. J Clin Microbiol 56:e00776-18. https://doi.org/10.1128/JCM.00776-18
9. Catalán-Nájera JC, Garza-Ramos U, Barrios-Camacho H. 2017. Hypervirulence and hypermucoviscosity: two different but complementary *Klebsiella* spp. phenotypes. Virulence 8:1111–1123. https://doi.org/10.1080/21505594.2017.1317412
10. Yu F, Lv J, Niu S, Du H, Tang YW, Pitout JDD, Bonomo RA, Kreiswirth BN, Chen L. 2018. Multiplex PCR analysis for rapid detection of *Klebsiella pneumoniae* carbapenem-resistant (sequence type 258 [ST258] and ST11) and hypervirulent (ST23, ST65, ST86, and ST375) strains. J Clin Microbiol 56:e00731-18. https://doi.org/10.1128/JCM.00731-18
11. Liao W, Long D, Huang Q, Wei D, Liu X, Wan L, Feng Y, Zhang W, Liu Y. 2020. Rapid detection to differentiate hypervirulent *Klebsiella pneumoniae* (hvKp) from classical *K. pneumoniae* by identifying *peg-344* with loop-mediated isothermal amplication (LAMP). Front Microbiol 11:1189. https://doi.org/10.3389/fmicb.2020.01189
12. Bulger J, MacDonald U, Olson R, Beanan J, Russo TA. 2017. Metabolite transporter PEG344 is required for full virulence of hypervirulent *Klebsiella pneumoniae* strain hvKP1 after pulmonary but not subcutaneous challenge. Infect Immun 85:e00093-17. https://doi.org/10.1128/IAI.00093-17
13. Gu D, Dong N, Zheng Z, Lin D, Huang M, Wang L, Chan EW-C, Shu L, Yu J, Zhang R, et al. 2018. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. Lancet Infect Dis 18:37–46. https://doi.org/10.1016/S1473-3099(17)30489-9
14. Yao B, Xiao X, Wang F, Zhou L, Zhang X, Zhang J. 2015. Clinical and molecular characteristics of multi-clone carbapenem-resistant hypervirulent (hypermucoviscous) *Klebsiella pneumoniae* isolates in a tertiary hospital in Beijing, China. Int J Infect Dis 37:107–112. https://doi.org/10.1016/j.ijid.2015.06.023
15. Zhang R, Lin D, Chan EW-C, Gu D, Chen G-X, Chen S. 2016. Emergence of carbapenem-resistant serotype K1 hypervirulent *Klebsiella pneumoniae* strains in China. Antimicrob Agents Chemother 60:709–711. https://doi.org/10.1128/AAC.02173-15
16. Brolund A, Lagerqvist N, Byfors S, Struelens MJ, Monnet DL, Albiger B, Kohlenberg A, European Antimicrobial Resistance Genes Surveillance

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

Network (EURGen-Net) capacity survey group. 2019. Worsening epidemiological situation of carbapenemase-producing Enterobacteriaceae in Europe, assessment by national experts from 37 countries, July 2018. Euro Surveill 24:1900123. https://doi.org/10.2807/1560-7917.ES.2019.24.9.1900123
17. Tang HL, Chiang MK, Liou WJ, Chen YT, Peng HL, Chiou CS, Liu KS, Lu MC, Tung KC, Lai YC. 2010. Correlation between *Klebsiella pneumoniae* carrying pLVPK-derived loci and abscess formation. Eur J Clin Microbiol Infect Dis 29:689–698. https://doi.org/10.1007/s10096-010-0915-1
18. Sanikhani R, Moeinirad M, Solgi H, Hadadi A, Shahcheraghi F, Badmasti F. 2021. The face of hypervirulent *Klebsiella pneumoniae* isolated from clinical samples of two Iranian teaching hospitals. Ann Clin Microbiol Antimicrob 20:58. https://doi.org/10.1186/s12941-021-00467-2
19. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. 2021. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nat Commun 12:4188. https://doi.org/10.1038/s41467-021-24448-3
20. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14
21. Martin MJ, Corey BW, Sannio F, Hall LR, MacDonald U, Jones BT, Mills EG, Harless C, Stam J, Maybank R, et al. 2021. Anatomy of an extensively drug-resistant *Klebsiella pneumoniae* outbreak in Tuscany, Italy. Proc Natl Acad Sci USA 118:e2110227118. https://doi.org/10.1073/pnas.2110227118
22. Di Pilato V, Henrici De Angelis L, Aiezza N, Baccani I, Niccolai C, Parisio EM, Giordano C, Camarlinghi G, Barnini S, Forni S, et al. 2022. Resistome and virulome accretion in an NDM-1-producing ST147 sublineage of *Klebsiella pneumoniae* associated with an outbreak in Tuscany, Italy: a genotypic and phenotypic characterisation. Lancet Microbe 3:e224–e234. https://doi.org/10.1016/S2666-5247(21)00268-8
23. Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925. https://doi.org/10.1080/106351501753462876
24. 2021. European centre for disease prevention and control. emergence of Hypervirulent Klebsiella pneumoniae St23 carrying Carbapenemase genes in EU/EEA countries. ECDC: Stockholm.
25. Thorpe HA, Booton R, Kallonen T, Gibbon MJ, Couto N, Passet V, López-Fernández S, Rodrigues C, Matthews L, Mitchell S, et al. 2022. A large-scale genomic snapshot of *Klebsiella* spp. isolates in Northern Italy reveals limited transmission between clinical and non-clinical settings. Nat Microbiol 7:2054–2067. https://doi.org/10.1038/s41564-022-01263-0
26. Beyrouthy R, Dalmasso G, Birer A, Robin F, Bonnet R. 2020. Carbapenem resistance conferred by OXA-48 in K2-ST86 hypervirulent *Klebsiella pneumoniae*, France. Emerg Infect Dis 26:1529–1533. https://doi.org/10.3201/eid2607.191490
27. Arena F, Menchinelli G, Di Pilato V, Torelli R, Antonelli A, Henrici De Angelis L, Coppi M, Sanguinetti M, Rossolini GM. 2022. Resistance and virulence features of hypermucoviscous *Klebsiella pneumoniae* from bloodstream infections: results of a nationwide Italian surveillance study. Front Microbiol 13:983294. https://doi.org/10.3389/fmicb.2022.983294
28. Hoashi K, Harada S, Ishii Y, Aoki K, Ishikawa S, Oshiro Y, Shinzato T. 2019. Community-acquired liver abscess caused by capsular genotype K2-ST375 hypervirulent *Klebsiella pneumoniae* isolates. IDCases 17:e00577. https://doi.org/10.1016/j.idcr.2019.e00577
29. Lapp Z, Crawford R, Miles-Jay A, Pirani A, Trick WE, Weinstein RA, Hayden MK, Snitkin ES, Lin MY. 2021. Regional spread of blaNDM-1-containing *Klebsiella pneumoniae* ST147 in post-acute care facilities. Clin Infect Dis 73:1431–1439. https://doi.org/10.1093/cid/ciab457

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

Rapid communication

Rapid spread of a novel NDM-producing clone of *Klebsiella pneumoniae* CC147, Northern Italy, February to August 2023

Irene Mileto[1,2,*], Greta Petazzoni[1,3,*], Marta Corbella[1], Stefano Gaiarsa[1], Cristina Merla[1], Angela Kuka[1], Marina Ramus[1,2], Cristina Terulla[4], Micaela Brandolini[5], Antonio Piralla[1], Patrizia Cambieri[1,], Fausto Baldanti[1,3,**]**

1. Microbiology and Virology Unit, IRCCS Fondazione Policlinico San Matteo, Pavia, Italy
2. Specialization School of Microbiology and Virology, University of Pavia, Pavia, Italy
3. Department of Clinical, Surgical, Diagnostic and Paediatric Sciences, University of Pavia, Pavia, Italy
4. ASST Pavia, Ospedale Civile di Voghera, Voghera, Italy
5. ASST Pavia, Ospedale Unificato Broni-Stradella, Internal Medicine Unit, Stradella, Italy

* These authors have contributed equally
** These authors contributed equally to the work and share the last authorship

Correspondence: Marta Corbella (m.corbella@smatteo.pv.it)

Citation style for this article:
Mileto Irene, Petazzoni Greta, Corbella Marta, Gaiarsa Stefano, Merla Cristina, Kuka Angela, Ramus Marina, Terulla Cristina, Brandolini Micaela, Piralla Antonio, Cambieri Patrizia, Baldanti Fausto. Rapid spread of a novel NDM-producing clone of Klebsiella pneumoniae CC147, Northern Italy, February to August 2023. Euro Surveill. 2023;28(42):pii=2300522. https://doi.org/10.2807/1560-7917.ES.2023.28.42.2300522

Article submitted on 01 Oct 2023 / accepted on 18 Oct 2023 / published on 19 Oct 2023

New Delhi metallo-beta-lactamase (NDM)-producing *Klebsiella pneumoniae* (*Kp*) ST147 caused a large multi-hospital outbreak in Italy from 2018 to 2021. We describe a new ST6668 NDM-producing *Kp* clone, belonging to CC147, which rapidly spread across hospitals in the Pavia province (Northern Italy) from February to August 2023. Genomic analyses revealed that ST6668 is different from ST147 and fast evolving. As shown here, genomic surveillance programmes are useful for tracking the spread of new clones with reduced susceptibility to most antibiotics.

The spread of New Delhi metallo-beta-lactamase (NDM)-producing *Klebsiella pneumoniae* (*Kp*) is of potential public health concern because these strains are transmitted in healthcare settings and are resistant to nearly all antibiotics. Cefiderocol or ceftazidime-avibactam in combination with aztreonam are the preferred antibiotic options for NDM-producing Enterobacterales [1].

Several sequence types (ST) of *Kp* worldwide, such as ST11 and ST147, have been reported to produce NDM carbapenemase. In particular, ST147 (and the clonal complex (CC) associated to it) is a high-risk international clone [2] that has caused a large outbreak in Italy (Tuscany region) in the period 2018 to 2021 [3,4]. Here we describe the isolation of a new ST of NDM-producing *Kp* (ST6668) belonging to CC147 that caused a multi-hospital outbreak from February to August 2023 in the Pavia province (Northern Italy).

Event description

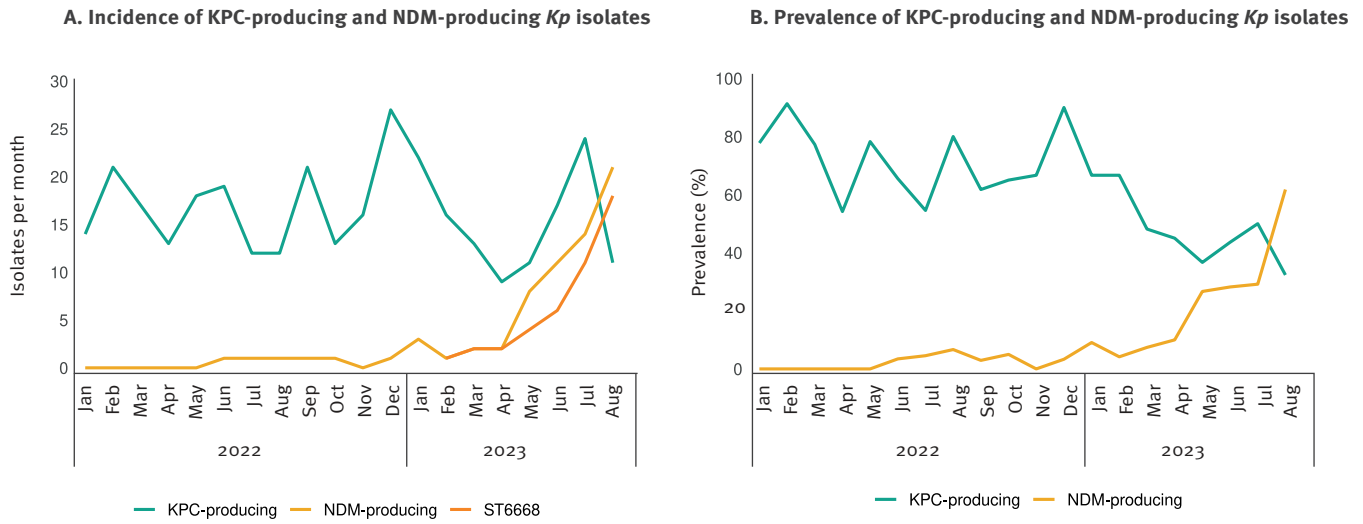
Until 2022, *K. pneumoniae* carbapenemase (KPC) has been the most detected carbapenemase in *Kp* at Fondazione IRCCS Policlinico San Matteo in Pavia (OSM), a 900-bed teaching hospital. The KPC incidence remained constant from August 2022 to August 2023 (16.3 ± 5.6 isolates/month), while the NDM incidence increased from April 2023 (Figure 1A). Indeed, among the carbapenem-resistant *Kp* isolates, the NDM prevalence reached 61.8% in August from 10% in April 2023 (Figure 1B).

From January to August 2023, 117 isolates of NDM-producing *Kp* were collected at OSM from 66 patients. Considering the first isolate of NDM-producing *Kp* for each patient, we performed whole genome sequencing of 66 isolates, using Illumina reads as in [5]. In addition, we sequenced the genomes of eight NDM-producing *Kp* isolated from patients admitted to Ospedale Civile di Voghera, a 200-bed hospital in the Pavia province (ca 30 km from Pavia), belonging to the Pavia Health Authority Unit (Pavia-HAU). Genomic analysis of all 74 isolates showed that 52 of them belonged to the same clone, which presented a novel MLST allelic profile. The profile was submitted to the BIGSdb service (https://bigsdb.pasteur.fr) and confirmed as a new clone, named ST6668. This new clone has since February 2023 been responsible for the rapid rise in NDM-producing *Kp* as shown in Figure 1A .

Of the 52 *Kp* ST6668 strains, 44 had been collected from patients in OSM and eight in Voghera hospital. Forty-four of them were isolated from rectal swabs

Figure 1

KPC and NDM-producing *Klebsiella pneumoniae* isolates, Fondazione IRCCS Policlinico San Matteo, Pavia, Northern Italy, January 2022–August 2023 (n = 394)



Kp: Klebsiella pneumoniae; KPC: *Kp* carbapenemase; NDM: New Delhi metallo-beta-lactamase.

and urine samples, five from blood cultures, and the remaining three from bronchoalveolar lavage, pus and drainage liquid. The patients had a mean age of 70 years (SD: ± 15.4), ranging from 31 to 90 years, and 31 of the 52 were male. At OSM, 13 patients were admitted to specialist medical care (pneumology, nephrology/rheumatology and infectious diseases), 10 to oncology wards, seven to intensive care units, five to surgery wards, three to general medicine, while the remaining six were outpatients. At OSM, 42 patients were screened for multidrug-resistant (MDR) Enterobacterales colonisation via rectal swab, as foreseen by internal procedures for high-risk patients and those hospitalised on critical wards. Of these, 36 were colonised at admission, while the remaining six became colonised during hospitalisation (mean value: 15 days after admission). Seven patients developed bloodstream or pulmonary infections. Moreover, three of the eight Voghera patients were later transferred to a small hospital of the Pavia-HAU (20 km from Pavia and Voghera).

Clone characterisation

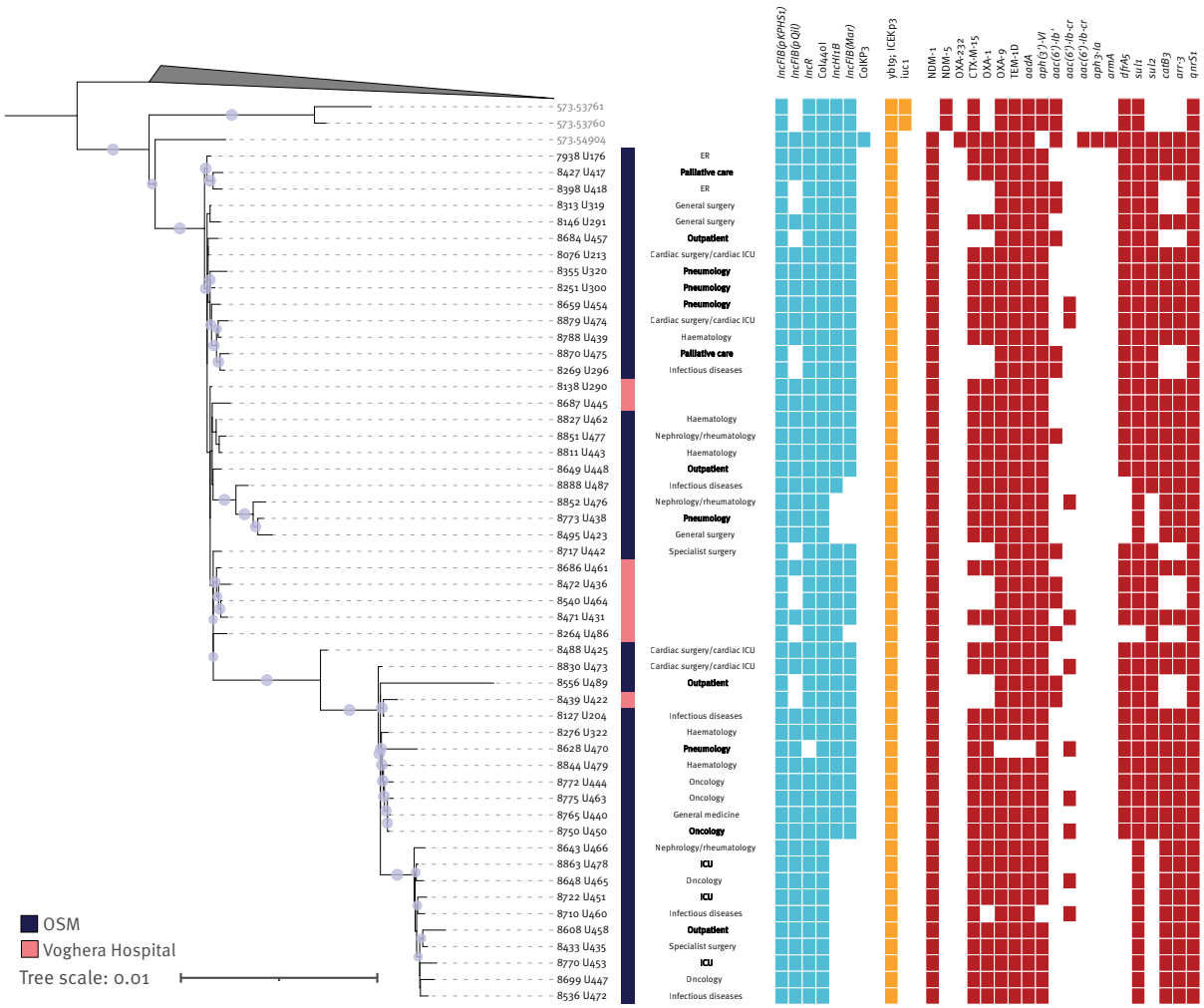
Antimicrobial susceptibility testing was performed using Sensititre DKMGN broth microdilution (ThermoFisher) and interpreted according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints [6]. All 52 isolates were resistant to most antibiotics including carbapenems, ceftazidime-avibactam and amikacin, while only two were resistant to cefiderocol. All isolates remained susceptible to gentamicin. The minimal inhibitory concentrations (MIC) are listed in the Supplementary Table. The combination of ceftazidime-avibactam with aztreonam had a synergistic effect on all the strains [7].

Illumina short reads were de novo assembled with Shovill (github.com/tseemann/shovill). We used P-DOR [8] to obtain the single nucleotide polymorphism (SNP)-based maximum-likelihood phylogeny of the 52 isolates described in this work and the 38 most similar high-quality genomes available from the BV-BRC database (bv-brc.org) (Figure 2). The 52 genomes from the Pavia area resulted to be monophyletic (SNP range: 4–51), suggesting that all strains were part of the same epidemic event. All outbreak isolates belonged to ST6668, which differs from ST147 by a point mutation (C414T) in the *phoE* locus; whereas the 38 genomes used as background belonged to ST147. The three closest genomes to the outbreak clade were isolated in Lebanon from blood (in 2020) and in the United States from urine and rectal swabs (in 2022). Moreover, none of the closest strains retrieved in the phylogeny was the ST147 clone responsible for the large outbreak in Tuscany [4]. This suggests that the ST6668 clone originated from a separate importation of CC147.

We used Kleborate [9] to search for determinants of antibiotic resistance and virulence, and PlasmidFinder [10] to determine the plasmid presence/absence profile. The Supplementary Table add information on genomic resistance determinants, virulence factors and plasmids presence. All outbreak strains harboured the bla_{NDM-1} gene conferring carbapenem resistance. In addition, all strains presented allele 9 of the *ybt* gene, encoding yersiniabactin, enclosed in the ICE*Kp*3 integrative island. Six plasmid sequences were detected in the outbreak. Two of them were common to all 52 isolates (*IncFIB(pKPHS1)*, *Col440I*), while the other four presented a presence/absence pattern, which

Figure 2

Phylogenetic analysis of carbapenemase-producing *Klebsiella pneumoniae* ST6668 isolated at two hospitals, Northern Italy, February–August 2023 (n = 52)



ER: emergency room; ICU: intensive care unit; Kp: Klebsiella pneumoniae; OSM: Fondazione IRCCS Policlinico San Matteo.

Maximum-likelihood phylogeny of 90 *Kp* strains inferred on coreSNPs with RAxML, including 52 from the outbreak (black labels) and 38 from the BV-BRC database (grey labels – 35 leaves have been collapsed). Bootstrap values above 70/100 are indicated with violet dots. Isolate information is reported alongside the tree, including the hospital of isolation (OSM/Voghera hospital). The OSM wards are also indicated. Presence/absence matrix shows the pattern of plasmids (blue), virulence determinants (yellow) and resistance determinants (red).

has a high level of concordance with that of antimicrobial resistance determinants (Figure 2). In particular, *bla* $_{\text{CTX-M-15}}$, *bla* $_{\text{OXA-1}}$, *catB3* and *arr-3* were associated with plasmid *IncFIB(pQil)*, while *dfrA5* and *sul2* were associated with plasmid sequences *IncHI1B* and *IncFIB(Mar)*. This result suggests a scenario of frequent exchange of gene content. The hypothesis is reinforced by the phylogeny: isolates that have lost plasmids *IncHI1B* and *IncFIB(Mar)* (and the associated resistance genes) are clustered into two high confidence monophyletic subclades (Figure 2).

Discussion

Here we report the isolation of a new *Kp* clone, ST6668, which has spread among a considerable number of patients of two hospitals in the Pavia province within few months. The new clone was isolated from 46 inpatients and six outpatients, suggesting that it is presently circulating also beyond the healthcare setting.

Phylogenetic analysis of the 52 epidemic isolates showed that the cluster of ST6668 genomes presents a few high confidence monophyletic subclades, even though internally not fully resolved. The eight isolates from Voghera hospital did not cluster in one single subclade, suggesting multiple exchange events between hospitals in the Pavia area.

The emergence of the new clone has caused a shift in the epidemiology of carbapenemases at OSM, where KPC had been the prevalent carbapenemase. Notably,

NDM detection has increased since April 2023, becoming the most detected carbapenemase at OSM in August 2023.

The presence in the new clone of a large number of resistance genes, including bla_{NDM-1} , could have facilitated the rapid spread of this clone [11], while no particular virulence factor was observed. Moreover, the high mean age of the patients involved suggests that they may have a history of frequent admissions to hospitals or long-term care facilities. This could have favoured the spread of the ST6668 over the Pavia area. Indeed, three patients were transferred from Voghera to another nearby hospital, belonging to the Pavia-HAU. This scenario underlines the need to establish local, regional and international surveillance and infection prevention strategies [12] in any healthcare setting. In particular, real-time genomic surveillance has proven essential in the timely detection of novel fast-spreading clones and variants of clinical importance [13,14].

Although most of the patients of this study were colonised by ST6668 *Kp* without developing infections, hospitalised patients tend to remain colonised with MDR strains for the duration of their stay, increasing the risk of transmission of MDR strains to other patients [15].

Conclusion

The fast spread of ST6668 and other NDM-producing clones should be controlled, since their reduced susceptibility to most antibiotics allows few therapeutic options. In light of how rapidly the new clone has spread within the healthcare setting in the Pavia area, we underline the importance of prompt communication between microbiologists and clinicians.

Ethical statement

The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (Fasc. 2023-3.11/105). The work described herein is a retrospective study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

Funding statement

This research was supported by EU funding within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. P Eoooooo07, INF-ACT) and by Ricerca Corrente n 08022218 "Real-Time bioinformatic surveillance of nosocomial infections and multidrug resistant bacteria to improve infection control".

Data availability

Sequence data of 52 ST6668 K. pneumoniae are available via BioProject N PRJNA1019273.

Conflict of interest

None declared.

Authors' contributions

Sample collection: IM, AK, MR, CT, MB. Conceptualisation: CM, PC. Formal analysis: IM, GP, SG, MC. Supervision: MC, PC. Writing—original draft: IM, GP, SG, CM, MC. Writing—review and editing: SG, CM, MC, AP, PC, FB. All authors have read and agreed to the published version of the manuscript.

References

1. Tamma PD, Aitken SL, Bonomo RA, Mathers AJ, van Duin D, Clancy CJ. Infectious Diseases Society of America 2023 guidance on the treatment of antimicrobial resistant Gram-negative infections. Clin Infect Dis. 2023;ciad428. https://doi.org/10.1093/cid/ciad428 PMID: 37463564
2. Peirano G, Chen L, Kreiswirth BN, Pitout JDD. Emerging antimicrobial-resistant high-risk Klebsiella pneumoniae clones ST307 and ST147. Antimicrob Agents Chemother. 2020;64(10):e01148-20. https://doi.org/10.1128/AAC.01148-20 PMID: 32747358
3. Di Pilato V, Henrici De Angelis L, Aiezza N, Baccani I, Niccolai C, Parisio EM, et al. Resistome and virulome accretion in an NDM-1-producing ST147 sublineage of Klebsiella pneumoniae associated with an outbreak in Tuscany, Italy: a genotypic and phenotypic characterisation. Lancet Microbe. 2022;3(3):e224-34. https://doi.org/10.1016/S2666-5247(21)00268-8 PMID: 35544076
4. Martin MJ, Corey BW, Sannio F, Hall LR, MacDonald U, Jones BT, et al. Anatomy of an extensively drug-resistant Klebsiella pneumoniae outbreak in Tuscany, Italy. Proc Natl Acad Sci USA. 2021;118(48):e2110227118. https://doi.org/10.1073/pnas.2110227118 PMID: 34819373
5. Merla C, Kuka A, Petazzoni G, Postiglione U, Zatelli M, Gaiarsa S, et al. Livestock-associated methicillin-resistant Staphylococcus aureus in inpatients: a snapshot from an Italian hospital. J Glob Antimicrob Resist. 2022;30:10-5. https://doi.org/10.1016/j.jgar.2022.05.013 PMID: 35644436
6. The European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters. Version 13.0. Växjö: EUCAST; 2023. Available from: https://www.eucast.org
7. Wenzler E, Deraedt MF, Harrington AT, Danizger LH. Synergistic activity of ceftazidime-avibactam and aztreonam against serine and metallo-β-lactamase-producing gram-negative pathogens. Diagn Microbiol Infect Dis. 2017;88(4):352-4. https://doi.org/10.1016/j.diagmicrobio.2017.05.009 PMID: 28602518
8. Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, et al. P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics. Bioinformatics. 2023;39(9):btad571. https://doi.org/10.1093/bioinformatics/btad571 PMID: 37701995
9. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. Nat Commun. 2021;12(1):4188. https://doi.org/10.1038/s41467-021-24448-3 PMID: 34234121
10. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895-903. https://doi.org/10.1128/AAC.02412-14 PMID: 24777092
11. Gullberg E, Albrecht LM, Karlsson C, Sandegren L, Andersson DI. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. MBio. 2014;5(5):e01918-14. https://doi.org/10.1128/mBio.01918-14 PMID: 25293762
12. Ciccolini M, Donker T, Köck R, Mielke M, Hendrix R, Jurke A, et al. Infection prevention in a connected world: the case for a regional approach. Int J Med Microbiol. 2013;303(6-7):380-7. https://doi.org/10.1016/j.ijmm.2013.02.003 PMID: 23499307
13. Forde BM, Bergh H, Cuddihy T, Hajkowicz K, Hurst T, Playford EG, et al. Clinical implementation of routine whole-genome sequencing for hospital infection control of multi-drug resistant pathogens. Clin Infect Dis. 2023;76(3):e1277-84. https://doi.org/10.1093/cid/ciac726 PMID: 36056896
14. Kwong JC, Lane CR, Romanes F, Gonçalves da Silva A, Easton M, Cronin K, et al. Translating genomics into practice for real-time surveillance and response to carbapenemase-producing

Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak. PeerJ. 2018;6:e4210. https://doi.org/10.7717/peerj.4210 PMID: 29312831

15. Vink JP, Otter JA, Edgeworth JD. Carbapenemase-producing Enterobacteriaceae - once positive always positive? Curr Opin Gastroenterol. 2020;36(1):9-16. https://doi.org/10.1097/MOG.0000000000000596 PMID: 31633563

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2023.

Bioinformatics, 2023, **39(9)**, btad571
https://doi.org/10.1093/bioinformatics/btad571
Advance Access Publication Date: 13 September 2023
Applications Note

OXFORD

Phylogenetics

P-DOR, an easy-to-use pipeline to reconstruct bacterial outbreaks using genomics

Gherard Batisti Biffignandi[1], Greta Bellinzona[1], Greta Petazzoni[2,3], Davide Sassera[1,4], Gian Vincenzo Zuccotti[5,6], Claudio Bandi[7], Fausto Baldanti[2,3], Francesco Comandatore [5,*,†], Stefano Gaiarsa [3,*,†]

[1]Department of Biology and Biotechnology, University of Pavia, Pavia, 27100, Italy
[2]Department of Medical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, 27100, Italy
[3]Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, 27100, Italy
[4]Fondazione IRCCS Policlinico San Matteo, Pavia, 27100, Italy
[5]Department of Biomedical and Clinical Sciences, Pediatric Clinical Research Center Romeo ed Enrica Invernizzi, University of Milan, Milan, 20157, Italy
[6]Pediatric Department, Buzzi Children's Hospital, Milan, 20154, Italy
[7]Department of Biosciences, Pediatric Clinical Research Center Romeo ed Enrica Invernizzi, University of Milan, Milan, 20133, Italy

*Department of Biomedical and Clinical Sciences, Pediatric Clinical Research Center Romeo and Enrica Invernizzi, University of Milan, Via Giovanni Battista Grassi 74, 20157, Milan, Italy. E-mail: francesco.comandatore@unimi.it (FC). Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Viale Camillo Golgi 19, 27100, Pavia, Italy. E-mail: s.gaiarsa@smatteo.pv.it (SG).

†These authors contributed equally to this study.

Associate Editor: Russell Schwartz

Abstract

Summary: Bacterial Healthcare-Associated Infections (HAIs) are a major threat worldwide, which can be counteracted by establishing effective infection control measures, guided by constant surveillance and timely epidemiological investigations. Genomics is crucial in modern epidemiology but lacks standard methods and user-friendly software, accessible to users without a strong bioinformatics proficiency. To overcome these issues we developed P-DOR, a novel tool for rapid bacterial outbreak characterization. P-DOR accepts genome assemblies as input, it automatically selects a background of publicly available genomes using k-mer distances and adds it to the analysis dataset before inferring a Single-Nucleotide Polymorphism (SNP)-based phylogeny. Epidemiological clusters are identified considering the phylogenetic tree topology and SNP distances. By analyzing the SNP-distance distribution, the user can gauge the correct threshold. Patient metadata can be inputted as well, to provide a spatio-temporal representation of the outbreak. The entire pipeline is fast and scalable and can be also run on low-end computers.

Availability and implementation: P-DOR is implemented in Python3 and R and can be installed using conda environments. It is available from GitHub https://github.com/SteMIDIfactory/P-DOR under the GPL-3.0 license.

1 Introduction

Bacterial infections are a constant threat to public health worldwide. When dealing with Healthcare-Associated Infections (HAIs) and outbreaks, timely epidemiological investigation is pivotal to establish effective infection control measures (Harris *et al.* 2013, Jiang *et al.* 2015, Raven *et al.* 2017, Balloux *et al.* 2018). Despite their wide use, conventional molecular typing techniques, such as Pulsed Field Gel Electrophoresis (PFGE) and Multi-Locus Sequence Typing (MLST), have a lower discriminatory capability in comparison to the modern Whole Genome Sequencing (WGS)-based typing, while maintaining similar costs and timescales.

Over the past few years, WGS-based typing has been increasingly adopted, first for research purposes and then as a routinary screening tool for infectious disease epidemiology in hospitals and public health settings. This approach leverages *in silico* techniques for isolates typing, antimicrobial profile determination, and outbreak reconstruction (Harris *et al.* 2013, Jiang *et al.* 2015, Onori *et al.* 2015, Raven *et al.* 2017, Balloux *et al.* 2018, Ferrari *et al.* 2019, Sherry *et al.* 2019). Several computational methods to analyze such datasets have been developed, which include database design (Zhou *et al.* 2020, Lam *et al.* 2021), epidemiological models via Bayesian inference (Jombart *et al.* 2014, De Maio *et al.* 2016, Campbell *et al.* 2018), network analysis (Worby *et al.* 2014, 2017), and phylogeny (Didelot *et al.* 2021).

Relationships among strains are mainly inferred using Single-Nucleotide Polymorphisms (SNPs) or k-mers. When reconstructing outbreaks, strains isolated from different sources (e.g. patients, fomites) and having SNP-distances below specific thresholds can be considered part of the same transmission cluster. The network of these genetically correlated strains can be used to reconstruct the pathogen transmission route. Although threshold-based methods are largely applied

Received: 1 June 2023; **Revised:** 24 August 2023; **Editorial Decision:** 10 September 2023; **Accepted:** 12 September 2023
© The Author(s) 2023. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/bioinformatics/article/39/9/btad571/7272610 by guest on 23 January 2026

in genomic epidemiology (Dallman *et al.* 2015, Octavia *et al.* 2015, Hatherell *et al.* 2016, David *et al.* 2019), they lack standardization (Duval *et al.* 2023). Indeed, threshold values can vary across bacterial species/clones because of their different genomic architectures (e.g. mutation rate, recombination). Also the duration of the epidemic event analyzed can influence the genetic variability in the bacterial population: a SNP-distance threshold set to disentangle a short outbreak can be inappropriate for a long-term genomic surveillance study (Duval *et al.* 2023). Furthermore, SNP distances can be affected (even by tenths or hundreds) by the SNP calling approach (e.g. mapping reads or aligning assembled contigs) and by the reference genome and software used. Finally, the sole use of genomic data without the inclusion of other information like clinical metadata (e.g. sample date/type, hospitalization ward) limits the comprehension of epidemic events (Jombart *et al.* 2014, De Maio *et al.* 2016, Stimson *et al.* 2019, Didelot *et al.* 2021, Duval *et al.* 2023).

Most of the software available for WGS-based epidemiological investigation is not user-friendly (De Maio *et al.* 2016, Campbell *et al.* 2018, Zhou *et al.* 2020, Didelot *et al.* 2021), not free (e.g. SeqSphere+ Ridom GmbH software), and/or does not encompass all the analyses required for a comprehensive study (De Maio *et al.* 2016, Zhou *et al.* 2020, Didelot *et al.* 2021). Most of the methods require the user to have a computational background, as they are composed of multiple command-line tasks that must be serially performed in succession, and often require format changes. This prevents most clinicians from performing genomic investigations in first person and limits their understanding of the results. Consequently, it also hampers them from making epidemiological conclusions in light of both clinical information and of their past experience on the field, which in turn would enable them to provide valuable feedback to developers. On the other hand, online tools are available, which are accessible to a wider usership, but lack the tunability that is required for most epidemiological investigations (e.g. Trifinopoulos *et al.* 2016) and are restricted to single tasks (e.g. phylogeny).

To answer the need for a comprehensive, tunable, and user-friendly tool, we developed P-DOR, a bioinformatic pipeline for rapid WGS-based bacterial outbreak detection and characterization. P-DOR integrates genomics and clinical metadata and uses a curated global genomic database to contextualize the strains of interest within the appropriate evolutionary frame. P-DOR is available at https://github.com/SteMIDIfactory/P-DOR.

2 P-DOR workflow

The inputs for the core P-DOR analysis are: (i) a folder containing the query genome assemblies; (ii) a reference genome for SNP extraction; (iii) a sketch database file in the Mash format (Ondov *et al.* 2016); and (iv) a table containing the patient metadata (i.e. hospitalization ward, date of admission and discharge). This last input is not mandatory, but when provided, it will be integrated in the analysis to add further clues on the epidemic event. The query genomes of the study must be in FASTA format and can be complete or draft assemblies.

Sketch files contain the genomic information of the strains from a Source Dataset (SD) chosen by the user. A sketch file is a vastly reduced representation of the genomes, which is produced via the *MinHash* algorithm to allow fast distance estimation using low memory and storage requirements. Regularly updated sketches for each of the ESKAPE members (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) are available in the P-DOR repository. Personalized SD sketches can alternatively be generated by the user using the "makepdordb.py" script. This script can automatically download the high-quality genomes of a species from the BV-BRC collection (Davis *et al.* 2020) or build a custom SD sketch starting from any set of genomes. The sketch files are used to compute the k-mer distances between each query genome and the SD genomes. Then, for each query genome the n most similar SD genomes are selected and joined in a Background Dataset (BD). Lastly, the query genomes are joined with the BD to obtain the Analysis Dataset (AD). Optionally, the entire AD can be scanned for the presence of antimicrobial resistance and virulence genes using AMRFinderPlus (Feldgarden *et al.* 2021).

After that, each genome of the resulting AD is aligned to the reference genome using Nucmer, the alignment obtained is polished using "Delta-Filter" and SNPs are called using "show-snps", all the commands being part of the Mummer4 package (Marçais *et al.* 2018). Lastly, the SNPs of all AD genomes are combined into a coreSNPs alignment using a Python script (Ferrari *et al.* 2019). Then, the coreSNPs alignment is used to infer a Maximum Likelihood (ML) phylogeny (correcting for the SNP ascertainment bias) via the IQ-TREE (Nguyen *et al.* 2015) software, which includes a prior step for the selection of the substitution model.

To infer relationships among the strains, epidemiological clusters are inferred on the basis of coreSNPs distances using a threshold value. The user can manually set the SNP threshold parameter according to previous studies in the literature, or by visualizing the SNP pairwise distances distribution plot provided among the outputs of P-DOR (Fig. 1A). The user should analyze the results, possibly tweak the threshold parameter and run P-DOR again.

3 Output

The main outputs of P-DOR are: (i) a SNP-based phylogenetic tree; (ii) a heatmap reporting the phylogenetic tree and presence/absence of resistance and virulence factors; (iii) a heatmap showing the coreSNP distance matrix; (iv) the histogram of the distribution of the SNP distances; and (v) a graph visualization of the epidemiological clusters, in which each pair of strains (nodes) is connected if the coreSNP distance between them is below the threshold. The epidemiological clusters are also highlighted on the phylogenetic tree. In addition, if patients metadata are provided, P-DOR creates a spatio-temporal representation of the outbreak and outputs a patients timeline plot where strains are placed based on the date of isolation and are connected on the basis of epidemic clusters.

4 Performance test

To test P-DOR, we simulated the genomic sequences of 11 *K.pneumoniae* isolates, involved in a complex epidemic event, with 2 distinct bacterial strains (both belonging to Sequence Type 258) circulating in a hospital in the same period. In detail, we obtained six simulated sequences starting from genome NJST258_1 and 5 sequences from genome NJST258_2

Downloaded from https://academic.oup.com/bioinformatics/article/39/9/btad571/7272610 by guest on 23 January 2026

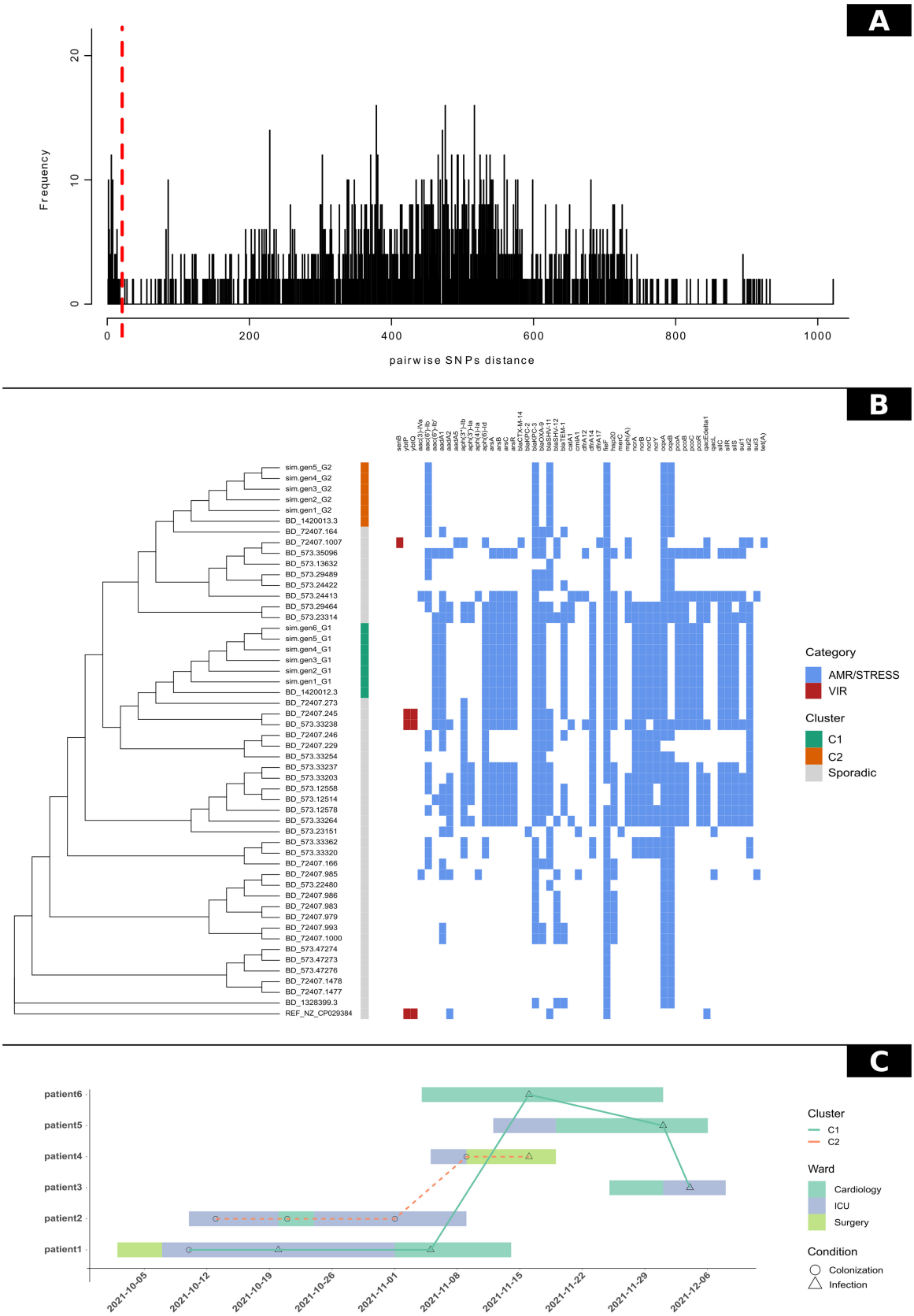


Figure 1. The P-DOR outputs of the test analysis. (A) Phylogenetic tree of the Analysis Dataset (AD). The first column shows epidemiological clusters of strains with SNP distances below the threshold set by the user. In addition, a heatmap representing the detection of resistance and virulence determinants is shown next to the tree for a better representation of the epidemic event. (B) Distribution of SNP distances calculated between all permutations of genome pairs in the AD. (C) Timeline depicting the movements of the patients during the hospitalization. Points indicate outbreak genomes and are shaped according to the isolation source of the corresponding strain. Samples are linked if their genetic distance in terms of SNPs is below the threshold.

Downloaded from https://academic.oup.com/bioinformatics/article/39/9/btad571/7272610 by guest on 23 January 2026

using the software simuG (Yue and Liti 2019). The genomes were simulated in a hierarchical way: i.e. each generated genome differs one to five SNPs from the parent. The simulated dataset was used as query to P-DOR. The SD was obtained from BV-BRC on 16 April 2023 using the script makepdordb.py. The BD was built selecting the 20 closest SD genomes for each query genome, as described above. After analyzing the distribution of SNP distances, the threshold was set at 21 SNPs (Fig. 1A). Finally, the complete genome of strain HS11286 (NZ_CP029384.2) was used as a reference for SNP calling. The phylogeny obtained correctly determined the presence of two outbreaking strains (Fig. 1B). P-DOR divided the simulated genomes into two monophyletic clusters labeled C1 (green) and C4 (orange). Both clusters also include a background genome (BV-BRC codes 1420013.3 and 1420012.3), which correspond to isolates NJST258_2 and NJST258_1, i.e. the genomes used as starting points for the generation of the simulated sequences. These results demonstrate the capability of the P-DOR pipeline to select a background suitable for epidemiological investigations and to identify outbreak clusters. They also show that P-DOR can identify the putative source of each epidemic cluster, when it is available in the SD. Supplementary Fig. S1 shows the SNP distances among all genomes in analysis and further confirms the results observed in the phylogeny (Fig. 1B). Furthermore, the timeline (Fig. 1C) can be used to hypothesize the chain of transmission based on the dates of isolation and the epidemiologic classification of the isolates.

The analysis was performed on a total of 50 genomes using a maximum of 313 Mb of RAM. The entire process required 1 h and 37 min, using four threads at 2.3 Ghz, 29 min when using 20 threads; these numbers drop respectively to 2.9 and 2.1 min when excluding the time-consuming AMRFinderPlus step (default setting).

We separately tested the performance and accuracy of our SNP-calling approach, which is based on the Mummer4 package (Marçais *et al.* 2018). We called point mutations using our method, Parsnp (Treangen *et al.* 2014), and Snippy (github.com/tseemann/snippy); we then inferred the phylogeny of the simulated dataset using all three alignments (IQ-TREE; Nguyen *et al.* 2015). Using two threads, the P-DOR built-in method yielded 8318 SNPs in 1.7 min, Parsnp yielded 10202 SNPs in 3.1 min, while Snippy took 31.2 min to output 9266 SNPs. We calculated the pairwise cophenetic correlation value among the three phylogenetic trees using the R library dendextend (Galili 2015). Phylogenies were highly concordant (Parsnp versus P-DOR: 0.996; Snippy versus P-DOR: 0.996; Snippy versus Parsnp: 0.999) and the monophyly of outbreak genomes was maintained among the three approaches.

To test the pipeline performance on a large-scale dataset, the P-DOR analysis on the simulated dataset was repeated using 1000 nearest genomes for each genome in the input folder. The AD was composed of 1388 total genomes and the analysis took 7 h and 50 min to complete, using a peak of 4.49 Gb of RAM on 20 threads (avoiding the AMRFinderPlus step; background genome download time excluded). These results show that P-DOR is a fast tool, which can be used in clinical contexts even when high informatic skills or resources are not available. Future efforts will be focused on developing a web-based interface, to further improve the ease of use and removing the need for computational resources.

Supplementary data

Supplementary data are available at *Bioinformatics* online.

Conflict of interest

None declared.

Funding

This work was supported by Ricerca Corrente [924-rcr2018i-34] (Italian Ministry of Health).

Data availability

All data are incorporated into the article and its online supplementary material.

References

Balloux F, Brønstad Brynildsrud O, van Dorp L *et al.* From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol* 2018;**26**:1035–48.

Campbell F, Didelot X, Fitzjohn R *et al.* outbreaker2: a modular platform for outbreak reconstruction. *BMC Bioinformatics* 2018;**19**:363.

Dallman TJ, Ashton PM, Byrne L *et al.* Applying phylogenomics to understand the emergence of Shiga-toxin-producing O157:H7 strains causing severe human disease in the UK. *Microb Genom* 2015;**1**:e000029.

David S, Reuter S, Harris SR *et al.*; ESGEM Study Group. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat Microbiol* 2019;**4**:1919–29.

Davis JJ, Wattam AW, Aziz RK *et al.* The PATRIC bioinformatics resource center: expanding data and analysis capabilities. *Nucleic Acids Res* 2020;**48**:D606–12.

De Maio N, Wu C-H, Wilson DJ *et al.* SCOTTI: efficient reconstruction of transmission within outbreaks with the structured coalescent. *PLoS Comput Biol* 2016;**12**:e1005130.

Didelot X, Kendall M, Xu Y *et al.* Genomic epidemiology analysis of infectious disease outbreaks using TransPhylo. *Curr Protoc* 2021;**1**:e60.

Duval A, Opatowski L, Brisse S *et al.* Defining genomic epidemiology thresholds for common-source bacterial outbreaks: a modelling study. *Lancet Microbe* 2023;**4**:e349–57.

Feldgarden M, Brover V, Gonzalez-Escalona N *et al.* AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci Rep* 2021;**11**:12728.

Ferrari C, Corbella M, Gaiarsa S *et al.* Multiple KPC clones contribute to an extended hospital outbreak. *Front Microbiol* 2019;**10**:2767.

Galili T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* 2015;**31**:3718–20.

Harris SR, Cartwright EJP, Török ME *et al.* Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis* 2013;**13**:130–6.

Hatherell H-A, Colijn C, Stagg HR *et al.* Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. *BMC Med* 2016;**14**:21.

Jiang Y, Wei Z, Wang Y *et al.* Tracking a hospital outbreak of KPC-producing ST11 *Klebsiella pneumoniae* with whole genome sequencing. *Clin Microbiol Infect* 2015;**21**:1001–7.

Jombart T, Cori A, Didelot X *et al.* Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLoS Comput Biol* 2014;**10**:e1003457.

Downloaded from https://academic.oup.com/bioinformatics/article/39/9/btad571/7272610 by guest on 23 January 2026

Lam MMC, Wick RR, Watts SC *et al.* A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat Commun* 2021;**12**:4188.

Marçais G, Delcher AL, Phillippy AM *et al.* MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol* 2018;**14**: e1005944.

Nguyen L-T, Schmidt HA, von Haeseler A *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Octavia S, Wang Q, Tanaka MM *et al.* Delineating community outbreaks of *Salmonella enterica* serovar typhimurium by use of whole-genome sequencing: insights into genomic variability within an outbreak. *J Clin Microbiol* 2015;**53**:1063–71.

Ondov BD, Treangen TJ, Melsted P *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016;**17**:132.

Onori R, Gaiarsa S, Comandatore F *et al.* Tracking nosocomial *Klebsiella pneumoniae* infections and outbreaks by whole-genome analysis: small-scale Italian scenario within a single hospital. *J Clin Microbiol* 2015;**53**:2861–8.

Raven KE, Gouliouris T, Brodrick H *et al.* Complex routes of nosocomial vancomycin-resistant *Enterococcus faecium* transmission revealed by genome sequencing. *Clin Infect Dis* 2017;**64**:886–93.

Sherry NL, Lane CR, Kwong JC *et al.* Genomics for molecular epidemiology and detecting transmission of Carbapenemase-Producing in Victoria, Australia, 2012 to 2016. *J Clin Microbiol* 2019;**57**: e00573–19.

Stimson J, Gardy J, Mathema B *et al.* Beyond the SNP threshold: identifying outbreak clusters using inferred transmissions. *Mol Biol Evol* 2019;**36**:587–603.

Treangen TJ, Ondov BD, Koren S *et al.* The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;**15**:524.

Trifinopoulos J, Nguyen L-T, von Haeseler A *et al.* W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* 2016;**44**:W232–5.

Worby CJ, Lipsitch M, Hanage WP *et al.* Shared genomic variants: identification of transmission routes using pathogen deep-sequence data. *Am J Epidemiol* 2017;**186**:1209–16.

Worby CJ, Lipsitch M, Hanage WP *et al.* Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. *PLoS Comput Biol* 2014;**10**: e1003549.

Yue J-X, Liti G. simuG: a general-purpose genome simulator. *Bioinformatics* 2019;**35**:4442–4.

Zhou Z, Alikhan N-F, Mohamed K *et al.*; Agama Study Group. The EnteroBase user's guide, with case studies on transmissions, phylogeny, and core genomic diversity. *Genome Res* 2020;**30**: 138–52.

Downloaded from https://academic.oup.com/bioinformatics/article/39/9/btad571/7272610 by guest on 23 January 2026

microorganisms

MDPI

Article

Gut Microbiota and B Cell Receptor (BCR) Inhibitors for the Treatment of Chronic Lymphocytic Leukemia: Is Biodiversity Correlated with Clinical Response or Immune-Related Adverse Event Occurrence? A Cross-Sectional Study

Valentina Zuccaro [1,*], Greta Petazzoni [2,3], Irene Mileto [2], Marta Corbella [2], Erika Asperges [1], Paolo Sacchi [1], Sara Rattotti [4], Marzia Varettoni [4], Irene Defrancesco [3,4], Patrizia Cambieri [2], Fausto Baldanti [2,3], Luca Arcaini [4,5] and Raffaele Bruno [1,3]

[1] Division of Infectious Diseases I, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
[2] Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
[3] Department of Medical, Surgical, Diagnostic and Pediatric Science, University of Pavia, 27100 Pavia, Italy
[4] Division of Hematology, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy
[5] Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy
* Correspondence: v.zuccaro@smatteo.pv.it

Abstract: Several studies have strengthened the link between the gut microbiota (GM) and the response to immunotherapy in patients with tumors, highlighting the potential role of GM as a biomarker of response. Targeted therapies including B-cell receptor (BCR) inhibitors (BCRi) represent the newest approach to the treatment of chronic lymphocytic leukemia (CLL); however, not all patients achieve a satisfactory response, and immune-related adverse events (irAEs) can also impact the efficacy. The aim of the study was to compare GM biodiversity in patients with CLL, treated with BCRi for at least 12 months. Twelve patients were enrolled: 10 patients in the *responder* group (R) and 2 patients in the *non-responder* group (NR). We identified seven patients (58.3%) who experienced adverse reactions (AE). Although we did not observe a significant difference across the study population in terms of relative abundance and alpha and beta diversity, we found a differing distribution of bacterial taxa between the analyzed groups. We noted a higher level of the class Bacteroidia and the order Bacteroidales in the R group, and an inversion in the Firmicutes and Bacteroidetes ratio in the AE group. No prior studies have focused on linking GM and response to BCRi in these patients. Although the analyses are preliminary, they provide suggestions to guide future research.

Keywords: gut microbiota; B cell receptor inhibitors; chronic lymphocytic leukemia; clinical response

check for updates

Citation: Zuccaro, V.; Petazzoni, G.; Mileto, I.; Corbella, M.; Asperges, E.; Sacchi, P.; Rattotti, S.; Varettoni, M.; Defrancesco, I.; Cambieri, P.; et al. Gut Microbiota and B Cell Receptor (BCR) Inhibitors for the Treatment of Chronic Lymphocytic Leukemia: Is Biodiversity Correlated with Clinical Response or Immune-Related Adverse Event Occurrence? A Cross-Sectional Study. *Microorganisms* **2023**, *11*, 1305. https://doi.org/10.3390/microorganisms11051305

Academic Editor: Markus M. Heimesaat

Received: 27 March 2023
Revised: 12 May 2023
Accepted: 15 May 2023
Published: 17 May 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Chronic lymphoproliferative disorders (LPD) are a large group of diseases characterized by the monoclonal expansion of lymphocytes leading to monoclonal lymphocytosis, lymphadenopathy and bone marrow infiltration [1]. Among LPD, chronic lymphocytic leukemia (CLL) is the most common leukemia observed in adults. According to the 2018 International Workshop on Chronic Lymphocytic Leukemia (iwCLL) updated consensus [2], at the time of CLL diagnosis, treatment is indicated only for active disease (i.e., peripheral cytopenias due to bone marrow failure, massive or symptomatic lymphadenopathy or splenomegaly, progressive lymphocytosis or disease-related symptoms). Many different molecules can be used, and more recently, targeted therapy for CLL has dramatically reshaped the landscape of treatment. For example, Bruton tyrosine kinase (BTK) inhibitor-based therapies have demonstrated that these compounds are highly effective in disease control [3]. Ibrutinib, the first-in-class inhibitor of Bruton tyrosine kinase (Btk-inhibitor), and Idelalisib, the first-in-class inhibitor of phosphatidylinositol 3-kinase δ (PI3K-inhibitor), have been approved for the treatment of indolent B-cell malignancies. At the time of our

study enrollment, these drugs were used especially used for previously unmet needs (i.e., patients with relapsed or refractory disease, high-risk cytogenetic or molecular abnormalities, or with comorbidities) [4]. Despite the high efficacy of BCRi, a portion of patients still do not respond to therapy or experience immune-related adverse events (irAEs) [2]. Since most irAEs of these new molecules are of gastrointestinal origin (e.g., colitis and diarrhea), researchers have suggested a potential role of the gut microbiota (GM) in modulating the response to these agents and in the development of such adverse reactions. Robust evidence regarding the cross talk of GM and host immune responses in melanoma patients has also accumulated, suggesting the potential involvement of the GM in modulating the efficacy of such anti-CTLA4- and anti-PD1-based therapies. The findings of a preclinical study by Vetizou et al. showed an association between the relative abundance of *Bacteroides fragilis* and *Burkholderia cepacia* and the efficacy of anti-CTLA4 antibodies in reducing sarcoma tumor growth in mice [5]. Following this, metagenomic analysis of melanoma patients' fecal gut microbiomes showed a higher relative abundance of *Clostridiales*, in particular those belonging to the *Ruminococcaceae* and *Faecalibacterium* families in anti-PD1 responders when compared to non-responders to therapy [6]. Even in hematological disorders, several studies have strengthened the bidirectional link between GM and the response to therapy: GM composition can influence drug efficacy and, conversely, chemotherapeutics can alter GM [7].

Although a lack of consensus exists, convincing evidence has increased, such as the findings of Hakim et al. regarding a correlation between GM composition and the development of chemotherapy-related adverse effects in patients affected by acute lymphocytic leukemia. For instance, the relative abundance of *Enterococcaceae* or *Streptococcaceae* is associated with infections, and that of *Proteobacteria* is associated with febrile neutropenia [8]. With this in mind, it is possible to consider GM composition as a biomarker of response and of adverse events [9] and, at the same time, to imagine the manipulation of gut microbiota composition as a new strategy to enhance the response to treatment. Concerning this, Baruch et al. noted in a phase I study (NCT03353402) a longer progression-free survival (PFS) in refractory metastatic melanoma patients who underwent fecal microbiota transplantation from a complete responder patient [10,11]. So far, no study has investigated the alteration of GM in patients diagnosed with CLL treated with BCRi and the potential association with therapy outcome and irAEs. We hypothesized that the immune-stimulatory and antitumor effects of BCRi (Btk- and PI3K-inhibitors, such as Ibrutinib and Idelalisib) in patients with CLL could be influenced by distinct gut microbiota compositions. Hence, the aim of this study was to investigate whether GM characteristics were correlated with clinical responses and the occurrence of irAEs in patients affected by CLL who underwent treatment with BCRi.

2. Materials and Methods

2.1. Study Setting

We designed our research as a cross-sectional study. All patients from the Hematology Unit of Fondazione "IRCCS Policlinico San Matteo di Pavia" affected by CLL who started BCRi therapy with Btk- or PI3K-inhibitors at least 12 months before the date of study enrollment were enrolled. We collected fecal samples in those patients at their 12-month visit after initiation of BCRi treatment. For patients whose BCRi was discontinued for any reason before the 12-month visit, a visit nearest the 12th month ($\pm$1 month) was scheduled in order to collect data. Fecal samples were examined with next-generation sequencing (NGS) techniques. All participating patients signed an informed consent form. We split the study population into a responder group (R) and a not-responder group (NR) according to the response definition of the 2018 iwCLL updated consensus as follows:

R Group: Patients meeting the clinical response criteria (no evidence of clinical progression criteria) and undergoing ongoing therapy with BCRi.
NR Group: Patients with clinical progression criteria and BCRi therapy discontinuation within 12 months of treatment initiation.

We further collected demographic data such as age, sex, clinical history, intake of antibiotic therapies during the last 3 months, hematological history (date of diagnosis of CLL, cytogenetic and molecular features, treatment lines, clinical and non-clinical response criteria, adverse events), patients' outcome, and cause of death. Special diets, such as celiac, lactose or casein-free, were also recorded. Response assessment was evaluated at least 2 months after the beginning of treatment and thereafter, as is clinical practice.

We further categorized the study population into two other groups: the AE group, patients who experienced irAEs at any time during the treatment, and the NAE group, patients who did not experience irAEs. The study was conducted according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for reporting observational studies [12], and was approved by the local Ethics Committee (Comitato Etico Area Pavia) and Institutional Review Board (P-20180107349, 7 December 2018).

2.2. Objectives

The primary objective of this study was to examine the GM composition in a population of patients with CLL in order to recognize differences between responders (R) and non-responders (NR). The secondary objective was to assess GM biodiversity in patients with/without irAEs.

2.3. Biological Samples and DNA Sequencing

At the first visit falling into the study period (at least 12 months from BRCi initation), patients were instructed to provide a stool sample at the following visit. All stool samples were collected and stored at −80 °C in the Laboratory of Microbiology and Virology Unit of Fondazione IRCCS Policlinico San Matteo (Pavia, Italy). Then, we extracted the genomic DNA using the DNeasy® PowerSoil Pro Kit (Qiagen, Hilden, DE, USA) according to the manufacturer's instructions and the V3-V6 hypervariable regions of the 16S rRNA gene were amplified using Microbiota Solution B Kit (Arrow Diagnostics s.r.l., Genova, Italy). Extracted DNA was finally sequenced using a paired-end 2 × 250 bp cycle on the Illumina MiSeq sequencing system (Illumina, CA, USA).

2.4. Statistical and Bioinformatic Analyses

Raw reads were processed with the MicrobAT (Microbiota Analysis Tool) system of the SmartSeq S.r.l. (Novara, Italy). In particular, the software filtered row reads for length and quality (data quality evaluation), then it aligned them against the RDP database (vr. 11.4) and it assigned them to a species if they met specific criteria (query coverage ≥80% and similarity ≥97%). Thus, the software produced the phylum, class, order, family, genus, and species level for each bacterium found in the samples.

Results were then processed using a dedicated pipeline in R software. Firstly, data were processed in order to identify and remove features that were unlikely to be useful when modeling the data. Features were filtered for low count, considering features with at least 4 counts in 20% of the samples (20% of prevalence cut-off), and low variance (based on the interquartile range). Finally, filtered features were used to compute the relative abundances of microbial taxa in each sample. Microbial profiles of taxa with a prevalence of at least >1% (most abundant) in one sample of the dataset were compared between groups: R vs. NR group and AE vs. NAE group, using the Mann–Whitney U-test. The significance threshold (p-value) was set at 0.05.

The within-sample diversity (α-diversity) was estimated at each taxonomic level using the phyloseq package in R [13]. The observed richness and Shannon indexes were compared between groups using the Mann–Whitney U-test (p-value threshold set at 0.05).

The diversity in composition among samples (β-diversity) was evaluated at all taxonomic levels, computing the dissimilarity matrices with the vegan package in R [14]. Bray–Curtis distances were applied and then visualized as principal coordinate analysis (PCoA). To statistically assess the differences between groups, permutational multivariate

analysis of variance (PERMANOVA) was performed with the threshold for *p*-value set at 0.05.

3. Results

Thirteen patients were enrolled from January 2019 to October 2019 in the Hematology Unit of Fondazione "IRCCS Policlinico San Matteo di Pavia". One was further excluded. The average age of patients was 63.6 years old, and 50% (6/12) were male. No patients reported a special diet. The hematologic diagnosis was CLL in all patients. Seven patients (58.3%) were treated with Ibrutinib and five (41.7%) were treated with Idelalisib. Only one patient was treated with BCRi as the first line of treatment. For all other patients, BCRi represented a later course of therapy.

According to the response definition of the 2018 iwCLL updated consensus [2], we split the study population into the *responder* group (R) and *non-responder* group (NR). We identified 10 patients who met the clinical response criteria (R) and 2 patients who met the clinical progression criteria (NR). Moreover, we identified seven patients (58.3%) who experienced adverse reactions (AE group), mainly diarrhea (three out of seven patients), followed by fever and liver toxicity. Of those patients, five (71.7%) were taking Idelalisib. Interestingly, the two NR patients were taking Ibrutinib. Although the two NR patients experienced irAEs, these were not the leading cause of discontinuation of the treatment. The demographic and hematological features of the study population are reported in Table 1.

Table 1. Patients' characteristics. B-cell receptor inhibitors used to treat patients are shown in the BCRi. Mutated (mut.) genes, unmutated (unmut.) genes, deletions of chromosomes and the trisomy recorded are listed in genomic features. irAEs are the immune-related adverse events experienced by each patient. Non-responder (NR) and responder (R) patients are shown in the outcome.

ID Patient	Age (YO)	Sex	BCRi	Genomic Features	Treatment Lines	Antibiotic Intake	irAEs	Outcome
01	40–50	M	Ibrutinib	TP53 mut., IGVH unmut.	3°	No	Yes	NR
03	60–70	M	Idelalisib	TP53 mut., IGVH unmut.	2°	Yes	Yes	R
04	60–70	F	Ibrutinib	TP53 mut., IGVH unmut., del(13q), del(17p)	2°	Yes	No	R
05	60–70	F	Ibrutinib	TP53 mut., IGVH unmut.	2°		No	R
06	60–70	M	Idelalisib	TP53 mut., IGVH unmut., del(17p), Trisomy 13	2°	Yes	Yes	R
07	60–70	F	Idelalisib	TP53 mut., IGVH unmut., del(13q)	4°	Yes	Yes	R
08	60–70	F	Idelalisib	TP53 mut., IGVH unmut., del(13q)	2°	Yes	Yes	R
09	60–70	M	Ibrutinib	TP53 mut., IGVH unmut., del(17p)	5°	Yes	Yes	NR
10	60–70	M	Ibrutinib	IGVH unmut., del(13q)	4°	Yes	No	R
11	70–80	F	Ibrutinib	TP53 mut., IGVH unmut., del(17p)	1°	No	No	R
12	70–80	F	Idelalisib	IGVH unmut., del(13q)	2°	No	Yes	R
13	60–70	M	Ibrutinib	IGVH unmut., Trisomy 13	3°	Yes	No	R

3.1. Taxonomic Composition and Diversity of Fecal Bacterial Communities in the Responder and Non-Responder Group

We examined the gut microbial signature across the *responder* and *non-responder* groups. The average relative abundance of bacterial phyla, classes and orders between the

two groups did not statistically differ, as shown in Figure 1. Concerning the phylum, the frequency of *Firmicutes* was 38.3% in R and 47.3% in NR, while that of *Bacteroidetes* was 35.7% and 28.2%, respectively. Although there was no significant difference, at the class level we found lower levels of fecal *Clostridia* (32.3%) and higher levels of fecal *Bacteroidia* (34.9%) in the R group, while the frequencies in the NR group were the opposite (40.9% and 27.9%, respectively) (Figure 1b). Regarding the order, despite the lack of statistical significance, we found a different average relative abundance between the two groups: *Bacteroidales* were the most abundant in the R group (34.9%) as *Clostridiales* were the most abundant in the NR group (40.1%) (Figure 1c).

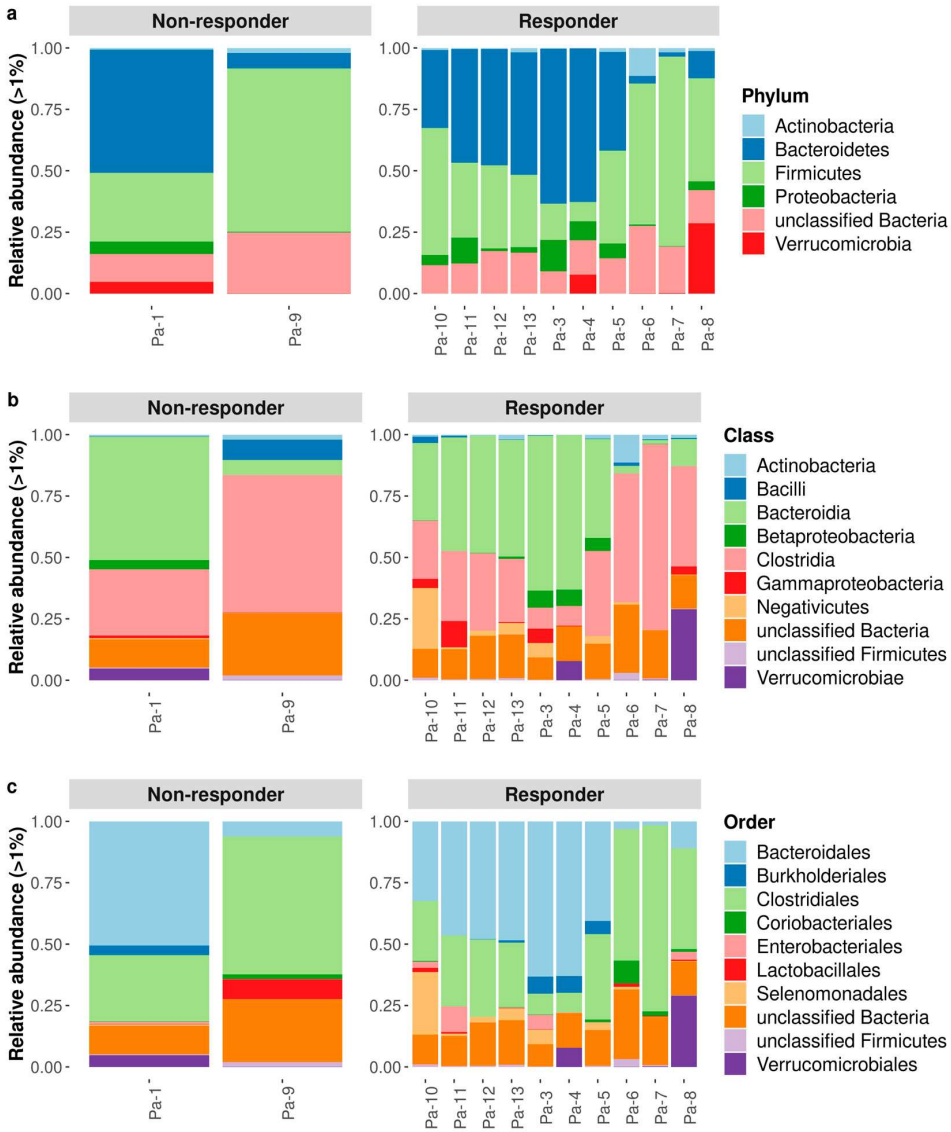


Figure 1. Taxonomic composition and diversity of fecal bacterial communities for each patient stratified for responder and non-responder group. The most represented phyla (**a**), classes (**b**), and orders (**c**) identified in the study groups are shown with relative abundance. Only taxa whose relative abundance was >1% in at least one group were included.

As for the species, the abundance of *Bacteroides* sp. S-18 was higher in the NR group compared with the R group ($p = 0.034$, Mann–Whitney U-test).

Only Ibrutinib-treated patients were selected for GM analysis to provide a homogeneous sample. Notably, the NR group only received Ibrutinib, while the responder group within the Ibrutinib-treated cohort (RI) differed. The taxonomic composition of the RI cohort

did not differ compared to the first one. At the phylum level, the *Firmicutes* frequency was 32.2% in RI, while that of *Bacteroidetes* was 46.3%. Similarly to the R group, at the class level, lower levels of fecal *Clostridia* (32.3%) and higher levels of fecal *Bacteroidia* (34.9%) were confirmed in the RI group when compared to the NR group (Supplementary Figure S1). Concerning the order, the trend was confirmed, and we found a greater average relative abundance for the RI group compared with the NR group: *Bacteroidales* were the most abundant in the RI group (44.9%). No significant difference was found between RI and NR patients in terms of α-diversity and β-diversity (Supplementary Figures S2 and S3).

α-Diversity and β-Diversity Statistical Analyses in the Responder and Non-Responder Groups

The α-diversity indexes (observed richness and Shannon) were computed at all taxonomic levels to analyze the within-sample diversity, and there were no significant differences among the groups (Figure 2).

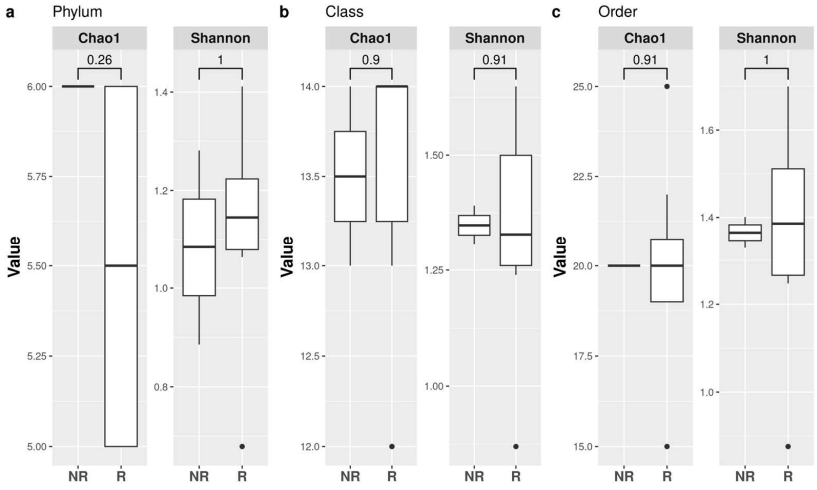


Figure 2. Alpha-diversity. Observed richness and Shannon indices are presented at the taxonomic level of phylum (**a**), class (**b**), and order (**c**). Significant ($p < 0.05$) comparisons between responder (R) and non-responder (NR) patients are indicated in the boxplot.

The β-diversity (diversity in composition among samples) was computed at all taxonomic levels, and we did not observe any significant differences, suggesting the lack of a different gut microbial community between the analyzed groups (Figure 3).

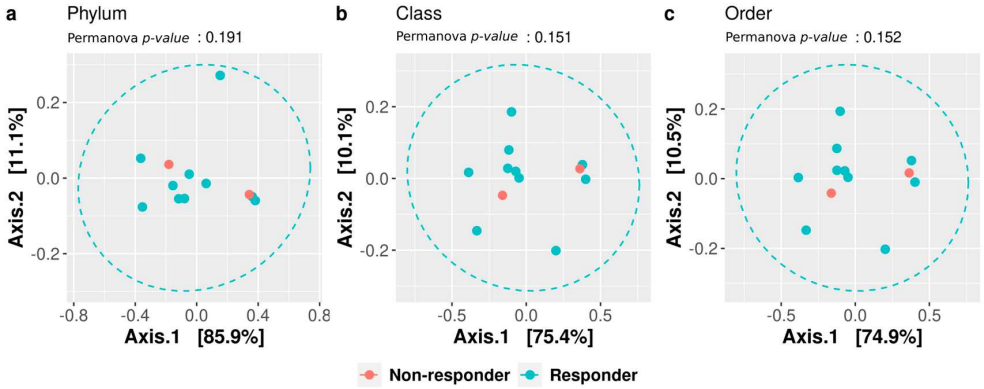


Figure 3. Beta-diversity. The microbiota distances were evaluated through the Bray-Curtis dissimilarity matrix at the taxonomic level of phylum (**a**), class (**b**), and order (**c**) and visualized through principal coordinate analysis (PCoA). Each point represents the microbiota composition of one sample stratified for outcome (responder and non-responder).

3.2. Taxonomic Composition and Diversity of Fecal Bacterial Communities between Patients with and without irAEs

The study population was later categorized according to the occurrence of irAEs: the AE group for patients who experienced irAEs and the non-adverse events (NAE) group for patients who did not experience such adverse events. Despite the lack of statistical significance, at the taxonomic rank of phylum, we observed a different average relative abundance across the AE group and NAE group. Specifically, we found a higher number of fecal *Firmicutes* (45.7%) and a lower number of fecal *Bacteroidetes* (26.1%) in the AE group, while the frequencies in the NAE group were the opposite (31.5% and 46.1%, respectively) (Figure 4a). Similarly, at the class level, the abundance of fecal communities differed between the two groups without statistical significance: *Clostridia* was the most abundant in the AE group (41%), while *Bacteroidia* was the most represented in NAE group (44.7%) (Figure 4b). At the order level, we found different average relative abundances between the two groups: *Bacteroidales* were the most abundant in the NAE group (44.7%), while *Clostridiales* were higher in the AE group (40.9%) (Figure 4c).

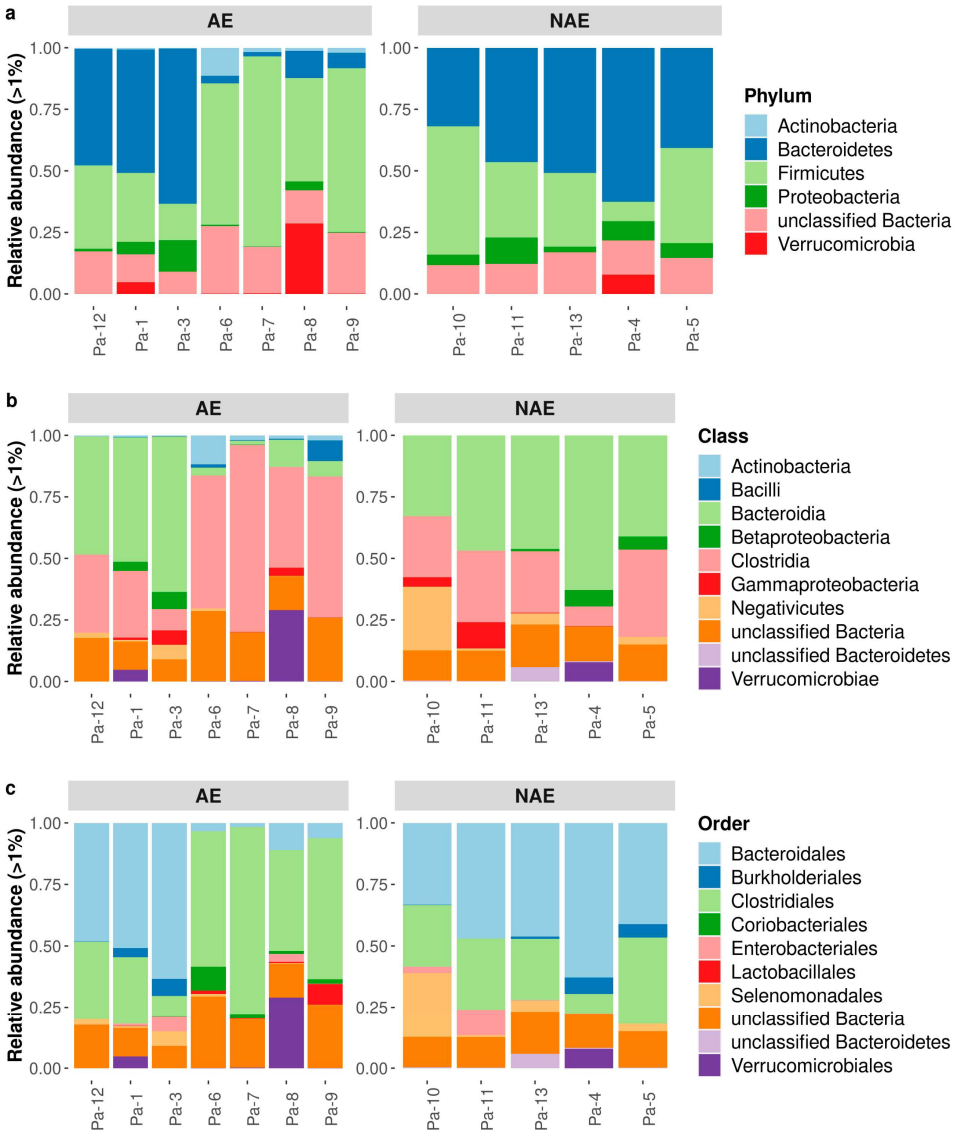


Figure 4. Taxonomic composition and diversity of fecal bacterial communities for each patient who experienced an irAE (AE) or did not (NAE). The most represented phyla (**a**), classes (**b**), and orders (**c**) identified in the study groups are shown with relative abundance. Only taxa whose relative abundance was >1% in at least one group were included.

Interestingly, some bacterial taxa were only found in one of the two groups, as shown in Table 2.

Table 2. Bacterial taxa found only in the AE or NAE group.

	AE Group	NAE Group
Phylum	Actinobacteria	-
Class	Actinobacteria Bacilli	unclassified Bacteroidetes
Order	*Coriobacteriales* *Lactobacillales*	unclassified Bacteroidetes
Family	*Coriobacteriaceae*	*Veillonellaceae* *Prevotellaceae*
Species	*Akkermansia muciniphila* ATCC BAA-835 *Blautia luti* *Ruminococcus obeum* *Barnesiella spp* EBA4-14	*Bacteroides eggerthii* *Veillonella spp* ICM51a *Faecalibacterium prausnitzii* *Bacteroides ovatus* *Butyrivibrio fibrisolvens* *Akkermansia muciniphila* *Bacteroidaceae bacterium* Smarlab 3301643 *Parabacteroides distasonis* ATCC 8503

α-Diversity and β-Diversity Statistical Analyses in Patients with or without irAEs

α-diversity indexes were also calculated according to irAE categorization, and we did not observe any significant differences between the two groups for all of the taxonomic ranks (Figure 5).

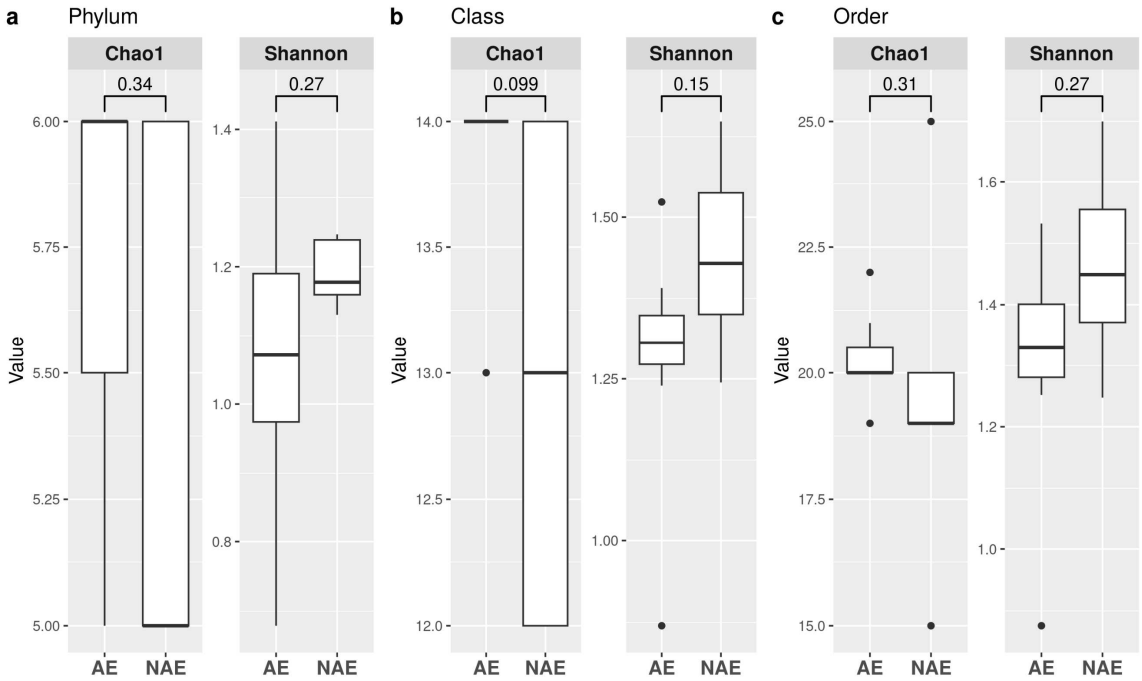


Figure 5. α-diversity according to the experience of an irAE (AE) or not (NAE). Observed richness and Shannon indices are presented at the taxonomic level of phylum (**a**), class (**b**), order (**c**). Significant ($p < 0.05$) comparisons between AE and NAE are indicated in the boxplot.

Similarly, we performed the between-sample diversity (β-diversity) evaluation, and the data suggested no significant differences in all bacterial consortia when comparing the samples according to irAE categorization (Figure 6).

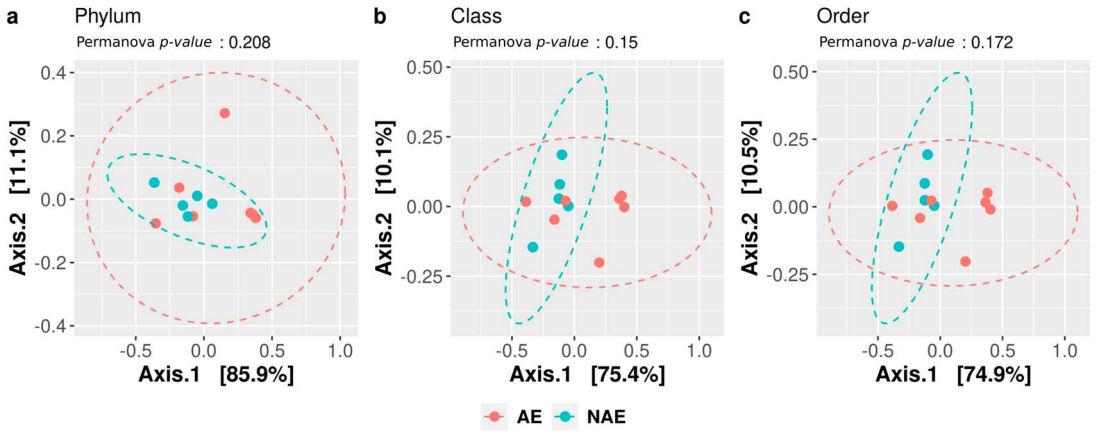


Figure 6. Beta-diversity according to the experience of an irAE (AE) or not (NAE). The microbiota distances were evaluated through the Bray-Curtis dissimilarity matrix at the taxonomic level of phylum (**a**), class (**b**), and order (**c**) and visualized through principal coordinate analysis (PCoA). Each point represents the microbiota composition of one sample.

4. Discussion

In this cross-sectional exploratory investigation, we reported an overall analysis of the gut microbiota of patients with CLL receiving BCRi treatment, and we speculated the that GM signature was associated with clinical response and the occurrence of irAEs. Although we did not observe a significant difference across the study population in terms of relative abundance and alpha and beta diversity, we found a differing distribution of bacterial taxa between the analyzed groups. In particular, we noted a higher number of bacteria from the class *Bacteroidia* and the order *Bacteroidales* in the R group when compared to NR, who are conversely enriched in the class Clostridia and the order *Clostridiales.* Similarly, in the AE group, we observed a higher number of bacteria from the class *Clostridia* and the order *Clostridiales*. These findings were consistent with data reported by Karpinets et al. about a population affected by cervical cancers, showing that a fecal community rich in *Clostridia* was characterized by pathways involved in stress response [15]. Not surprisingly, the emerging view based on numerous studies is that collective interactions within the bacterial community drive disease through the dysregulation of mucosal immunity and the disruption of the mucosal barrier. In the AE group, we also found an inversion in the *Firmicutes* and *Bacteroidetes* ratio. As is widely known, GM produces bio-products, and short chain fatty acids (SCFAs), such as acetate, propionate and butyrate, are the most relevant of these, such as in inflammation pathways. Butyrate, mainly produced by the phylum *Firmicutes*, provides for the maintenance of the integrity of colonocytes and plays a key role in the control of inflammation [16]. With this in mind, the lower abundance of *Firmicutes*, which means lower butyrate production, is not an unexpected finding in the AE group. Among *Firmicutes*, the butyrate-producing *Faecalibacterium prausnitzii* had been widely proved to be associated with a good response to immune checkpoint inhibitors (ICIs) and a higher risk of developing irAEs (in particular, immune-related colitis) in metastatic melanoma patients [17]. On the other hand, among *Bacteroidetes*, *Akkermansia muciniphila* seemingly plays a protective role in the occurrence of colitis [16]. In agreement with literature data, in our cohort, we did not find *Akkermansia muciniphila* in the AE group and we failed to demonstrate the association between *Faecalibacterium prausnitzii* and irAEs; conversely, we found this species only in the NAE group. An explanation of these contradictory data could be the nonhomogeneity of irAEs other than colitis that we observed. Moreover, we speculated that the metabolites secreted by *F. prausnitzii*, blocking NF-κB activation and IL-8 secretion [18], might play a role in attenuating the manifestation of irAEs, as demonstrated by Liu et al. in patients treated with anti-PD-1 inhibitors: a higher abundance of *Faecalibacterium* was associated with mild irAEs [19]. Several confounding factors might further impact the interpretability of our findings,

especially the study design and the small sample size. As discussed above, we found different signatures across the study population, and we tried to understand which factors could drive these differences and how. Firstly, most of the study population had already received chemotherapy and many of them additionally received antibiotic prophylaxis. In this regard, the main challenge is to assess the impact of antibiotics in modulating GM, and then the antitumor effect. We already mentioned that chemotherapeutics can alter GM; for example, in preclinical studies, cyclophosphamide altered GM composition through the selective translocation of species of Gram-positive bacteria into secondary lymphoid organs, leading to tumor invasion of pathogenic TH17 cells [7,20]. With this in mind, in a large German cohort of patients affected by CLL and treated with cyclophosphamide, according to Viad's findings, Pflug et al. showed that treatment with anti-Gram-positive reduced anticancer efficacy [21].

As discussed above, our study has several limitations that might impact the interpretation of our findings. Furthermore, the small sample size prevented us from reaching statistical significance. Further validation of our findings in prospective cohorts or analyses conducted in association with randomized controlled trials should be promoted. Despite these limitations, the strength of our study comes from a strict definition of patient characteristics, and, to the best of our knowledge, no previous research on this topic is available.

5. Conclusions

In summary, despite the lack of statistical significance, we found a different distribution of bacterial taxa between the analyzed groups, and our findings are consistent with the currently well-known evidence regarding the cross-talk of GM and host immune responses in melanoma patients. In the era of precision medicine, further studies in this group of patients are needed in order to better define the prognostic role of the GM signature and to develop targeted interventions such as the manipulation of GM through fecal microbial transplantation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/microorganisms11051305/s1, Figure S1: Taxonomic composition and diversity of Fecal Bacterial Communities for each patient treated with Ibrutinib stratified for Responder and Non-responder group; Figure S2: α-diversity. Figure S3: Beta-diversity

Author Contributions: Conceptualization, V.Z. and I.D.; methodology, G.P., I.M. and M.C.; software, G.P.; validation, V.Z.; formal analysis, M.C., G.P. and I.M.; investigation, V.Z., E.A., S.R., M.V. and I.D.; resources, V.Z., E.A., S.R., M.V. and I.D.; data curation V.Z., S.R., M.V. and I.D.; writing—original draft preparation V.Z., E.A., P.S., M.C. and G.P.; writing—review and editing V.Z., E.A., P.S., M.C. and G.P.; visualization, V.Z.; supervision, R.B., L.A., F.B. and P.C.; project administration, V.Z.; funding acquisition, R.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from Ricerca Corrente 2018, Fondazione IRCCS Policlinico San Matteo, Pavia (Grant No. 08020418).

Data Availability Statement: All the data supporting the findings of this study can be found within the article.

Acknowledgments: We are grateful for the help and time of all the staff involved in the sampling process.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Justiz Vaillant, A.A.; Stang, C.M. *Lymphoproliferative Disorders*; NCBI Bookshelf; StatPearls: Treasure Island, FL, USA, 2022.
2. Hallek, M.; Cheson, B.D.; Catovsky, D.; Caligaris-Cappio, F.; Dighiero, G.; Döhner, H.; Hillmen, P.; Keating, M.; Montserrat, E.; Chiorazzi, N.; et al. IwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL. *Blood* **2018**, *131*, 2745–2760. [CrossRef] [PubMed]
3. Gianfelici, V.; Levato, L.; Molica, S. The Evolution of Targeted Therapies in Chronic Lymphocytic Leukaemia. *Curr. Hematol. Malig. Rep.* **2020**, *15*, 343–349. [CrossRef] [PubMed]

4. de Weerdt, I.; Koopmans, S.M.; Kater, A.P.; van Gelder, M. Incidence and Management of Toxicity Associated with Ibrutinib and Idelalisib: A Practical Approach. *Haematologica* **2017**, *102*, 1629–1639. [CrossRef] [PubMed]
5. Vétizou, M.; Pitt, J.M.; Daillère, R.; Lepage, P.; Waldschmitt, N.; Flament, C.; Rusakiewicz, S.; Routy, B.; Roberti, M.P.; Duong, C.P.M.; et al. Anticancer Immunotherapy by CTLA-4 Blockade Relies on the Gut Microbiota. *Science* **2015**, *350*, 1079–1084. [CrossRef] [PubMed]
6. Gopalakrishnan, V.; Spencer, C.N.; Nezi, L.; Reuben, A.; Andrews, M.C.; Karpinets, T.V.; Prieto, P.A.; Vicente, D.; Hoffman, K.; Wei, S.C.; et al. Gut Microbiome Modulates Response to Anti–PD-1 Immunotherapy in Melanoma Patients. *Science* **2018**, *359*, 97–103. [CrossRef] [PubMed]
7. Uribe-Herranz, M.; Klein-González, N.; Rodríguez-Lobato, L.G.; Juan, M.; Fernández de Larrea, C. Gut Microbiota Influence in Hematological Malignancies: From Genesis to Cure. *Int. J. Mol. Sci.* **2021**, *22*, 1026. [CrossRef] [PubMed]
8. Hakim, H.; Dallas, R.; Wolf, J.; Tang, L.; Schultz-Cherry, S.; Darling, V.; Johnson, C.; Karlsson, E.A.; Chang, T.-C.; Jeha, S.; et al. Gut Microbiome Composition Predicts Infection Risk During Chemotherapy in Children With Acute Lymphoblastic Leukemia. *Clin. Infect. Dis.* **2018**, *67*, 541–548. [CrossRef] [PubMed]
9. Huang, J.; Jiang, Z.; Wang, Y.; Fan, X.; Cai, J.; Yao, X.; Liu, L.; Huang, J.; He, J.; Xie, C.; et al. Modulation of Gut Microbiota to Overcome Resistance to Immune Checkpoint Blockade in Cancer Immunotherapy. *Curr. Opin. Pharmacol.* **2020**, *54*, 1–10. [CrossRef] [PubMed]
10. Lee, S.M.; Donaldson, G.P.; Mikulski, Z.; Boyajian, S.; Ley, K.; Mazmanian, S.K. Bacterial Colonization Factors Control Specificity and Stability of the Gut Microbiota. *Nature* **2013**, *501*, 426–429. [CrossRef] [PubMed]
11. Baruch, E.N.; Youngster, I.; Ortenberg, R.; Ben-Betzalel, G.; Katz, L.H.; Lahat, A.; Barshack, I.; Dick-Necula, D.; Mamtani, R.; Bloch, N.; et al. Fecal Microbiota Transplant Promotes Response in Immunotherapy-Refractory Melanoma Patients. *Science* **2021**, *371*, 602–609. [CrossRef] [PubMed]
12. Cuschieri, S. The STROBE Guidelines. *Saudi J. Anaesth.* **2019**, *13*, S31–S34. [CrossRef] [PubMed]
13. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef]
14. Oksanen, J.; Simpson, G.L.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.R.; O'Hara, R.; Solymos, P.; Stevens, M.H.; Szoecs, E.; et al. Vegan: Community Ecology Package. 2022. Available online: https://cran.r-project.org/web/packages/vegan/index.html (accessed on 20 March 2023).
15. Karpinets, T.V.; Wu, X.; Solley, T.; El Alam, M.B.; Sims, T.T.; Yoshida-Court, K.; Lynn, E.; Ahmed-Kaddar, M.; Biegert, G.; Yue, J.; et al. Metagenomes of Rectal Swabs in Larger, Advanced Stage Cervical Cancers Have Enhanced Mucus Degrading Functionalities and Distinct Taxonomic Structure. *BMC Cancer* **2022**, *22*, 945. [CrossRef] [PubMed]
16. Portincasa, P.; Bonfrate, L.; Vacca, M.; De Angelis, M.; Farella, I.; Lanza, E.; Khalil, M.; Wang, D.Q.-H.; Sperandio, M.; Di Ciaula, A. Gut Microbiota and Short Chain Fatty Acids: Implications in Glucose Homeostasis. *Int. J. Mol. Sci.* **2022**, *23*, 1105. [CrossRef]
17. Effendi, R.M.R.A.; Anshory, M.; Kalim, H.; Dwiyana, R.F.; Suwarsa, O.; Pardo, L.M.; Nijsten, T.E.C.; Thio, H.B. Akkermansia Muciniphila and Faecalibacterium Prausnitzii in Immune-Related Diseases. *Microorganisms* **2022**, *10*, 2382. [CrossRef] [PubMed]
18. Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermúdez-Humarán, L.G.; Gratadoux, J.-J.; Blugeon, S.; Bridonneau, C.; Furet, J.-P.; Corthier, G.; et al. *Faecalibacterium Prausnitzii* Is an Anti-Inflammatory Commensal Bacterium Identified by Gut Microbiota Analysis of Crohn Disease Patients. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16731–16736. [CrossRef] [PubMed]
19. Liu, W.; Ma, F.; Sun, B.; Liu, Y.; Tang, H.; Luo, J.; Chen, H.; Luo, Z. Intestinal Microbiome Associated With Immune-Related Adverse Events for Patients Treated With Anti-PD-1 Inhibitors, a Real-World Study. *Front. Immunol.* **2021**, *12*, 756872. [CrossRef] [PubMed]
20. Viaud, S.; Saccheri, F.; Mignot, G.; Yamazaki, T.; Daillere, R.; Hannani, D.; Enot, D.P.; Pfirschke, C.; Engblom, C.; Pittet, M.J.; et al. The Intestinal Microbiota Modulates the Anticancer Immune Effects of Cyclophosphamide. *Science* **2013**, *342*, 971–976. [CrossRef] [PubMed]
21. Pflug, N.; Kluth, S.; Vehreschild, J.J.; Bahlo, J.; Tacke, D.; Biehl, L.; Eichhorst, B.; Fischer, K.; Cramer, P.; Fink, A.-M.; et al. Efficacy of Antineoplastic Treatment Is Associated with the Use of Antibiotics That Modulate Intestinal Microbiota. *OncoImmunology* **2016**, *5*, e1150399. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Spectrum

RESEARCH ARTICLE

The COVID-19 Pandemic Sparked Off a Large-Scale Outbreak of Carbapenem-Resistant *Acinetobacter baumannii* from the Endemic Strains at an Italian Hospital

Greta Petazzoni,[a] Greta Bellinzona,[b] Cristina Merla,[a] Marta Corbella,[a] Vincenzina Monzillo,[a,c] Ørjan Samuelsen,[d,e] Jukka Corander,[f,g] Davide Sassera,[b] Stefano Gaiarsa,[a] Patrizia Cambieri[a]

[a]Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[b]Department of Biology and Biotechnology, University of Pavia, Pavia, Italy
[c]Department of Internal Medicine and Medical Therapy, University of Pavia, Pavia, Italy
[d]Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway
[e]Institute of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway
[f]Department of Biostatistics, University of Oslo, Oslo, Norway
[g]Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

Greta Petazzoni and Greta Bellinzona contributed equally to this article. The order was determined by the corresponding author after negotiation.

ABSTRACT *Acinetobacter baumannii* is a nosocomial pathogen that poses a serious threat due to the rise of incidence of multidrug-resistant (MDR) strains. During the COVID-19 pandemic, MDR *A. baumannii* clones have caused several outbreaks worldwide. Here, we describe a detailed investigation of an MDR *A. baumannii* outbreak that occurred at Policlinico San Matteo (Pavia, Italy). A total of 96 *A. baumannii* strains, isolated between January and July 2020 from 41 inpatients (both SARS-CoV-2 positive and negative) in different wards, were characterized by phenotypic and genomic analyses combining Illumina and Nanopore sequencing. Antibiotic susceptibility testing revealed that all isolates were resistant to carbapenems, and the sequence analysis attributed this to the carbapenemase gene bla_{OXA-23}. Virulence factor screening unveiled that all strains carried determinants for biofilm formation, while plasmid analysis revealed the presence of two plasmids, one of which was ~100 kbp long and encoded a phage sequence. A core genome-based phylogeny was inferred to integrate outbreak strain genomes with background genomes from public databases and the local surveillance program. All strains belonged to the globally disseminated sequence type 2 (ST2) clone and were mainly divided into two clades. Isolates from the outbreak clustered with surveillance isolates from 2019, suggesting that the outbreak was caused by two strains that were already circulating in the hospital before the start of the pandemic. The intensive spread of *A. baumannii* in the hospital was enhanced by the extreme emergency situation of the first COVID-19 pandemic wave that resulted in reduced attention to infection prevention and control practices.

IMPORTANCE The COVID-19 pandemic, especially during the first wave, posed a great challenge to the hospital management and generally promoted nosocomial pathogen dissemination. MDR *A. baumannii* can easily spread and persist for a long time on surfaces, causing outbreaks in health care settings. Infection prevention and control practices, epidemiological surveillance, and microbiological screening are fundamental in order to control such outbreaks. Here, we sequenced the genomes of 96 isolates from an outbreak of MDR *A. baumannii* strains using both short- and long-read technology in order to reconstruct the outbreak events in fine detail. The sequence data demonstrated that two endemic clones of MDR *A. baumannii* were the source of this large hospital outbreak during the first COVID-19 pandemic wave, confirming the

Editor Philip N. Rather, Emory University School of Medicine

Copyright © 2023 Petazzoni et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Stefano Gaiarsa, s.gaiarsa@smatteo.pv.it.

The authors declare no conflict of interest.

Received 5 November 2022
Accepted 1 March 2023
Published 23 March 2023

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

effect of COVID-19 emergency disrupting the protection provided by the use of the standard prevention procedures.

KEYWORDS *Acinetobacter baumannii*, COVID-19, Illumina, Nanopore, outbreak, phylogeny, multidrug resistance, plasmids

Acinetobacter baumannii is an opportunistic Gram-negative pathogen most frequently associated with health care-associated infections (HAIs). This pathogen is a member of the ESKAPE group, despite causing only 2% of HAIs (1), due to its virulence and the substantial prevalence of antibiotic resistance.

A. baumannii can evolve rapidly by acquiring determinants of resistance to a wide range of antibiotics (including front-line agents like carbapenems). Multidrug-resistant (MDR) strains are increasingly isolated in hospitals worldwide; this is a critical issue in southern Europe, especially in Italy, where the percentage of carbapenem-resistant *A. baumannii* (CRAB) strains during 2020 was 80.8% (2). In conjunction with antibiotic resistance, *A. baumannii* is concerning because of its exceptional capability to persist for long periods in the environment, even in the hospital setting, aided by biofilm production. This persistence on abiotic and biotic surfaces leads to chronic infections and facilitates its spread (3), especially in intensive care units (ICUs) (4). Indeed, *A. baumannii* infections are typically related to medical equipment such as catheters and ventilators (i.e., pneumonia and bloodstream infections), while urinary infections and those affecting skin, soft tissues, and surgical sites are less common (5). In light of this, the World Health Organization included MDR *A. baumannii* in the critical group, a list of bacteria that pose the greatest threat to human health and for which the development of novel antibiotics is urgently needed (6).

The COVID-19 pandemic has been posing an arduous challenge for hospitals worldwide, especially in the early period. The increasing COVID-19-related hospitalizations led to shortages in personnel, personal protective equipment, and medical equipment. These, in turn, often resulted in the impossibility of maintaining strict infection prevention and control (IPC) practices and, consequently, in an increase in bacterial and fungal infections. In fact, several *A. baumannii* outbreaks have been described during the first pandemic wave worldwide; e.g., in references 7 and 8. Fondazione IRCCS Policlinico San Matteo (Pavia, Italy) played a key role in the management of the COVID-19 pandemic in North Italy, especially during the first wave, being one of the largest hospitals (~900 beds) in the Lombardy region. An active genomic surveillance program and IPC containment strategies for ESKAPE pathogens were already in place in the hospital before COVID-19. Thanks to surveillance measures, the prevalence of *A. baumannii* in the 2018–2019 biennium was less than one isolate per 1,000 days of hospitalization. However, during the first wave of COVID-19, a large outbreak of MDR *A. baumannii* emerged (~2.8 isolates/1,000 days of hospitalization). Here, we report a detailed genomic characterization of this outbreak.

RESULTS

Data set characterization. During the first COVID-19 wave, an increase in incidence of *A. baumannii* strains was registered in the study hospital; the cases raised from a baseline of less than one isolate per 1,000 days of hospitalization in the 2018–2019 biennium to 2.8 isolates during the outbreak. We characterized 96 strains collected in this period (from January to July 2020) from 41 individuals hospitalized in the ICUs (n = 37), pneumology (n = 2), and otorhinolaryngology (n = 1) wards. Three samples were collected from one patient who was hospitalized in another care institute in Pavia (2 samples) and later admitted to the emergency department of the study hospital (1 sample). Patients had a mean age of 61 (standard deviation [SD], ±10 years), ranging from 30 to 83 years old, and were predominantly male (87.8%; n = 36). The median length of stay was 47 days (interquartile range [IQR], 25 to 74 days), and 43.9% (n = 18) of patients died (median length of stay, 52 days; IQR, 37 to 85 days). Thirty patients tested positive and seven negative for SARS-CoV-2; the remaining four were

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

hospitalized before the pandemic period. Of the *A. baumannii* isolates, 33.3% (n = 32) were related to lower respiratory tract infections and 22.9% (n = 22) to bloodstream infections; 3.1% of strains were isolated from the urinary tract (n = 3) and the other 40.6% from rectal swabs (n = 38) and nasal swab (n = 1). Table S1 in the supplemental material summarizes isolate and patient information.

All 96 isolates were sequenced using both long- and short-read technology. The hybrid assemblies enabled us to obtain 8 complete and 87 high-quality genomes (mean length of the genomes, 4 Mbp; mean number of contigs, 11.3 (Table S2)). The remaining one was discarded (sample 4964_2020) and excluded from all downstream analyses. Genome analysis revealed that all 95 isolates belonged to sequence type 2 (ST2), which is part of the globally disseminated international clone II (IC2) (9).

Phylogenetic analysis. In order to further characterize the outbreak, we performed a phylogenetic analysis using the 95 *A. baumannii* genomes from the outbreak, 23 from the hospital surveillance program, and 367 from the PATRIC database as background. The inferred phylogeny (Fig. 1) depicts a situation in which the *A. baumannii* outbreak in the hospital was caused by multiple strains. The majority of the outbreak isolates (n = 93) are divided into two monophyletic clades, one larger (cluster 1; 71 outbreak strains in a clade of 80 total genomes; single nucleotide polymorphism (SNP) distance among cluster 1 genomes, 0 to 100) and one smaller (cluster 2; 23 outbreak strains; SNP distance among cluster 2 genomes, 0 to 43). Cluster 1 also includes nine isolates from the hospital surveillance program (one from 2019, four from 2020, and four from 2021) and one from the PATRIC DB. Cluster 2 is exclusively composed of 2020 isolates; nonetheless, isolate 4200_2019 from 2019 is basal to this cluster. A background isolate from Belgium (470.13595 of PATRIC DB) interposes between the 2019 surveillance isolate and cluster 2; however, the common branch to these genomes has very limited support (29/100 bootstraps). We measured the SNP distance between all genomes of cluster 1 and each one of cluster 2, which ranged between 274 and 356.

Interestingly, the genome of one of the two isolates from January 2020 not included in the main clusters (4636_2020) formed a monophyletic group with four 2019 surveillance genomes, while the second one (4614_2020) clustered with one surveillance genome from late 2020.

Antibiotic resistance and virulence. All 95 outbreak isolates were resistant to carbapenems (100% to imipenem and meropenem). Almost all isolates that presented resistance to aminoglycosides: 97.9% (n = 93) were found to be resistant to gentamicin and 98.9% (n = 94) to amikacin. Moreover, only 8.9% (n = 7) of isolates tested were resistant to tigecycline, and eight isolates were susceptible with increased exposure to this antibiotic. Additionally, the totality of isolates tested for ciprofloxacin (n = 90) were resistant. None of the isolates was found resistant to colistin. The antibiotic resistance factors predicted in the genomes were concordant with the phenotypes. Resistance to carbapenems was confirmed by the presence of $bla_{\text{OXA-23}}$. The intrinsic $bla_{\text{OXA-51}}$-variant, $bla_{\text{OXA-66}}$, was also detected, but no ISAba1 sequence insertion was found upstream. All strains carried a gene of the $bla_{\text{ADC-25}}$ family: the majority (n = 89) presented $bla_{\text{ADC-73}}$, another five presented $bla_{\text{ADC-56}}$, and the remaining one had $bla_{\text{ADC-30}}$ (Table S3).

Virulence determinants were also found in outbreak isolates, particularly those involved in biofilm formation. In detail, all isolates presented the *ompA* gene, which is important for the development of robust biofilms on abiotic surfaces (10), and the *bap* gene, which ensures a strong biofilm formation (11). Ninety-four out of 95 isolates had *csuA/B*, *csuA*, *csuB*, *csuC*, *csuD*, and *csuE* that encode the Chaperone-Usher secretion (CUS) system, fundamental for the first bacterial attachment on abiotic surfaces (12). The sole strain not carrying the CUS system was isolated in January 2020 (4614_2020).

Strains also had genes for BfmS/BfmR, a two-component system that plays a key role in CUS system regulation and *ompA* (13). Finally, 23.3% (n = 25) of the isolates presented the entire cluster of four genes (*pgaA*, *pgaB*, *pgaC*, and *pgaD*) responsible for the production of poly-β-(1,6)-*N*-acetylglucosamine (PNAG), a component of biofilm (12). Interestingly, these isolates included two from January 2020 (4636_2020, 4614_2020)

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

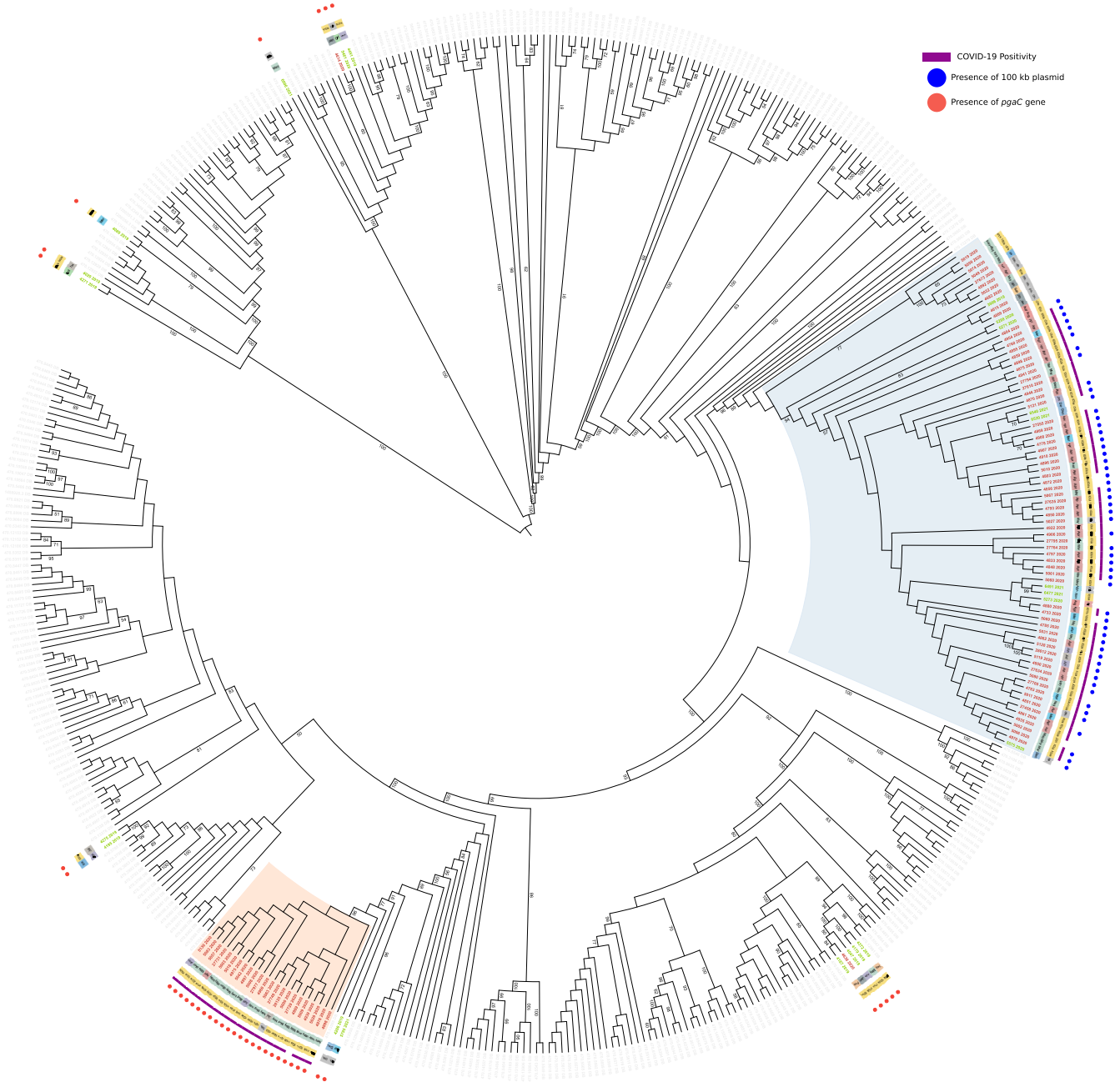


FIG 1 Maximum-likelihood phylogeny of 485 *A. baumannii* strains, including 95 outbreak genomes (red labels), 23 surveillance genomes (green labels), and 367 PATRIC genomes (gray labels) inferred on coreSNPs with RAxML. Date of isolation (innermost crown), ward (second crown), and SARS-CoV-2 positivity (third crown) were retrieved from the hospital and microbiological database. The presence of 100-kbp plasmid (blue dots) and *pgaC* gene (red dots) were determined by genomic analyses. The two main outbreak clusters are highlighted in blue (cluster 1) and red (cluster 2). Bootstrap values above 50/100 are indicated on tree branches. This figure was obtained using iTol (https://itol.embl.de/).

and those in cluster 2. In cluster 1 genomes (n = 70), the *pgaC* gene was instead interrupted by the ISAba1 insertion sequence (Table S4).

Outbreak strains and those from the genomic surveillance program were tested for biofilm formation capability. The majority (n = 64) were found to be weak biofilm producers, while 40 were moderate producers, 13 strong producers, and only 1 was not a biofilm producer (Fig. S1). Interestingly, the isolates in the two large phylogenetic clusters tended to produce a weaker biofilm than the other isolates described in this work: in cluster 1, 67.50% are classified as weak, 28.75% are moderate, and 3.75% are strong producers; in cluster 2, 43.48% are weak producers, 47.82% are moderate, and 8.70%

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

are strong; and the remaining isolates are 6.66% nonproducers, 6.66% weak, 33.33% moderate, and 53.33% strong producers.

Plasmids. We investigated the presence of plasmids in the outbreak genomes and found that all the samples shared a plasmid of 8,731 bp, which perfectly matches GenBank accession no. NZ_CP008708 containing 12 coding sequences, 10 being hypothetical proteins. Most of the genomes of cluster 1 (51/71) presented also a second, 106,175-bp-long, plasmid (highly similar to GenBank accession no. CP081146). Neither of the two plasmids encoded known virulence or resistance genes. We analyzed the ~100-kbp plasmid in depth; 83 out of its 131 coding sequences were annotated as hypothetical, 20 of which were reported to be phage associated. Indeed, the majority of the plasmid sequence consisted of a prophagic region of ~85 kbp (9261 to 95595). The phylogeny in Fig. 1 indicates that this plasmid is present only in cluster 1. Genomes in the basal position of the cluster do not carry the plasmid, suggesting that it was acquired with a single transfer event at the beginning of the outbreak. Also, the pattern of plasmid presence within the cluster indicates multiple losses. Lastly, we investigated the presence of the ~8-kbp and the ~100-kbp plasmids among the overall *A. baumannii* population (n = 8,969 genomes) and found them to be differentially distributed (Fig. S2). The first one was present in 32.1% (n = 2,882), and the second one was in 7% (n = 636) of the population.

DISCUSSION

In this work, we present the data from a hospital-wide outbreak of MDR *A. baumannii* that occurred during the first wave of the COVID-19 pandemic. A rise in the incidence of HAIs during COVID-19 pandemic waves was observed in the study hospital and also in other hospitals and countries worldwide (14–16). This is particularly true for *A. baumannii*, which was the most frequently reported cause of hospital outbreaks, together with methicillin-resistant *Staphylococcus aureus* and fungal infections (17–19). For example, during the first wave, Perez et al. reported a hospital-acquired carbapenem-resistant *A. baumannii* (CRAB) outbreak of 34 isolates harboring $bla_{OXA\text{-}23}$, which involved 34 COVID-19 patients admitted to ICUs in New Jersey hospital (7). Also, Gottesman and colleagues (8) reported an outbreak in Israel of five CRAB carrying $bla_{OXA\text{-}24}$ and belonging to IC2 in ICUs patients.

In our study, we were able to identify two large phylogenetic clusters (Fig. 1), which included most of the outbreak strains. The inclusion of genomes sequenced before and after the outbreak by the hospital surveillance program revealed that multiple strains were circulating in the hospital before the start of the pandemic, including the two (3996_2019 and 4200_2019) that caused the outbreak.

We detected two plasmids within our genomes, one of ~8 kbp and another of ~100 kbp. The second one is specific to a subclade of cluster 1 and contains a large prophage sequence carrying mostly uncharacterized genes. Four strains (4682_2020, 3996_2019, 4615_2020, and 4669_2020), which are basal to the subclade, were isolated in 2019 and early 2020, and their genomes did not contain the ~100-kbp plasmid. This suggests that the plasmid was acquired later during the outbreak, possibly within the hospital. The hypothesis is reinforced by the presence of another subclade of this cluster (n genomes = 8) in which the plasmid is absent. Moreover, this phage-encoding plasmid is supposedly specific to *A. baumannii*, as it is not present in other species (BLASTN on the nucleotide [nt] database performed on 4 October 2022). This observation suggests that the source of acquisition could have been another strain of *A. baumannii*, not detected by our surveillance program. Despite its fitness advantage not being clear, we can observe a tendency to preserve the plasmid in the genomes analyzed. This trend can be observed both in the local phylogeny (Fig. 1) and in the global one (see Fig. S2 in the supplemental material).

Mapping patient information on the tree revealed that SARS-CoV-2 coinfection was approximately equally distributed between the two phylogenetic clusters. Thus, no association between *A. baumannii* clusters and viral infection can be hypothesized. The

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

possibility of transmission of *A. baumannii* between COVID-19-positive and -negative patients in the absence of viral cotransmission is interesting and can be explained by the high capability of *A. baumannii* to survive on abiotic surfaces due to biofilm formation (20). In turn, this could also explain why this bacterial species thrived during the emergency period, both in the study hospital and in other institutions worldwide. Moreover, no association was found between clusters and wards, outcome, or age of patients.

The phylogeny we inferred allowed us to identify eight separate clusters of study genomes (including the two large ones), four of which contained strains isolated months before and after the epidemic event. This observation unveils the presence of at least four endemic clones in the study hospital, two of which gave origin to the large outbreak described in this work. Interestingly, no genomic factor that is known to enhance bacterial spread was found in the study genomes. Moreover, the biofilm formation capability of the isolates in the two large clusters was generally weaker than that of the other isolate described in this work (Fig. S1). This further suggests that the strains responsible for the outbreak did not have peculiar characteristics which could have favored their prolification. In the genomes of cluster 1, we observed the interruption of the *pgaC* gene by the mobile element *ISAba1*. This gene contributes to the formation of PNAG (21), and it was observed that the inactivation of this gene leads to a weaker adhesion capability in *Klebsiella pneumoniae* (22). In our work, we observed that the isolates in cluster 1 tend to be weaker biofilm producers than all others (including those in cluster 2), which have a functional *pgaC* gene. Cluster 1 is the largest, and it possibly includes the most diffusive clone in this study. Yet this cannot be attributed to a peculiar biofilm production capability. All in all, these results indicate that the spread of the bacterial pathogen was enhanced solely by the extreme emergency situation caused by the first wave of the COVID-19 pandemic. Personnel and bed shortages, and the prioritization of self-preservation from the viral infection of health care workers, led to unavoidable diminished attention toward the standard patient-handling measures (e.g., changing gloves between patients). In accordance with this hypothesis, the incidence of *A. baumannii* decreased to lower than 1 isolate per 1,000 days of hospitalizations after the first pandemic wave and remained under control during the following waves, thanks to better management of COVID-19 patients and a better knowledge of the viral disease.

MATERIALS AND METHODS

Data collection. This retrospective study was conducted at IRCCS Fondazione Policlinico San Matteo in Pavia (Italy), where 151 isolates of *A. baumannii* were collected from 6 January 2020 to 30 July 2020. Ninety-six isolates were chosen for use in this study; the choice was made favoring the inclusion of both samples from colonization and infection materials for each patient (where available) and of multiple samples in case of long hospitalizations (1 sample/7 days/material). The isolates selected were identified with matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS; Bruker Daltonik, Bremen, Germany) equipped with Bruker Biotyper 3.1 software. Antibiotic susceptibility was tested with both the Sensititre Gram-negative EUMDROXF plate (Thermo Fisher Scientific, Rodano, Italy) and Phoenix NMIC-402 panel loaded on a BD Phoenix M50 instrument. Values were interpreted according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) clinical breakpoints (23). Tigecycline was interpreted following Clinical and Laboratory Standards Institute (CLSI) breakpoints (24).

Patient information was retrieved from the hospital database. COVID-19 positivity was defined at the time of admission, consulting the results of real-time reverse transcriptase PCR tests for the presence of SARS-CoV-2.

DNA extraction and genome characterization. The genomic DNA of the 96 strains was extracted with DNeasy blood and tissue kit (Qiagen) for short-read sequencing using an Illumina platform and with MagAttract high-molecular-weight (HMW) DNA kit (Qiagen) for long-read sequencing using Nanopore technology. Short and long reads were combined with the Unicycler software (25) to perform hybrid assemblies. The genomes obtained were then *in silico* sequence typed (ST) using the Pasteur multilocus sequence typing (MLST) scheme (26) with an in-house script. The presence of virulence and resistance genes was assessed using ABRicate (https://github.com/tseemann/abricate) with the NCBI AMRFinderPlus (27), CARD (28), Resfinder (29), and VirulenceFinder (30) databases. Results were manually curated and double-checked using BLAST and long reads. The presence of insertion sequences was determined with MobileElementFinder (31).

The presence of plasmids was evaluated with PlasmidFinder (32), and a more in-depth plasmid search was performed by visualizing the assembly graph in Bandage (33) for the least fragmented isolates, checking for circularity and higher depth coverage. The presence of the plasmids identified was

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

assessed in all the remaining assemblies using BLASTN. Plasmids were annotated combining Prokka (34), NCBI Conserved Domain Database (35), and possibly blastp on the nonredundant (nr) database. PHASTER (36) was then used to identify and annotate prophage sequences. In order to understand the prevalence of the plasmids found with our approach in the *A. baumannii* overall population, we created a collection of *A. baumannii* genomes, joining those from this study and all high-quality genomes from the PATRIC DB (sequences were retrieved in July 2022, using the makepdordb.py script of the P-DOR pipeline [https://github.com/SteMIDIfactory/P-DOR]). All assemblies were then scanned for the presence of all plasmid genes; genomes were considered to carry a plasmid when they encoded the vast majority of its genes (≥11 genes for the ~8 kbp; ≥120 genes for the ~100 kbp). Finally, FastTree (37) was used to infer the global phylogeny of the species, thus allowing us to study the distribution of the plasmids among the population.

Surveillance program genomes. Strains selected by the genomic surveillance program of the hospital are routinely processed for short-read sequencing as follows: DNA is extracted using DNeasy blood and tissue kit (Qiagen), sequenced with an Illumina platform, and assembled using Unicycler (25) with default parameters. We retrieved all the ST2 genomes obtained by the surveillance program ($n = 23$) in the months before and after the outbreak period (12 from 2019, 5 from 2020, and 6 from 2021), and we added them to our data set.

Biofilm formation assay. All *A. baumannii* isolates ($n = 119$) were tested for the ability to produce biofilm following the protocol described by Kumari and colleagues (38). *Pseudomonas aeruginosa* ATCC 27853 and *Staphylococcus epidermidis* 30678 from the internal laboratory collection (characterized to be a strong biofilm producer) were used as positive controls. Sterile medium was used as a negative control. Interpretation of optical density (OD) measurements was performed as previously described in reference 39. In detail, the cutoff value (c) was calculated as three standard deviations (SDs) above the average optical density (OD) of the negative control. Strains were classified into the following four categories based upon mean OD values: "no producer" when mean OD is $\leq c$, "weak producer" when c is less than the mean OD, which is $\leq 2c$, "moderate producer" when $2c$ is less than the mean OD, which is $\leq 4c$, and "strong producer" when mean OD is $> 4c$.

Phylogenetic analyses. In order to characterize the outbreak of *A. baumannii*, a genomic background was constructed. In detail, the 20 high-quality genomes most similar to those in our set according to k-mer content similarity (MASH [40]) were retrieved from the PATRIC DB (https://www.patricbrc.org). Then, the P-DOR pipeline (https://github.com/SteMIDIfactory/P-DOR; version beta1.5) was used to align all genomes in our data set to an internal reference (the complete genome of strain 4946) and to extract core SNPs. Maximum-likelihood phylogeny was inferred with RAxML (41) (using 100 bootstraps resamples) on the resulting core SNP alignment using the general time reversible model, as suggested by ModelTest-NG (42), with ascertainment bias correction (43).

Ethical statement. The study was designed and conducted in accordance with the Helsinki Declaration and approved by the Ethics Committee of Fondazione IRCCS Policlinico San Matteo in Pavia, Italy (internal project code, 08022298/13; project number 733-rcr2013-34). The work described herein is a retrospective study performed on bacterial isolates from human samples that were obtained as part of routine hospital care.

Data availability. Assembled genomes of outbreak and surveillance *A. baumannii* isolates are deposited in the National Center for Biotechnology Information (NCBI) under the BioProject accession no. PRJNA844013.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.05 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 5, PDF file, 1 MB.

ACKNOWLEDGMENTS

We acknowledge Anna K. Pöntinen and Francois Cleon for experimental assistance and the Genomics Support Centre Tromsø, UiT The Arctic University of Norway, for the short- and long-read sequencing.

J.C. was supported by ERC grant 742158. S.G. was supported by Ricerca Corrente grant 924-rcr2018i-34 (Italian Ministry of Health).

REFERENCES

1. Lob SH, Hoban DJ, Sahm DF, Badal RE. 2016. Regional differences and trends in antimicrobial susceptibility of Acinetobacter baumannii. Int J Antimicrob Agents 47:317–323. https://doi.org/10.1016/j.ijantimicag.2016.01.015.
2. Yang C-H. 2022. WHO regional office for Europe/European centre for disease prevention and control. antimicrobial resistance surveillance in Europe 2022 to 2020 data. WHO Regional Office for Europe, Copenhagen, Denmark.
3. Yang C-H, Su P-W, Moi S-H, Chuang L-Y. 2019. Biofilm formation in Acinetobacter baumannii: genotype-phenotype correlation. Molecules 24:1849. https://doi.org/10.3390/molecules24101849.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

4. Antunes LCS, Visca P, Towner KJ. 2014. Acinetobacter baumannii: evolution of a global pathogen. Pathog Dis 71:292–301. https://doi.org/10.1111/2049-632X.12125.
5. Wong D, Nielsen TB, Bonomo RA, Pantapalangkoor P, Luna B, Spellberg B. 2017. Clinical and pathophysiological overview of Acinetobacter infections: a century of challenges. Clin Microbiol Rev 30:409–447. https://doi.org/10.1128/CMR.00058-16.
6. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outterson K, Patel J, Cavaleri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18:318–327. https://doi.org/10.1016/S1473-3099(17)30753-3.
7. Perez S, Innes GK, Walters MS, Mehr J, Arias J, Greeley R, Chew D. 2020. Increase in hospital-acquired carbapenem-resistant Acinetobacter baumannii infection and colonization in an acute care hospital during a surge in COVID-19 admissions—New Jersey, February-July 2020. MMWR Morb Mortal Wkly Rep 69:1827–1831. https://doi.org/10.15585/mmwr.mm6948e1.
8. Gottesman T, Fedorowsky R, Yerushalmi R, Lellouche J, Nutman A. 2021. An outbreak of carbapenem-resistant Acinetobacter baumannii in a COVID-19 dedicated hospital. Infect Prev Pract 3:100113. https://doi.org/10.1016/j.infpip.2021.100113.
9. Dijkshoorn L, Aucken H, Gerner-Smidt P, Janssen P, Kaufmann ME, Garaizar J, Ursing J, Pitt TL. 1996. Comparison of outbreak and nonoutbreak Acinetobacter baumannii strains by genotypic and phenotypic methods. J Clin Microbiol 34:1519–1525. https://doi.org/10.1128/jcm.34.6.1519-1525.1996.
10. Gaddy JA, Tomaras AP, Actis LA. 2009. The Acinetobacter baumannii 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. Infect Immun 77:3150–3160. https://doi.org/10.1128/IAI.00096-09.
11. Azizi O, Shahcheraghi F, Salimizand H, Modarresi F, Shakibaie MR, Mansouri S, Ramazanzadeh R, Badmasti F, Nikbin V. 2016. Molecular analysis and expression of bap gene in biofilm-forming multi-drug-resistant Acinetobacter baumannii. Rep Biochem Mol Biol 5:62–72.
12. Tomaras AP, Dorsey CW, Edelmann RE, Actis LA. 2003. Attachment to and biofilm formation on abiotic surfaces by Acinetobacter baumannii: involvement of a novel chaperone-usher pili assembly system. Microbiology (Reading) 149:3473–3484. https://doi.org/10.1099/mic.0.26541-0.
13. Kishii K, Hamada M, Aoki K, Ito K, Onodera J, Ishii Y, Tateda K. 2020. Differences in biofilm formation and transcription of biofilm-associated genes among Acinetobacter baumannii clinical strains belonging to the international clone II lineage. J Infect Chemother 26:693–698. https://doi.org/10.1016/j.jiac.2020.02.017.
14. Garcia-Vidal C, Sanjuan G, Moreno-García E, Puerta-Alcalde P, Garcia-Pouton N, Chumbita M, Fernandez-Pittol M, Pitart C, Inciarte A, Bodro M, Morata L, Ambrosioni J, Grafia I, Meira F, Macaya I, Cardozo C, Casals C, Tellez A, Castro P, Marco F, García F, Mensa J, Martínez JA, Soriano A, COVID-19 Researchers Group. 2021. Incidence of co-infections and superinfections in hospitalized patients with COVID-19: a retrospective cohort study. Clin Microbiol Infect 27:83–88. https://doi.org/10.1016/j.cmi.2020.07.041.
15. Cusumano JA, Dupper AC, Malik Y, Gavioli EM, Banga J, Berbel Caban A, Nadkarni D, Obla A, Vasa CV, Mazo D, Altman DR. 2020. Staphylococcus aureus bacteremia in patients infected with COVID-19: a case series. Open Forum Infect Dis 7:ofaa518. https://doi.org/10.1093/ofid/ofaa518.
16. Contou D, Claudinon A, Pajot O, Micaëlo M, Longuet Flandre P, Dubert M, Cally R, Logre E, Fraissé M, Mentec H, Plantefève G. 2020. Bacterial and viral co-infections in patients with severe SARS-CoV-2 pneumonia admitted to a French ICU. Ann Intensive Care 10:119. https://doi.org/10.1186/s13613-020-00736-x.
17. Pascale R, Bussini L, Gaibani P, Bovo F, Fornaro G, Lombardo D, Ambretti S, Pensalfine G, Appolloni L, Bartoletti M, Tedeschi S, Tumietto F, Lewis R, Viale P, Giannella M. 2022. Carbapenem-resistant bacteria in an intensive care unit during the coronavirus disease 2019 (COVID-19) pandemic: a multicenter before-and-after cross-sectional study. Infect Control Hosp Epidemiol 43:461–466. https://doi.org/10.1017/ice.2021.144.
18. O'Toole RF. 2021. The interface between COVID-19 and bacterial healthcare-associated infections. Clin Microbiol Infect 27:1772–1776. https://doi.org/10.1016/j.cmi.2021.06.001.
19. Kundu R, Singla N. 2022. COVID-19 and plethora of fungal infections. Curr Fungal Infect Rep 16:47–54. https://doi.org/10.1007/s12281-022-00432-2.
20. Eze EC, Chenia HY, El Zowalaty ME. 2018. Acinetobacter baumannii biofilms: effects of physicochemical factors, virulence, antibiotic resistance determinants, gene regulation, and future antimicrobial treatments. Infect Drug Resist 11:2277–2299. https://doi.org/10.2147/IDR.S169894.
21. Choi AHK, Slamti L, Avci FY, Pier GB, Maira-Litrán T. 2009. The pgaABCD locus of Acinetobacter baumannii encodes the production of poly-beta-1–6-N-acetylglucosamine, which is critical for biofilm formation. J Bacteriol 191:5953–5963. https://doi.org/10.1128/JB.00647-09.
22. Chen K-M, Chiang M-K, Wang M, Ho H-C, Lu M-C, Lai Y-C. 2014. The role of pgaC in Klebsiella pneumoniae virulence and biofilm formation. Microb Pathog 77:89–99. https://doi.org/10.1016/j.micpath.2014.11.005.
23. The European Committee on Antimicrobial Susceptibility Testing. 2020. Breakpoint tables for interpretation of MICs and zone diameters. version 10.0.
24. Clinical and Laboratory Standards Institute. 2020. Performance Standards for Antimicrobial Susceptibility Testing. CLSI, CLSI supplement M100, 30th ed, Wayne, PA.
25. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
26. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S. 2010. The population structure of Acinetobacter baumannii: expanding multiresistant clones from an ancestral susceptible genetic pool. PLoS One 5:e10034. https://doi.org/10.1371/journal.pone.0010034.
27. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 63:e00483-19. https://doi.org/10.1128/AAC.00483-19.
28. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 45:D566–D573. https://doi.org/10.1093/nar/gkw1004.
29. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261.
30. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510. https://doi.org/10.1128/JCM.03617-13.
31. Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. 2021. Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. J Antimicrob Chemother 76:101–109. https://doi.org/10.1093/jac/dkaa390.
32. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.
33. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383.
34. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
35. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39:D225–D229. https://doi.org/10.1093/nar/gkq1189.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

36. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–21. https://doi.org/10.1093/nar/gkw387.
37. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. https://doi.org/10.1093/molbev/msp077.
38. Kumari M, Bhattarai NR, Rai K, Pandit TK, Khanal B. 2022. Multidrug-resistant Acinetobacter: detection of ESBL, MBL, blaNDM-1 genotype, and biofilm formation at a tertiary care hospital in eastern Nepal. Int J Microbiol 2022:8168000. https://doi.org/10.1155/2022/8168000.
39. Stepanović S, Vuković D, Hola V, Di Bonaventura G, Djukić S, Cirković I, Ruzicka F. 2007. Quantification of biofilm in microtiter plates: overview of testing conditions and practical recommendations for assessment of biofilm production by staphylococci. APMIS 115:891–899. https://doi.org/10.1111/j.1600-0463.2007.apm_630.x.
40. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17:132. https://doi.org/10.1186/s13059-016-0997-x.
41. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
42. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291–294. https://doi.org/10.1093/molbev/msz189.
43. Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925. https://doi.org/10.1080/106351501753462876.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Spectrum

RESEARCH ARTICLE

Combining Genome Surveillance and Metadata To Characterize the Diversity of *Staphylococcus aureus* Circulating in an Italian Hospital over a 9-Year Period

U. Postiglione,[a] G. Batisti Biffignandi,[a] M. Corbella,[b] C. Merla,[b] E. Olivieri,[c] G. Petazzoni,[b] E. J. Feil,[d] C. Bandi,[e] P. Cambieri,[b] S. Gaiarsa,[b] M. Brilli,[e] D. Sassera[a,f]

[a]Department of Biology and Biotechnology, University of Pavia, Pavia, Italy
[b]Microbiology and Virology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[c]Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Pavia, Italy
[d]The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom
[e]Department of Bioscience, University of Milan, Milan, Italy
[f]Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

U. Postiglione and G. Batisti Biffignandi contributed equally. Author order was determined based on seniority.
M. Brilli and D. Sassera contributed equally. Author order was determined based on seniority.

ABSTRACT *Staphylococcus aureus* is an opportunistic pathogen and a leading cause of morbidity and mortality worldwide. Genomic-based surveillance has greatly improved our ability to track the emergence and spread of high-risk clones, but the full potential of genomic data is only reached when used in conjunction with detailed metadata. Here, we demonstrate the utility of an integrated approach by leveraging a curated collection of clinical and epidemiological metadata of *S. aureus* in the San Matteo Hospital (Italy) through a semisupervised clustering strategy. We sequenced 226 sepsis *S. aureus* samples, recovered over a period of 9 years. By using existing antibiotic profiling data, we selected strains that capture the full diversity of the population. Genome analysis revealed 49 sequence types, 16 of which are novel. Comparative genomic analyses of hospital- and community-acquired infection ruled out the existence of genomic features differentiating them, while evolutionary analyses of genes and traits of interest highlighted different dynamics of acquisition and loss between antibiotic resistance and virulence genes. Finally, highly resistant clones belonging to clonal complexes (CC) 8 and 22 were found to be responsible for abundant infections and deaths, while the highly virulent CC30 was responsible for rare but deadly episodes of infections.

IMPORTANCE Genome sequencing is an important tool in clinical microbiology, as it allows in-depth characterization of isolates of interest and can propel genome-based surveillance studies. Such studies can benefit from *ad hoc* methods of sample selection to capture the genomic diversity present in a data set. Here, we present an approach based on clustering of antibiotic resistance profiles that allows optimal sample selection for bacterial genomic surveillance. We apply the method to a 9-year collection of *Staphylococcus aureus* from a large hospital in northern Italy. Our method allows us to sequence the genomes of a large variety of strains of this important pathogen, which we then leverage to characterize the epidemiology in the hospital and to perform evolutionary analyses on genes and traits of interest. These analyses highlight different dynamics of acquisition and loss between antibiotic resistance and virulence genes.

KEYWORDS *Staphylococcus aureus*, antibiotic resistance, community acquired, genomic surveillance, hospital acquired, risk factors

Editor Se-Ran Jun, University of Arkansas for Medical Sciences

Copyright © 2023 Postiglione et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to M. Brilli, matteo.brilli@unimi.it, or D. Sassera, davide.sassera@unipv.it.

The authors declare no conflict of interest.

Received 15 March 2023
Accepted 18 June 2023
Published 17 July 2023

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

Staphylococcus aureus is a globally disseminated bacterial species capable of thriving both free-living in the environment and associated with eukaryotic hosts (1). In humans, *S. aureus* is present in the nasal microbiome of about 30% of the healthy population (2), where it usually is a harmless commensal; however, it can also act as an opportunistic pathogen on immune-deficient hosts or body sites exposed during surgery. *S. aureus* can cause a plethora of clinical manifestations, ranging from superficial skin infections to invasive, difficult to eradicate, and potentially fatal conditions, such as bacteremia, pneumonia, endocarditis, and osteomyelitis (3). A key factor proposed to explain the success of *S. aureus* both as a colonizer and as a pathogen is the ability to acquire novel genes and elements (e.g., transposons, bacteriophages, insertion sequences, pathogenicity islands) by horizontal gene transfer (HGT), although chromosomal mutations in the core genome can also play a key role (4). The frequent gain (or loss) of pathogenicity and resistance genes underpins the rapid evolution of high-risk clones, such as those resistant to methicillin (methicillin-resistant *S. aureus* [MRSA]) (5) that are now widespread in both health care and community settings.

A central question of *S. aureus* epidemiology concerns the extent of overlap between resistant clones circulating in health care settings, which are frequently exposed to antibiotics, and those in the community, which are broadly considered to be more susceptible to antibiotics (6). Early studies that compared community- and healthcare-acquired (CA and HA) strains highlighted key molecular differences, such as differential expression of the cytotoxin Panton-Valentine leukocidin (PVL, more common in CA strains) (7), different staphylococcal chromosomal cassette (SCC) *mec* types (types IV and V are more common in, but not restricted to, CA strains) (6–8). In addition to these genotypic characteristics, it was shown that CA strains tend to affect healthier and younger patients, mostly causing minor skin and soft tissue infections (9). With the advent of whole-genome sequencing (WGS), more exhaustive comparisons have been performed which challenged the view that CA and HA strains can be clearly distinguished on the basis of genomic features (10–16).

S. aureus was the first species for which population-scale genome sequencing was exploited to understand local and global epidemiology of high-risk clones (17). Subsequent studies helped to determine which clones were more frequent in specific settings and worldwide (18–20) and in environments (21, 22). These population-wide studies provide important baseline data and complement studies focusing on specific lineages at global or local levels (23, 24). These data have shed light on broad aspects of *S. aureus* epidemiology, such as hypervirulence, specific lineages, and the reconstruction of transmission chains during outbreaks (25–27).

Studies based on sequencing representative samples of the *S. aureus* population "'snapshots"' (20) and studies focused on individual lineages are therefore both valuable for understanding how the frequency of specific clones may change over time and from region to region. However, these approaches may not capture the full diversity of the population, and rare genotypes may be overlooked. Here, we take an alternative approach, aimed to maximize the diversity of strains for sequencing in hospital settings by exploiting archived antibiograms as proxies of genetic relatedness in a semisupervised sampling. We apply our strategy on a 9-year data set from an Italian hospital, and we show that it allowed us to sequence more rare strains, providing a more complete picture of the diversity existing in the hospital. We then combined the detected genomic diversity with the available metadata to examine potential differences between CA and HA strains and to investigate risk factors and phyletic patterns of genes and traits of interest.

RESULTS AND DISCUSSION

Data set summary and antibiotic resistance. Our data set comprises a total of 7,523 records from different hospital wards and clinical samples, from 5,233 patients who tested positive for *S. aureus* in 2011 to 2019. Of these patients, 2,807 stayed at the hospital at least overnight and are thus considered inpatients, while the remaining 2,426 were classified as outpatients. Of the 7,523 records, 1,404 (26.8%) are methicillin resistant (MRSA), a resistance that was more common in samples from inpatients (32.2%) than in outpatients (20.5%). MRSA isolates from inpatients were enriched ($P < 2.2e-16$) in blood (22%) and respiratory

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

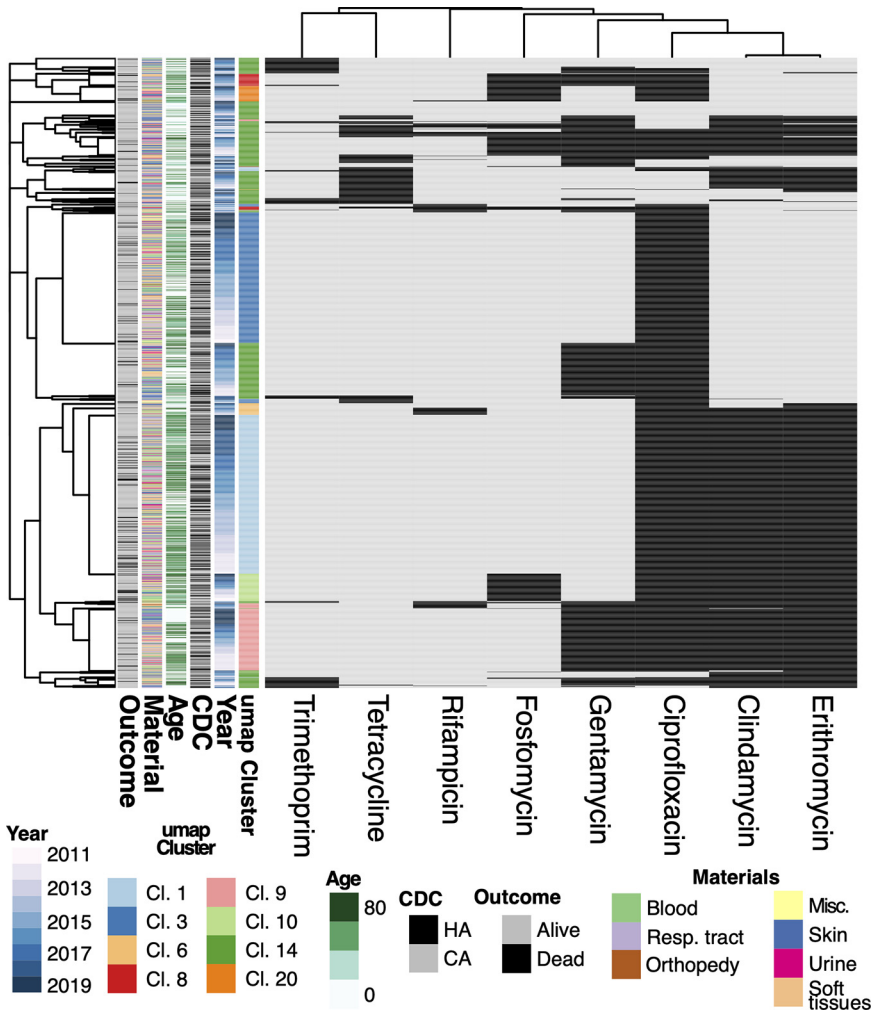


FIG 1 Summary of the MRSA samples grouped on the basis of similarities in resistance profiles for 8 antibiotics, chosen because the percentage of resistant strains varied between 5% and 99%. Black squares on the right indicate resistance, while gray squares indicate susceptibility. Additional metadata are shown in the six columns on the left, highlighting a relatively homogeneous distribution for outcome, material, age, CDC classification, and year of sampling with respect to the patterns of resistance.

tract (22%) samples, while those from outpatients were more often from skin (39%) and soft tissue (35%) (see Fig. S1 in the supplemental material). Fig. 1 shows a clustering of the MRSA isolates based on profiles of resistance to the eight antibiotics for which the percentage of resistant strains varied between 5% and 99%.

Linear regression estimated MRSA isolates to have decreased from 34.8% in 2011 to 23.5% in 2019 ($P = 0.0015$). This change in percentage is not due to a reduction of MRSA in absolute numbers, but to the increase of yearly methicillin-susceptible *S. aureus* (MSSA) isolates, which in the analyzed time span increased from 73 in 2011 to 523 in 2019 (beta = 16.3, $P = 0.002$).

Resistance to penicillin, ciprofloxacin, erythromycin, clindamycin and gentamicin were the most common across both MRSA and MSSA isolates, with the former characterized on average by a larger number of resistances per isolate (Table S1). Resistance to teicoplanin, daptomycin, and linezolid were found in less than 1% of isolates (generally regarded as false positives; were not used for downstream analyses), while resistance to vancomycin was not detected. Linear regression highlighted temporal trends in the relative abundance of resistance. MRSA strains resistant to clindamycin, erythromycin, fosfomycin, ciprofloxacin, and rifampicin decreased in time, while the percentage of MSSA strains resistant to clindamycin and erythromycin increased significantly, and those resistant to penicillin decreased (Table S2).

Selection of strains for sequencing. To select strains for sequencing that captured the maximum amount of diversity, we reasoned that the antibiotic resistance profiles

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

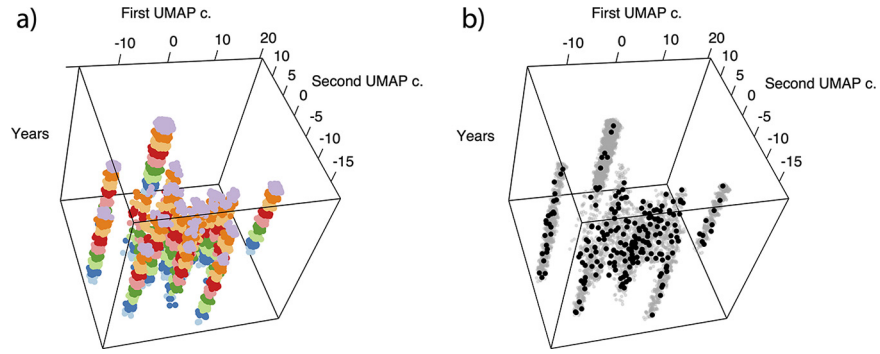


FIG 2 UMAP transformation of the data and integration with sampling year. The *x* and *y* axes correspond to the first two UMAP components and are therefore a function of the original resistance profiles, while the *z* axis corresponds to the year of sampling ($z = 0$ is 2011). (a and b) In panel a the colors indicate the sampling years (2011 to 2019), while in panel b the black dots indicate the isolates selected for genome sequencing.

available for all isolates provides a valid proxy of genomic relatedness/divergence. On this premise, we clustered isolates based on resistance profiles and selected isolates from each cluster, while also maximizing the temporal spread. This aim was achieved through a two-step procedure: the uniform manifold approximation and projection (UMAP) algorithm was used to reduce the dimensionality of the antibiograms, and DBSCAN was used to partition observations into groups of similar profiles.

This procedure resolved the 243 unique resistance profiles of the 7,523 isolates into 24 clusters (Fig. 2), 16 of which contained only MSSA isolates. Isolates were selected for sequencing to include, where possible, at least one isolate per cluster per year of study. This resulted in the selection of 226 samples, encompassing 79 different resistance profiles, including some below 1% frequency (see Table S3 for a detailed description of the resistance profiles and corresponding sequenced isolates). By comparison, random sampling resulted in a significantly smaller number of unique resistance profiles (49 ± 4).

Genomic analysis. The 226 genomes were sequenced and assembled (see Materials and Methods). *In silico* multilocus sequence typing (MLST) identified 8 clonal complexes (CCs) and 49 sequence types (STs), including 16 novel ones STs 6438 to 6452 and ST 6461 that were submitted to PubMLST (28). We identified 87 Spa types and 4 staphylococcal cassette chromosome *mec* (SCC*mec*) types: classical nosocomial SCC*mec* types I ($n = 11$) and II ($n = 6$) represented a minority, while cassette IV, historically community associated ($n = 74$), made up 76% of the 97 MRSA isolates. SCC*mec* V, another community-associated cassette, was detected in only 6 genomes. A detailed description of the isolates and their genomic characteristics is given in Table S4.

The sampling strategy used allowed us to detect rare strains that otherwise would have remained unseen with a sequencing effort based on random sampling. Indeed, we found 37 locally rare STs (i.e., represented by one or two genomes in the entire data set), 16 of which are completely novel, for which we had to define novel STs. The phylogenetic tree of our entire data set shows that we sequenced genomes from 13 rare and phylogenetically divergent lineages, which based on branch length can be clearly seen as not closely related to the common ones (Fig. S2).

Although our sampling strategy intrinsically prevents the use of genomic data in conjunction with information regarding the relative frequencies of the resistance profiles, this approach maximizes phylogenetic breadth and thus allows us to investigate how often virulence and antibiotic resistance genes (ARGs) are lost or gained over evolutionary time. ARGs impose a fitness cost in the absence of the antibiotic, and this may provide strong selective pressure for their rapid loss after treatment (29). We explore these dynamics using the phylogenetic signal (here indicated by D) as previously implemented in R (30) (Table S7). Briefly, this metric provides a means to gauge the concordance between the presence/absence of a given trait with the branching pattern of the underlying tree; a negative value reflects high concordance, while a positive value indicates independence from the phylogenetic topology

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

and, therefore, dispersion of a trait. A trait in this context can be a gene, a phenotype, or any other binary feature (presence/absence) associated with the isolates. As a control, we calculated the same metric for a trait that is, by definition, highly consistent with the tree, i.e., belonging to a clonal complex (CC22).

The D value of all-cause mortality in this analysis is 1.5, indicating a very low correlation of this trait with the phylogenetic tree. This suggests that the outcome of *S. aureus* sepsis could be largely independent from the *S. aureus* lineages and that other factors are likely accountable for the death of the patients (such as starting health condition of the patient, nature of the illness/wound, age). This topic was further explored with the *ad hoc* analysis presented below.

The presence/absence of all virulence factors is consistent with the underlying tree ($D < -1$), suggesting identity by descent to be the dominant mode of evolution, rather than frequent horizontal transfer. Genes linked within a single pathogenicity island (e.g., *splA* and *splB*) show highly consistent results, as expected for linked genes inherited together.

D values for antibiotic resistance genes are also mostly negative, but closer to zero than those for virulence genes, suggesting a faster rate of gains and losses, which is to be expected, as many resistance genes are carried on plasmids. We speculate that larger or positive D values for antibiotic resistance genes reflect variable levels of antibiotic exposure, combined with strong selection pressures for the gain and loss of these genes. Resistance genes may impose a fitness cost in the absence of the antibiotic, and the plasmid carrying these genes can be lost quickly if compensatory mutations are not acquired. Alternatively, resistant and sensitive strains may coexist in the environment.

Rifampicin and ciprofloxacin show extreme values; the former is the only character highly over dispersed on the tree, which is consistent with very frequent losses and gains ($D = 1.86$). In contrast, ciprofloxacin resistance is by far the most stable ($D = -20$), which is consistent with previous findings (31). While both these resistances are due to mutations in housekeeping chromosomal genes, there are important differences between them. Resistance to rifampicin is typically caused by single mutations in the *rpoB* gene (32), while resistance to ciprofloxacin requires multiple mutations in different genes (mainly in genes encoding DNA gyrase and topoisomerase IV, the targets of the antibiotic) (33). Additionally, mutations conferring resistance to rifampicin are associated with a significant fitness cost in the absence of antibiotic, while mutations conferring resistance to ciprofloxacin confer a comparatively mild decrease in fitness (29, 34). Both the fitness cost and the resistance level can be reduced by additional mutations at the same or different loci in the absence of antibiotic pressure (35).

We conclude that resistance to rifampicin is easy to develop but costly to maintain, while resistance to ciprofloxacin is more difficult to acquire (multiple mutations required) but less expensive. Indeed, mathematical modeling has previously shown that even small differences in fitness can translate into very different survival times after removal of selective pressure (e.g., 36). According to the modeling by Nowak, if rifampicin-resistant strains have a 25% decrease in fitness with respect to ciprofloxacin (29, 34), they will be replaced more than four times faster by wild-type strains once antibiotic treatment is removed.

Additionally, we performed a birth-death analysis to explore the relative rates of gains and losses of these two resistance traits compared to the presence/absence of a virulence gene, *lukED* (see Materials and Methods; Fig. 3b). This confirmed that resistance tends to be lost more commonly than gained, while the opposite is true for the more stable virulence gene, *lukED*. Rifampicin resistance is more rapidly lost than ciprofloxacin resistance, which again is consistent with the former conferring a greater fitness cost.

Community- and hospital-acquired infections. *S. aureus* infections are usually defined as hospital or community acquired (HA and CA) based on epidemiological criteria. While most early studies suggested genetic characteristics that may differentiate HA and CA, later comparative studies have challenged this view (14).

According to the CDC epidemiological criteria, out of the 2,807 inpatients positive for *S. aureus* in our data set, 51.7% are classified as CA (while the 2,426 outpatients were clearly all classified as CA). HA isolates were more common in blood (27%) and respiratory

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

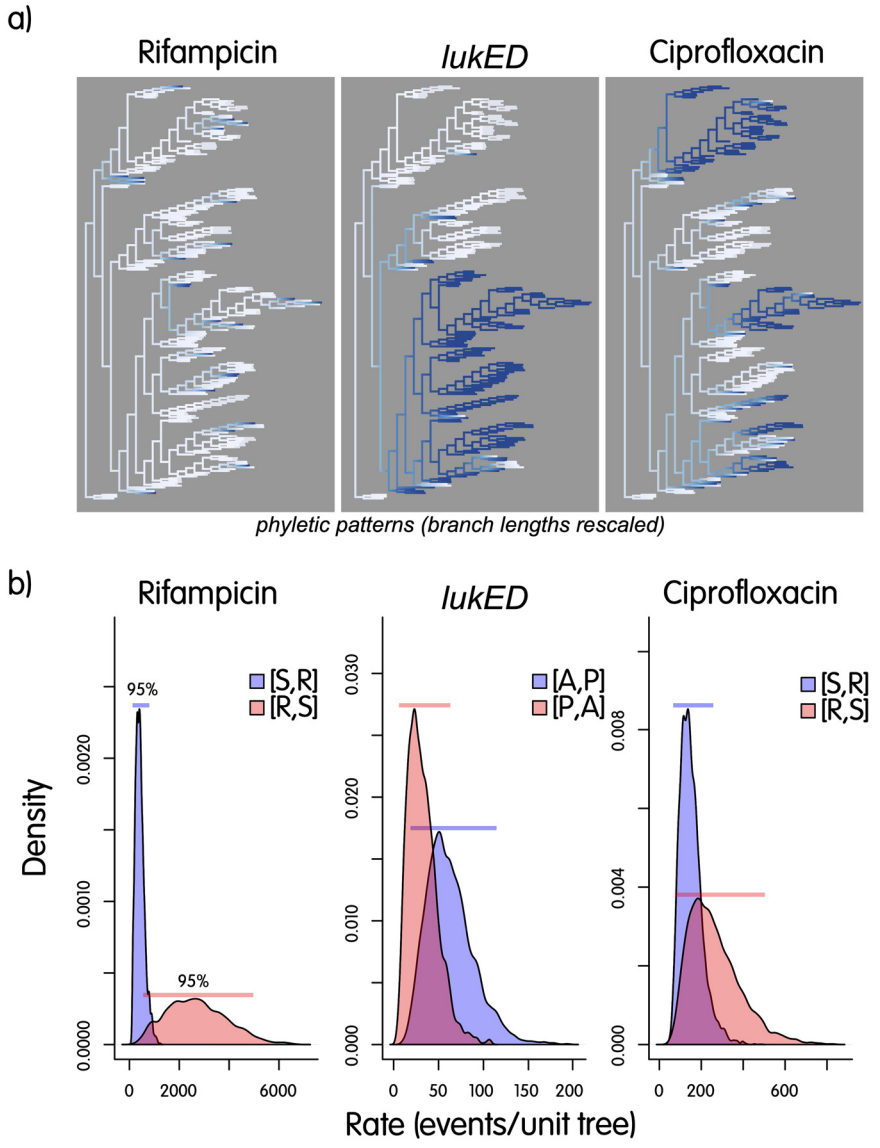


FIG 3 (a) Phyletic patterns with ancestral state reconstruction for two resistance phenotypes and one virulence gene. Rifampicin resistance is highly scattered across the tree and is lost quickly, consistent with a high fitness cost in the absence of the antibiotic. This contrasts with the ciprofloxacin resistance, which is more common and carries a lower fitness cost. The distribution of *LukED* closely matches the phylogeny. (b) Evolutionary rates for the same traits, indicating that reversion rates (from resistant to sensitive) for antibiotics are on average higher than the rate of loss of virulence genes. The trees have branch lengths rescaled for graphical purposes.

tract samples (29%), while CA isolates were more commonly recovered from skin (24%) and soft tissue infections (33%) (Fig. S1). During this period, CA infections increased in yearly numbers (from 39 in 2011 to 182 in 2019), while HA remained stable, leading to an increase of CA from 38% to 54.4% of the total (linear regression $P < 0.01$, beta = 1.8), becoming the primary source of infections in the hospital. By additionally partitioning HA and CA infections into MR and MS, we observed significant changes in the prevalence of CA-MSSA and HA-MRSA in inpatients. The percentage of CA-MSSA increased from 27% to 42.1%, while HA-MRSA declined from 30.2% to 16% (see Table S5 for raw numbers and regression values). Comparison of resistance patterns of HA and CA samples indicates that the former tend to have more resistances than the latter, with the largest differences being for methicillin (37.89% and 27.05%, respectively) and ciprofloxacin (39.29% and 27.8%, respectively) (Table S6). However, since MRSA and MSSA are nonhomogeneously distributed in the two groups, we explored how the number of different antibiotic resistances is

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

TABLE 1 Significant ($P < 0.05$) predictors of worse outcome when patients older than 60 years are removed from the analysis and the predictors indicated in Materials and Methods[a]

Predictor	Estimate	SE	Z value	Significance
Intercept	−4.77	0.85	−5.59	2.22E-08
Methicillin	1.62	0.70	2.32	2.02E-02
Fosfomycin	1.90	0.84	2.25	2.43E-02
Ciprofloxacin	−1.73	0.76	−2.28	2.25E-02
Blood	2.69	0.68	3.97	7.09E-05
Respiratory tract	3.07	0.68	4.51	6.42E-06
Urine	2.41	0.98	2.45	1.43E-02

[a]This filtered data set highlights additional predictors that may be important when infections involve younger patients.

distributed across the four groups, CA-MSSA, HA-MSSA, CA-MRSA, and HA-MRSA (Fig. S3), confirming that the main differences reflect the MRSA/MSSA partition. As expected, and previously reported (20), isolates sensitive to all or most antibiotics are more often MSSA, while multidrug-resistant isolates tend to be MRSA (Table S2).

Considering the 226 sequenced genomes of sepsis isolates, 35% were HA and 65% CA. We checked for the presence of markers known to be associated with CA, such as the SCC*mec* type and the presence of the PVL gene. However, no trait was found to be significantly associated with either HA or CA strains, confirming previous findings that showed that the two epidemiological classes are not distinct at the genomic level.

Mortality rate. Through the studied period, *S. aureus* infections can be connected to at least 435 all-cause deaths (8.3%; mean age, 71.8 years), 418 of which were inpatients. We used generalized linear models (GLM) to identify predictive factors of mortality. At first, we used the whole metadata associated with samples and patients, including antibiotic resistances, site of infection/sampling, and age as predictors (Table S7). This resulting model highlighted that the most significant predictor associated with fatal outcome is age, together with two source materials, blood and respiratory tract, which are associated with invasive infections and higher mortality rates (42.5% and 30%, respectively). Two antibiotic resistances, to methicillin and fosfomycin, are also positively significantly associated with mortality. Resistance to ciprofloxacin is associated with increased survivability. Mortality rate due to MR infections was higher that due to MS infections (15% to 5.9%), a difference that may be associated with the higher average number of resistances in MR isolates (Fig. S3). Additionally, the average length of stay (LOS) for MR infections is significantly longer than for MS infected inpatients (31.2 and 18.21, respectively). Older patients usually have weakened immune systems, which makes any infection potentially lethal; this might mask the smaller contribution of other predictors. For this reason, we also built a model removing isolates from patients more than 60 years old (Table 1). This model confirms the one above and additionally identifies urine infections as leading to an increased risk of death.

Lastly, the mortality of CA and HA strains was 13.2% and 16.7%, respectively. For the latter, we report a significant mortality reduction ($P = 0.03$, beta = –0.7), from 20.5% in 2011 to 13.1% in 2019, while casualties associated with CA strains were relatively stable (standard deviation = 2.2%). By exploiting the additional MR/MS partitioning, we detected a significant ($P = 0.02$, beta = –0.4) decline over time in the number of HA-MR casualties.

When focusing on the isolates selected for genomic sequencing, the overall mortality rate was 29.7% ($n = 53$). The attributable mortality rate for MSSA ($n = 29$) and MRSA ($n = 24$) blood infections was 28.7% and 31.1%, respectively. None of the CCs were found to be significantly associated with an increased risk of death, even though we found CC22, CC8, and CC30 to have a rate of mortality above the mean, standing at 40.9%, 39.3%, and 33.3%, respectively.

This result agrees with that described above, with the distribution of mortality on the phylogenetic tree being best explained by a model that does not account for phylogenetic structure (Table S8). The two results indicate that mortality is more dependent on characteristics of the host/infection site than on genomic properties of the specific *S. aureus* strain. To explore this issue in more detail, we applied phylogenetic regression to mortality using

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

as predictors both hospital and genomics data. In contrast to the generalized models described above, in this case we integrated the structure of the phylogenetic tree to take into account the nonindependence of the observations. The model identifies several predictors that are significantly associated with the outcome; the most important is age, followed by contracting the infection once inside the hospital (CDC-HA). We also identified some antibiotic resistances and virulence factors that are associated with the outcome (Table S9).

It must be noted that our analysis presents limitations. First, we were not able to evaluate the association between genotype and important clinical outcomes outside mortality (e.g., readmission, prolonged bacteremia), as we had no access to these metadata. Additionally, our numbers are limited; this could explain discrepancies with previous studies suggesting a significant relationship of genotype and mortality (e.g., 37, 38). It should be considered, however, that studies such as those cited above use multivariate regression without taking into account the fact that genomes are not independent observations. This makes all studies where this is not considered potentially flawed. The major technical novelty of our paper is indeed that we account for phylogenetic distances in the regression analysis, which enables us to remove the confounding effect of an uneven sampling of the different CCs. We posit that future studies, with higher numbers, following our unbiased sampling approach and considering genomes as not independent observations could be the key to better understanding whether and how much genotypes and outcomes can be associated.

Clonal complexes. Almost half ($n = 109$) of the 226 sequenced strains belong to CC5, CC8, and CC22, which are known to be globally abundant, and we therefore inspected them in more detail, together with CC30, a complex characterized by a high all-cause mortality rate (33%). We used core single-nucleotide polymorphism (SNP)-based phylogenetic analysis and profiling of virulence and antibiotic resistance genes, and the results are summarized in Fig. 4 (see also Fig. S4 and S5 for trees without and with branch lengths). Figure S6 shows the heatmaps of the SNP distances between pairs of isolates, highlighting the presence of clusters of genomes with high levels of similarity.

The phylogenetic analyses, which include genomes from public database of the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) (39) (see Table S10), provide the means to determine whether monophyletic clades of closely related genomes can be found in our data set. Indeed, such clades can be spotted in the phylogenetic tree of CC8 and CC22 and encompass isolates dating across the 9 years of our data set, indicating that these strains are endemic to the hospital. The phylogenetic tree of CC5 is also characterized by several small monophyletic clades of highly related genomes, but these are more limited in time, indicating that there were independent introductions in the hospital with consequent outbreaks that were quickly eradicated. The phylogenetic tree of CC30 reveals interesting differences with respect to the other three CCs: genomes are more diverse and spread on the tree with no monophyletic clades, suggesting sporadic but deadly infections, with no strain apparently persisting in the hospital. This lower persistence of CC30 could be explained by an average lower resistance to antibiotics of these isolates (Fig. 4), as also previously reported by Aanensen and colleagues (20).

Besides being resistant to penicillin, and less commonly to tetracycline, CC5, CC8, and CC22 are also usually resistant to ciprofloxacin and oxacillin/cefoxitin, these two being rare in other CCs, and to a lesser degree also gentamicin and erythromycin, which are shared with several other CCs (Fig. S5a). Isolates belonging to CC30 show much lower levels of antibiotic resistance. CC5, CC22, and CC30 possess the enterotoxin gene cluster (*ecg*, light blue in Fig. S5b), genes that are only sparsely present in other CCs, except for CC45. They also code for hemolysins, chemotaxis inhibitory protein (Chp), and staphylococcal complement inhibitor (Scn). CC5 and CC8 show the sporadic presence of several additional toxins (light purple in Fig. S5b) that are not found in any of the other CCs, including some that were annotated as superantigens (*ser* and *sel*) and share the presence of a serine protease operon (represented here by the *splABE* genes) together with the pore-forming toxin gene *lukED*. A portion of CC30 isolates also harbor *tst*, an association consistent with previous genomics studies (20). This gene has been implicated in toxic shock syndrome in susceptible hosts (40). This serious and often fatal condition is triggered by superantigen toxins causing a

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

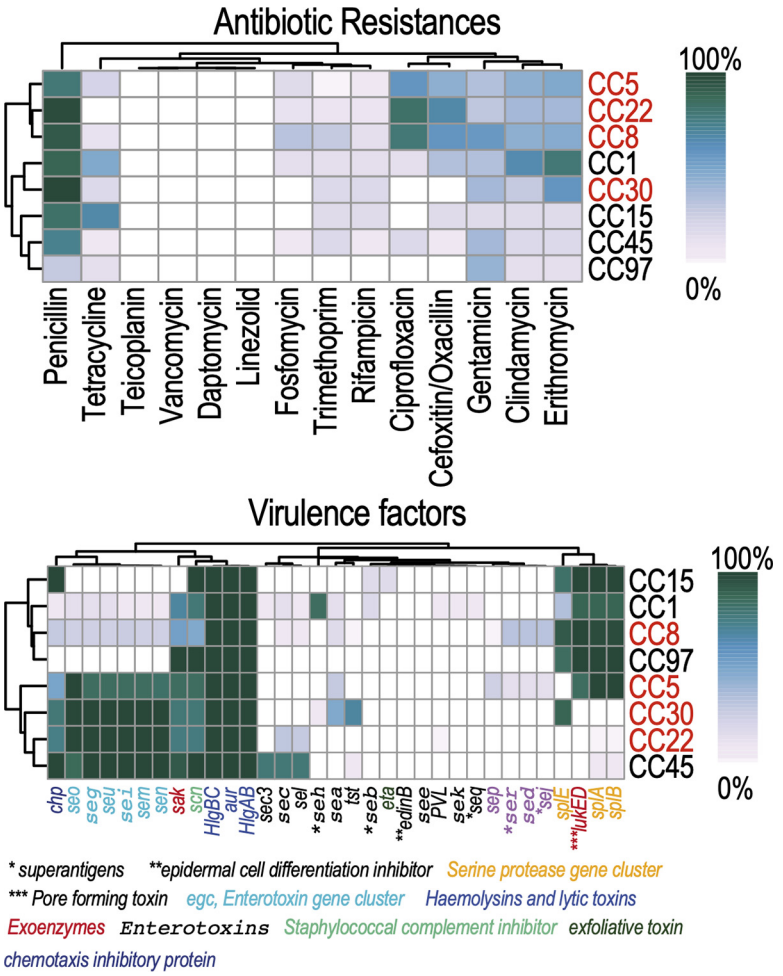


FIG 4 (a and b) Summary of antibiotic resistances (a) and virulence factors (b) in a clonal complex perspective. Cells in the heatmap correspond to the frequency of strains possessing a resistance or virulence factor in a given CC. Functional annotation of virulence genes is also indicated.

cytokine avalanche by T cells manifesting as fever, rash, shock, and rapid and multiple organ failure (41).

Conclusions. In this work we present a strategy for maximizing the diversity of clinical strains that can be characterized through genomics in real-world settings with finite resources. The strategy exploits antibiograms as a proxy of genetic diversity to select isolates for sequencing, as these data are usually generally available for all isolates in clinical settings. We validated our strategy on an *S. aureus* collection, demonstrating how it can reveal rare lineages by reducing the redundancy that characterizes bacterial populations, especially in hospitals, where outbreaks of identical clones are common.

This approach provides a means to avoid the oversampling of very similar isolates from common STs, allowing us to obtain genomes from rarer strains, which would be less likely to be detected through a random population snapshot. We posit that obtaining the genomes of rare strains provides a more complete genomic and phylogenetic picture of the epidemiology of the hospital. This, in turn, grants advantages such as additional power to forecasting strategies, which can benefit from the early detection of rare strains that could be rising in prevalence and cause outbreaks in the future.

MATERIALS AND METHODS

Determination of antimicrobial susceptibility by MIC. *Staphylococcus aureus* isolates from patients admitted to Fondazione IRCCS Policlinico San Matteo Hospital (Pavia, Italy) are routinely investigated by the automated system Phoenix (Becton, Dickinson, Franklin Lakes, NJ, USA) using PMIC-88 panels for the detection of antimicrobial resistances and determination of MIC values. MICs are interpreted with the EUCAST

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

breakpoints (42) to define which strains are susceptible, intermediate, or resistant to each antibiotic. This work collects and exploits available metadata collected in the hospital for all the isolates of *S. aureus* and the respective patients, from 1 January 2011 to 31 December 2019. For all isolates, resistance profiles were collected with respect to 15 antibiotic compounds of clinical relevance: penicillin, teicoplanin, clindamycin, daptomycin, cotrimoxazole, oxacillin, cefoxitin, erythromycin, linezolid, rifampicin, vancomycin, gentamicin, fosfomycin, ciprofloxacin, and tetracycline. Oxacillin and cefoxitin results are used to define methicillin resistance. For all patients, the following data were retrieved and used: isolation material, patient age, admission date, outcome. It should be noted that due to technical issues occurring within the hospital servers/digital infrastructure before the beginning of this project, for the years 2011 and 2015 an unspecified number of samples were lost.

CDC classification. Community- and health care-acquired infections are currently defined based on criteria provided by the Centers for Disease Control and Prevention (CDC) (11). An isolate is considered CA if the patient fulfills the following criteria: (i) testing positive for an infection within 48 h of admission, (ii) not having undergone surgery in the 48 h since admission, (iii) not residing in a long-term-care facility before admission, and (iv) not having undergone hemodialysis/peritoneal dialysis in the past year or during the current admission. Instead, it is considered HA if a subsequently positive wound culture was taken after 48 h from the hospital admission or if one or more of the CA criteria are not met.

Antibiogram-based isolate selection for genome sequencing. All antibiograms with missing MICs, MICs classified as intermediate, or ambiguous MIC values were removed from the data set. The remaining MIC values were coded into binary variables R (resistant) and S (susceptible). Data set dimensionality was first reduced with a technique called uniform manifold approximation and projection (UMAP) using the R package uwot (43). Subsequently, DBSCAN enabled the identification of compact clusters in UMAP coordinates using the R package DBSCAN (44). This analysis informed the selection of 226 *S. aureus* isolates that capture all the cluster diversity.

DNA extraction and sequencing. DNA was extracted from 226 blood isolates using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany), following the manufacturer's recommendation and adding an initial step of incubation of the bacterial pellet with lysozyme (20 mg/mL) for 2 h; gel electrophoresis (Tris-acetate-EDTA [TAE], 1.5% agarose) was used to check DNA quantity and quality. Whole-genome sequencing was performed in paired-end mode (2 × 150 bp) on a NovaSeq machine (Illumina). Reads were quality checked with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and assembled with SPAdes v.3.14.1 (45), and quality was assessed with assembly-stats (github.com/sanger-pathogens/assembly-stats). In order to evaluate the completeness of our genome assemblies, we performed a previously described check (46). Briefly, we searched for the presence of core genome MLST (cgMLST) genes (determined in reference 47) using BLAST and verified that all genomes contained more than 95% of them (Table S4).

***In silico* typing and annotation.** *In silico* multilocus sequence typing (MLST) was performed by comparing our genomes with the *S. aureus* MLST database (PubMLST) (28), using an in-house python script. Typing of the staphylococcal cassette chromosome *mec* (SCCmec) element was done with the web implementation of SCCmecFinder (48) with the options min_coverage = 80 and percentage identity = 90. Spa types were assigned with Ridom SeqSphere+ (48, 49). Presence/absence profiles of 32 virulence factors, chosen because they are commonly associated with bloodstream infections and involved in different clinical syndromes, were identified with VirulenceFinder v.2.0 (50), with the settings min_coverage = 80 and percentage identity = 90.

Statistical analyses. As recommended by the Clinical and Laboratory Standards Institute (CLSI) guidelines M39-A4 (51) and many recent papers (e.g., 52), statistical analyses were performed on the first isolate per patient only. R (v3.2.4) was used for all statistical analyses. Antibiotic resistance data were analyzed using the chisq.test function, while population proportions were assessed with the proportion test (Z test) through the function prop.test. The lm function was used for linear regression using time as a predictor; a significant regression coefficient (beta) for time was taken as an indication of a temporal trend in the data. For generalized models, we used the glm function and a combination of numerical and categorical variables as detailed in Results and Discussion. For models with outcome as the dependent variable considering all isolates, we included antibiotic resistance profiles, age, material, and CDC classification as predictors.

Phylogenetic analyses. A core SNP-based phylogeny and an SNP distance matrix were generated for each of the selected clonal complexes (CC) using the P-DOR pipeline (53). Briefly, the pipeline contextualizes the query genomes with background data sets, based on genetic similarity. Core SNPs are then called, aligning the genomes to a reference. The following reference genomes were used: GenBank accession no. NC_017763.1 (CC22), BX571856.1 (CC30), CP026068.1 (CC8), and CP021105.1 (CC5); the BV-BRC (39) genomes used for contextualization are available in Table S10. The resulting core SNP multialignment was fed to IQ-TREE (54) to obtain phylogenetic trees for each CC. P-DOR was also used to generate the phylogenetic tree of the entire novel local data set (GenBank accession no. BX571858), skipping the step of retrieval of database genomes from BV-BRC (39).

We estimated the transition rates of a simple birth-and-death model based on the phyletic pattern of interest, whose states are given by sensitive (or absent, S or A) and resistant (or present, R or P) and are on the core SNP-based phylogenetic tree (Fig. 3b). The mcmcMk function from the R package phytools (55) applied here uses the tree and the phyletic pattern to estimate the transition rates among the two states and was run for 50,000 generations; the function contMap from the same package was used to reconstruct and map ancestral states on the phylogenetic tree; the phylogenetic signal was estimated using the function phylo.D from the package caper (https://rdrr.io/cran/caper/) and 1,000 permutations. Linear phylogenetic regression was performed with phylolm (56) to explicitly consider the nonindependence of observations through the relationships summarized by the phylogenetic tree. Before using presence/absence profiles of virulence genes as predictors in phylogenetic regression, we removed colinear profiles that are known to affect parameter estimation in a negative way. This was done by identifying genes with profiles 90% identical or

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

more and then using one of the profiles as a representative in regression, resulting in 18 virulence gene patterns over 26 genes.

Data availability. The genome sequences are available at NCBI under BioProject accession number PRJNA797572. Accession numbers of BV-BRC genomes used for contextualization are listed in Table S10.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.01 MB.

SUPPLEMENTAL FILE 2, PDF file, 0.4 MB.

SUPPLEMENTAL FILE 3, PDF file, 0.02 MB.

SUPPLEMENTAL FILE 4, PDF file, 0.1 MB.

SUPPLEMENTAL FILE 5, JPG file, 0.2 MB.

SUPPLEMENTAL FILE 6, PNG file, 0.3 MB.

SUPPLEMENTAL FILE 7, DOCX file, 0.02 MB.

SUPPLEMENTAL FILE 8, XLSX file, 0.1 MB.

ACKNOWLEDGMENT

The project was funded by Ricerca Corrente grant 924-rcr2018i-34 (Italian Ministry of Health) to S.G.

REFERENCES

1. Balasubramanian D, Harper L, Shopsin B, Torres VJ. 2017. Staphylococcus aureus pathogenesis in diverse host environments. Pathog Dis 75:ftx005. https://doi.org/10.1093/femspd/ftx005.
2. Kluytmans JAJW, Kluytmans JAJ, Wertheim HFL. 2005. Nasal carriage of Staphylococcus aureus and prevention of nosocomial infections. Infection 33:3–8. https://doi.org/10.1007/s15010-005-4012-9.
3. Tong SYC, Davis JS, Eichenberger E, Holland TL, Fowler VG Jr. 2015. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev 28:603–661. https://doi.org/10.1128/CMR.00134-14.
4. Haaber J, Penadés JR, Ingmer H. 2017. Transfer of antibiotic resistance in Staphylococcus aureus. Trends Microbiol 25:893–905. https://doi.org/10.1016/j.tim.2017.05.011.
5. Jani M, Sengupta S, Hu K, Azad RK. 2017. Deciphering pathogenicity and antibiotic resistance islands in methicillin-resistant Staphylococcus aureus genomes. Open Biol 7:170094. https://doi.org/10.1098/rsob.170094.
6. Tsuji BT, Rybak MJ, Cheung CM, Amjad M, Kaatz GW. 2007. Community- and health care-associated methicillin-resistant Staphylococcus aureus: a comparison of molecular epidemiology and antimicrobial activities of various agents. Diagn Microbiol Infect Dis 58:41–47. https://doi.org/10.1016/j.diagmicrobio.2006.10.021.
7. Baggett HC, Hennessy TW, Rudolph K, Bruden D, Reasonover A, Parkinson A, Sparks R, Donlan RM, Martinez P, Mongkolrattanothai K, Butler JC. 2004. Community-onset methicillin-resistant Staphylococcus aureus associated with antibiotic use and the cytotoxin Panton-Valentine leukocidin during a furunculosis outbreak in rural Alaska. J Infect Dis 189:1565–1573. https://doi.org/10.1086/383247.
8. Berglund C, Ito T, Ikeda M, Ma XX, Söderquist B, Hiramatsu K. 2008. Novel type of staphylococcal cassette chromosome mec in a methicillin-resistant Staphylococcus aureus strain isolated in Sweden. Antimicrob Agents Chemother 52:3512–3516. https://doi.org/10.1128/AAC.00087-08.
9. David MZ, Daum RS. 2010. Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev 23:616–687. https://doi.org/10.1128/CMR.00081-09.
10. Miller LG, Perdreau-Remington F, Bayer AS, Diep B, Tan N, Bharadwa K, Tsui J, Perlroth J, Shay A, Tagudar G, Ibebuogu U, Spellberg B. 2007. Clinical and epidemiologic characteristics cannot distinguish community-associated methicillin-resistant Staphylococcus aureus infection from methicillin-susceptible S. aureus infection: a prospective investigation. Clin Infect Dis 44:471–482. https://doi.org/10.1086/511033.
11. David MZ, Glikman D, Crawford SE, Peng J, King KJ, Hostetler MA, Boyle-Vavra S, Daum RS. 2008. What is community-associated Methicillin-resistant Staphylococcus aureus? J Infect Dis 197:1235–1243. https://doi.org/10.1086/533502.
12. McCarthy NL, Sullivan PS, Gaynes R, Rimland D. 2010. Health care-associated and community-associated methicillin-resistant Staphylococcus aureus infections: a comparison of definitions. Am J Infect Control 38:600–606. https://doi.org/10.1016/j.ajic.2010.01.006.
13. Mediavilla JR, Chen L, Mathema B, Kreiswirth BN. 2012. Global epidemiology of community-associated methicillin resistant Staphylococcus aureus (CA-MRSA). Curr Opin Microbiol 15:588–595. https://doi.org/10.1016/j.mib.2012.08.003.
14. Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt BG, Moore CE, Day NPJ. 2003. How clonal is Staphylococcus aureus? J Bacteriol 185:3307–3316. https://doi.org/10.1128/JB.185.11.3307-3316.2003.
15. Calfee DP. 2017. Trends in community versus health care-acquired methicillin-resistant Staphylococcus aureus infections. Curr Infect Dis Rep 19:48. https://doi.org/10.1007/s11908-017-0605-6.
16. Pang R, Wu S, Zhang F, Huang J, Wu H, Zhang J, Li Y, Ding Y, Zhang J, Chen M, Wei X, Zhang Y, Gu Q, Zhou Z, Liang B, Li W, Wu Q. 2020. The genomic context for the evolution and transmission of community-associated ST59 through the food chain. Front Microbiol 11:422. https://doi.org/10.3389/fmicb.2020.00422.
17. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science 327:469–474. https://doi.org/10.1126/science.1182395.
18. Monaco M, Pimentel de Araujo F, Cruciani M, Coccia EM, Pantosti A. 2017. Worldwide Epidemiology and Antibiotic Resistance of Staphylococcus aureus. Curr Top Microbiol Immunol 409:21–56. https://doi.org/10.1007/82_2016_3.
19. Stefani S, Chung DR, Lindsay JA, Friedrich AW, Kearns AM, Westh H, Mackenzie FM. 2012. Meticillin-resistant Staphylococcus aureus (MRSA): global epidemiology and harmonisation of typing methods. Int J Antimicrob Agents 39:273–282. https://doi.org/10.1016/j.ijantimicag.2011.09.030.
20. Aanensen DM, Feil EJ, Holden MTG, Dordel J, Yeats CA, Fedosejev A, Goater R, Castillo-Ramírez S, Corander J, Colijn C, Chlebowicz MA, Schouls L, Heck M, Pluister G, Ruimy R, Kahlmeter G, Åhman J, Matuschek E, Friedrich AW, Parkhill J, Bentley SD, Spratt BG, Grundmann H, European SRL Working Group. 2016. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive Staphylococcus aureus in Europe. mBio 7:e00444-16. https://doi.org/10.1128/mBio.00444-16.
21. Matuszewska M, Murray GGR, Harrison EM, Holmes MA, Weinert LA. 2020. The Evolutionary Genomics of Host Specificity in Staphylococcus aureus. Trends Microbiol 28:465–477. https://doi.org/10.1016/j.tim.2019.12.007.
22. Merla C, Kuka A, Petazzoni G, Postiglione U, Zatelli M, Gaiarsa S, Mariani B, Corbella M, Marone P, Sassera D, Cambieri P. 2022. Livestock-associated methicillin-resistant Staphylococcus aureus in inpatients: a snapshot from an Italian hospital. J Glob Antimicrob Resist 30:10–15. https://doi.org/10.1016/j.jgar.2022.05.013.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.

23. Abdelbary MMH, Feil EJ, Senn L, Petignat C, Prod'hom G, Schrenzel J, François P, Werner G, Layer F, Strommenger B, Pantosti A, Monaco M, Denis O, Deplano A, Grundmann H, Blanc DS. 2020. Phylogeographical analysis reveals the historic origin, emergence, and evolutionary dynamics of methicillin-resistant ST228. Front Microbiol 11:2063. https://doi.org/10.3389/fmicb.2020.02063.
24. Reuter S, Török ME, Holden MTG, Reynolds R, Raven KE, Blane B, Donker T, Bentley SD, Aanensen DM, Grundmann H, Feil EJ, Spratt BG, Parkhill J, Peacock SJ. 2016. Building a genomic framework for prospective MRSA surveillance in the United Kingdom and the Republic of Ireland. Genome Res 26:263–270. https://doi.org/10.1101/gr.196709.115.
25. Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J, Peacock SJ. 2013. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis 13:130–136. https://doi.org/10.1016/S1473-3099(12)70268-2.
26. Stein M, Navon-Venezia S, Chmelnitsky I, Kohelet D, Schwartz O, Agmon O, Somekh E. 2006. An outbreak of new, nonmultidrug-resistant, methicillin-resistant Staphylococcus aureus strain (SCCmec type iiia variant-1) in the neonatal intensive care unit transmitted by a staff member. Pediatr Infect Dis J 25:557–559. https://doi.org/10.1097/01.inf.0000219407.31195.44.
27. Manara S, Pasolli E, Dolce D, Ravenni N, Campana S, Armanini F, Asnicar F, Mengoni A, Galli L, Montagnani C, Venturini E, Rota-Stabelli O, Grandi G, Taccetti G, Segata N. 2018. Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of Staphylococcus aureus strains in a paediatric hospital. Genome Med 10:82. https://doi.org/10.1186/s13073-018-0593-7.
28. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.
29. Melnyk AH, Wong A, Kassen R. 2015. The fitness costs of antibiotic resistance mutations. Evol Appl 8:273–283. https://doi.org/10.1111/eva.12196.
30. Fritz SA, Purvis A. 2010. Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. Conservation Biology 24:1042–1051. https://doi.org/10.1111/j.1523-1739.2010.01455.x.
31. Jones ME, Boenink NM, Verhoef J, Köhrer K, Schmitz F-J. 2000. Multiple mutations conferring ciprofloxacin resistance in Staphylococcus aureus demonstrate long-term stability in an antibiotic-free environment. J Antimicrob Chemother 45:353–356. https://doi.org/10.1093/jac/45.3.353.
32. Goldstein BP. 2014. Resistance to rifampicin: a review. J Antibiot (Tokyo) 67:625–630. https://doi.org/10.1038/ja.2014.107.
33. Jacoby GA. 2005. Mechanisms of resistance to quinolones. Clin Infect Dis 41:S120–S126. https://doi.org/10.1086/428052.
34. Dunai A, Spohn R, Farkas Z, Lázár V, Györkei Á, Apjok G, Boross G, Szappanos B, Grézal G, Faragó A, Bodai L, Papp B, Pál C. 2019. Rapid decline of bacterial drug-resistance in an antibiotic-free environment through phenotypic reversion. Elife 8:e47088. https://doi.org/10.7554/eLife.47088.
35. O'Neill AJ, Huovinen T, Fishwick CWG, Chopra I. 2006. Molecular genetic and structural modeling studies of Staphylococcus aureus RNA polymerase and the fitness of rifampin resistance genotypes in relation to clinical prevalence. Antimicrob Agents Chemother 50:298–309. https://doi.org/10.1128/AAC.50.1.298-309.2006.
36. Nowak MA. 2006. Evolutionary dynamics: exploring the equations of life. Harvard University Press, Cambridge, MA.
37. Kempker RR, Farley MM, Ladson JL, Satola S, Ray SM. 2010. Association of methicillin-resistant Staphylococcus aureus (MRSA) USA300 genotype with mortality in MRSA bacteremia. J Infect 61:372–381. https://doi.org/10.1016/j.jinf.2010.09.021.
38. Souli M, Ruffin F, Choi S-H, Park LP, Gao S, Lent NC, Sharma-Kuinkel BK, Thaden JT, Maskarinec SA, Wanda L, Hill-Rorie J, Warren B, Hansen B, Fowler VG. 2019. Changing characteristics of Staphylococcus aureus bacteremia: results from a 21-year, prospective, longitudinal study. Clin Infect Dis 69:1868–1877. https://doi.org/10.1093/cid/ciz112.
39. Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, Davis JJ, Dempsey DM, Dickerman A, Dietrich EM, Kenyon RW, Kuscuoglu M, Lefkowitz EJ, Lu J, Machi D, Macken C, Mao C, Niewiadomska A, Nguyen M, Olsen GJ, Overbeek JC, Parrello B, Parrello V, Porter JS, Pusch GD, Shukla M, Singh I, Stewart L, Tan G, Thomas C, VanOeffelen M, Vonstein V, Wallace ZS, Warren AS, Wattam AR, Xia F, Yoo H, Zhang Y, Zmasek CM, Scheuermann RH, Stevens RL. 2023. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res 51:D678–D689. https://doi.org/10.1093/nar/gkac1003.
40. Spaulding AR, Salgado-Pabón W, Kohler PL, Horswill AR, Leung DYM, Schlievert PM. 2013. Staphylococcal and streptococcal superantigen exotoxins. Clin Microbiol Rev 26:422–447. https://doi.org/10.1128/CMR.00104-12.
41. Silversides JA, Lappin E, Ferguson AJ. 2010. Staphylococcal toxic shock syndrome: mechanisms and management. Curr Infect Dis Rep 12:392–400. https://doi.org/10.1007/s11908-010-0119-y.
42. European Committee on Antimicrobial Susceptibility Testing (EUCAST). 2019. Breakpoint tables for interpretation of MICs and zone diameters, version 9.0. https://www.eucast.org/ast_of_bacteria/previous_versions_of_documents.
43. McInnes L, Healy J, Saul N, Großberger L. 2018. UMAP: uniform manifold approximation and projection. JOSS 3:861. https://doi.org/10.21105/joss.00861.
44. Hahsler M, Piekenbrock M, Doran D. 2019. DBSCAN: fast density-based clustering with R. J Stat Softw 91:1–30. https://doi.org/10.18637/jss.v091.i01.
45. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
46. Effelsberg N, Stegger M, Peitzmann L, Altinok O, Coombs GW, Pichon B, Kearns A, Randad PR, Heaney CD, Bletz S, Schaumburg F, Mellmann A. 2020. Global epidemiology and evolutionary history of Staphylococcus aureus ST45. J Clin Microbiol 59:e02198-20. https://doi.org/10.1128/JCM.02198-20.
47. Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A. 2014. Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. J Clin Microbiol 52:2365–2370. https://doi.org/10.1128/JCM.00262-14.
48. Kaya H, Hasman H, Larsen J, Stegger M, Johannesen TB, Allesøe RL, Lemvigh CK, Aarestrup FM, Lund O, Larsen AR. 2018. SCCmecFinder, a web-based tool for typing of staphylococcal cassette chromosome mec in Staphylococcus aureus using whole-genome sequence data. mSphere 3:e00612-17. https://doi.org/10.1128/mSphere.00612-17.
49. Jünemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, Mellmann A, Goesmann A, von Haeseler A, Stoye J, Harmsen D. 2013. Updating benchtop sequencing performance comparison. Nat Biotechnol 31:294–296. https://doi.org/10.1038/nbt.2522.
50. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510. https://doi.org/10.1128/JCM.03617-13.
51. Clinical and Laboratory Standards Institute (CLSI). 2014. Analysis and presentation of cumulative antimicrobial susceptibility test data, 4th ed. CLSI, Wayne, PA.
52. Kim S, Yoo SJ, Chang J. 2020. Importance of susceptibility rate of "the first" isolate: evidence of real-world data. Medicina 56:507. https://doi.org/10.3390/medicina56100507.
53. Batisti Biffignandi G, Bellinzona G, Petazzoni G, Sassera D, Zuccotti GV, Bandi C, Baldanti F, Comandatore F, Gaiarsa S. 2023. P-DOR, an easy-to-use pipeline to reconstruct outbreaks using pathogen genomics. BioRxiv https://doi.org/10.1101/2023.05.30.542810.
54. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. https://doi.org/10.1093/molbev/msaa015.
55. Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.
56. Ho L Si T, Ané C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. Syst Biol 63:397–408. https://doi.org/10.1093/sysbio/syu005.

Downloaded from https://journals.asm.org/journal/spectrum on 23 January 2026 by 2.37.205.135.