CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

LOW-COST RAPID DIAGNOSTICS AND TYPING FOR CLINICAL MICROBIOLOGY USING FT-IR

by

Rita Fonseca da Silva Azevedo Martins

September 2025

CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

LOW-COST RAPID DIAGNOSTICS AND TYPING FOR CLINICAL MICROBIOLOGY USING FT-IR

Thesis presented to *Escola Superior de Biotecnologia* of the *Universidade Católica Portuguesa* to fulfill the requirements of Master of Science degree in Microbiologia Aplicada

by

Rita Fonseca da Silva Azevedo Martins

Supervisor: Dra. Ângela Novais

Tutor: Dra. Clara Sousa

September 2025

RESUMO

As infeções associadas aos cuidados de saúde e a resistência antimicrobiana representam desafios urgentes para a saúde pública global, afetando desfechos clínicos, recursos hospitalares e controlo infecioso. Contudo, as ferramentas de diagnóstico atualmente disponíveis apresentam limitações em termos de rapidez, precisão e simplicidade, aspetos críticos para a deteção precoce de patogénios, a tomada de decisões terapêuticas eficazes, a prevenção de infeções e a contenção da sua disseminação em contextos clínicos.

Esta tese centrou-se em desafios diagnósticos específicos com o objetivo de melhorar a deteção de patogénios e de mecanismos de resistência antimicrobiana de forma precisa, rápida e simples, bem como a expandir as aplicações da espectroscopia de infravermelho com transformada de Fourier (FT-IR) na deteção de espécies do complexo *Enterobacter cloacae*, *Klebsiella pneumoniae* multirresistente e *K. pneumoniae* hipervirulenta. Para concretizar este objetivo, i) foi utilizado um caso clínico real para identificar limitações associadas aos métodos moleculares de rotina na deteção de mecanismos de resistência raros (produção de IMP-22 em *K. pneumoniae*); ii) avaliou-se o potencial da espectroscopia FT-IR para diferenciar e identificar espécies clinicamente relevantes do complexo *E. cloacae*, em comparação com métodos de referência; iii) foi expandida a aplicação da FT-IR para a deteção de *K. pneumoniae* hipervirulenta. Métodos fenotípicos, genotípicos e moleculares foram combinados com a espectroscopia FT-IR e análise de dados multivariada, aplicada à interpretação de padrões espectrais e à construção de modelos de classificação.

Este trabalho reforçou a necessidade de integrar abordagens fenotípicas e moleculares direcionadas para uma vigilância abrangente das carbapenemases. Demonstrou, adicionalmente, a eficácia da FT-IR para a diferenciação a nível de espécie e subespécie, destacando o seu papel como ferramenta diagnóstica rápida e fiável, essencial para o tratamento atempado, a contenção da disseminação de patogénios multirresistentes e a melhoria dos cuidados prestados aos doentes.

PALAVRAS-CHAVE: Diagnóstico, espetroscopia FT-IR, Resistência antimicrobiana, *Klebsiella pneumoniae*, Complexo *Enterobacter cloacae*

ABSTRACT

Healthcare-associated infections (HAIs) and antimicrobial resistance (AMR) represent urgent challenges for global public health, with impactful consequences in patient outcomes and hospital resources. However, current diagnostic tools present limitations in terms of speed, accuracy and simplicity, critical for early pathogen detection, effective treatment decisions, infection prevention and dissemination in clinical settings.

This thesis focused on specific diagnostic challenges to improve accurate, quick and simple pathogen detection and antimicrobial resistance mechanisms and extend Fourier transform infrared (FT-IR) spectroscopy applications for the detection of clinically relevant *Ec*SC, MDR-*Kp* and Hv-*Kp*. To accomplish this task, we i) used a real clinical case to identify limitations associated with routine molecular methods for detection of rare mechanisms of resistance (production of IMP-22 in *Klebsiella pneumoniae*); ii) assessed the potential of FT-IR spectroscopy to differentiate and identify clinically-relevant *Enterobacter cloacae* species compared to reference methods; iii) extended FT-IR application to detect hypervirulent *K. pneumoniae*. Phenotypic, genotypic and molecular methods were combined with FT-IR spectroscopy and multivariate data analysis, used to interpret spectral patterns and build classification models.

This work underscored the need to integrate phenotypic and targeted molecular approaches for comprehensive carbapenemase surveillance. It further demonstrated FT-IR's effectiveness in species and subspecies differentiation, underscoring its role as a rapid and reliable diagnostic tool for prompt treatment, minimizing the spread of multidrug-resistant pathogens, and improving patient care.

KEYWORDS: Diagnostic, FT-IR spectroscopy, Antimicrobial resistance, *Klebsiella pneumoniae*, *Enterobacter cloacae* complex

ACKNOWLEDGMENTS

First and foremost, I owe a tremendous debt of gratitude to my family, whose unwavering support and encouragement have been essential throughout my educational journey. To my parents, thank you for your endless love and belief in me—this achievement would not have been possible without you.

I would also like to extend a special acknowledgment to my boyfriend, Bruno, and my best friend, Mariana, who were a constant source of strength and motivation. Your understanding and encouragement during the most challenging moments have made a profound impact on my success.

Also, a special thanks to Márcia, Liliane and especially Beatriz for their generous assistance and support during my time in the laboratory. Your kindness and professionalism have truly inspired me, and without you I wouldn´t be able to finish this.

My profound gratitude to the Bill and Melinda Gates Foundation for funding this work and me through Grand Challenges (reference INV060596), iH4B, and UCIBIO, which also played a huge role through FCT (UIDP/04378/2020, UIDB/04378/2020, LA/P/0140/2020).

Lastly, I wish to acknowledge my advisor, Professor Dra. Ângela Novais, for her patience, dedication, and insightful guidance in bringing this work to fruition. I am also grateful to Professor Dra. Luísa Peixe, for the opportunity of working in her research group in the first place.

To all those mentioned and those who have supported me in various ways, my deepest appreciation goes out to you all.

CONTENTS

INDEX OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

AMR—Antimicrobial resistance

ATR— Attenuated total reflection

BLAST—Basic Local Alignment Search Tool

CPS—Capsular polysaccharides

DALYs—Disability-adjusted life years

ECDC—European Centre for Disease Prevention and Control

*Ec*SC—*Enterobacter cloacae* species complex

EPS—Extracellular polysaccharides

EUCAST— European Committee on Antimicrobial Susceptibility Testing

EU/EEA—European Union and European Economic Area

FT-IR—Fourier Transform Infrared

HAIs—Healthcare-Associated Infections

HCA—Hierarchical cluster analysis

Hv-*Kp*—Hypervirulent *K. pneumoniae*

HMV—Hypermucoviscous

IPC—Infection prevention and control

*Kp*SC—*Klebsiella pneumoniae* species complex

LightGBM—Light Gradient-Boosting Machine

LMICs—Low- and middle-income countries

LPS— Lipopolysaccharides

MALDI-TOF MS—Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

MBLs—Metallo-beta-lactamases

MDR—Multidrug-resistant

MGEs—Mobile genetic elements

MDR-Hv-*Kp*—Convergent *K. pneumoniae*

MDR-*Kp*—Multidrug-resistant *K. pneumoniae*

MLST—Multi-Locus Sequence Typing

NPV—Negative predictive value

PCA—Principal component analysis

PCR—Polymerase chain reaction

PDR—Pandrug-resistant

PFGE—Pulsed-field gel electrophoresis

PLS-DA—Partial least square discriminant analysis

PPV— Positive predictive value

RF—Random Forest

SMOTE—Synthetic minority oversampling technique

SMV—Support vector machines

SNV— standard normal variate

SPs—Surface polysaccharides

ST—Sequence type

WHO—World Health Organization

WGS—Whole-genome sequencing

XDR—Extensively drug-resistant

ZnSe— Zinc selenide

1. INTRODUCTION

1.1 Healthcare-Associated Infections (HAIs)

Healthcare-associated bacterial infections (HAIs) are infections acquired by patients during hospital stays, typically emerging 48 hours or more after admission. Most (86%) HAIs predominantly affect the urinary tract, the respiratory tract or the skin (1). HAIs generally arise from bacteria colonizing hospitalized patients or from contaminated environments, significantly contributing to the persistence and dissemination of pathogens within healthcare facilities, as well as at regional or global levels (2, 3) (Figure 1).

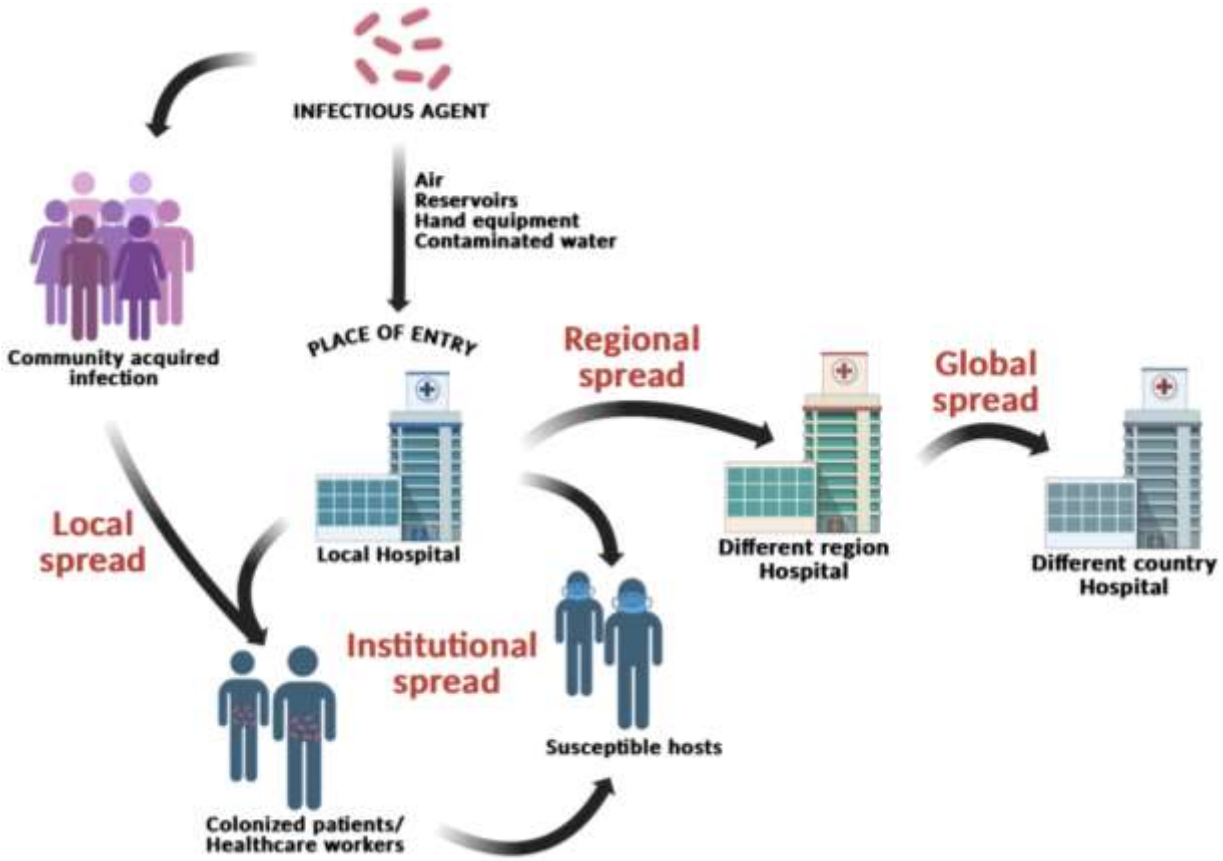


Figure 1. Transmission pathways of HAI infections within healthcare settings

(Created with BioRender resources).

In the early 2000s, HAIs accounted for approximately 37,000 attributable deaths annually in Europe and around 99,000 deaths in the United States (4). Currently, HAIs represent a significant and escalating Public Health concern, with recent estimates indicating that as many as 9.5 million deaths occur globally each year (5). In the European Union and European Economic Area (EU/EEA) alone, more than 2.5 million cases of HAIs are reported each year, leading to an estimated 2.5 million disability-adjusted life years (DALYs) lost annually. Besides, it is associated with extended hospital stays, increased antibiotic resistance and, consequently, substantial financial burdens on healthcare systems worldwide. Annual estimates point to USD 147 million in Brazil (6), £774 million in the United Kingdom (6), USD 9.8 billion in the United States (7), USD 13 billion in Africa (8), and USD 955 million in New Zealand (9). The estimated global prevalence of HAIs is approximately 14%, with low- and middle-income countries (LMICs) being disproportionately affected. Europe has the lowest rate (8%), followed by the Americas region (9.6%), the Western Pacific Region (9.7%), the Eastern Mediterranean Region (12.5%), the South-East Asia Region (12.9%) and finally, with the highest prevalence in Africa (27%) (10).

According to the European Centre for Disease Prevention and Control (ECDC), HAIs are mainly caused by ten bacterial species with *Escherichia coli* and *Klebsiella pneumoniae* being the most prevalent ones (Figure 2) (1).

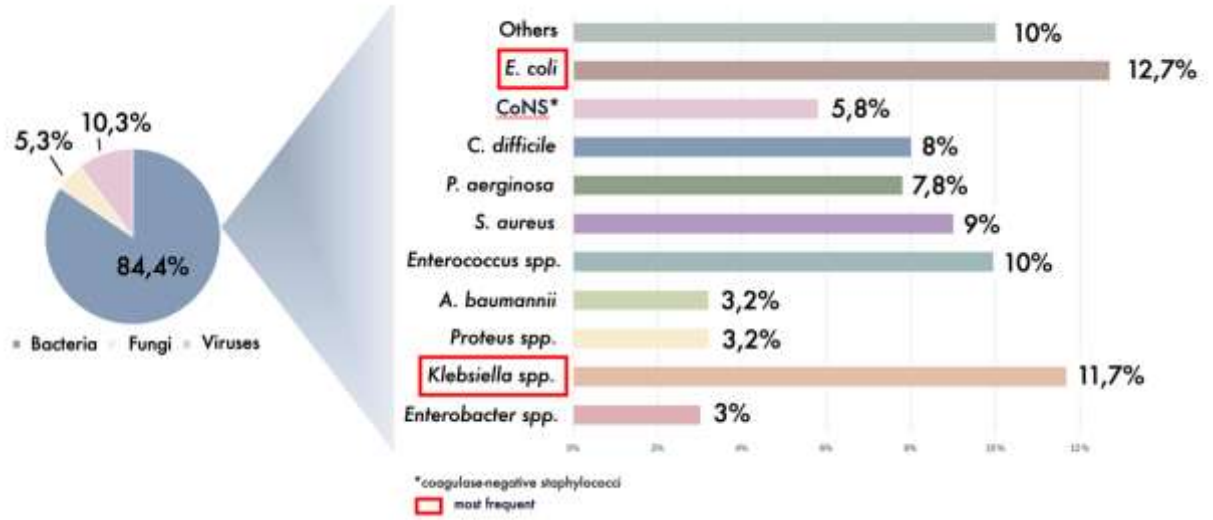


Figure 2. The most common microorganisms causing HAIs in hospitals in Europe (based on data from ECDC).

Among pathogens causing HAIs, infections caused by multidrug-resistant (MDR) bacteria are of particular concern, not only because they pose the greatest threat due to limited treatment options and worse clinical outcomes, but also because they represent 23% of all HAIs (1). The World Health Organization (WHO) and the ECDC have established a list of 24 pathogens considered as "priority" to guide research, surveillance, and development of new treatments, divided into three priority categories (critical, high, and medium) (11). This list includes, but is not limited to, the ESKAPE group, an acronym introduced by Louis Rice in 2008, which refers to six highly problematic bacteria: *Enterococcus faecium, Staphylococcus aureus, K. pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter* spp., notorious for their remarkable ability to develop resistance to multiple antibiotics (12).

Antimicrobial resistance (AMR) arises from both intrinsic and acquired resistance mechanisms. Intrinsic resistance, naturally present in certain bacterial species, results from inherent structural or functional features that render specific antimicrobials. In contrast, acquired resistance emerges when bacteria acquire mutations or new genetic traits that confer resistance to antimicrobials to which they were previously susceptible. The acquisition of new genetic traits is frequently mediated by mobile genetic elements (MGEs), such as plasmids and transposons, and these mechanisms are spread through horizontal gene transfer. A critical aspect is the frequent colocalization of multiple resistance genes on the same genetic platform, a phenomenon known as co-resistance (13, 14). This genetic linkage enables the simultaneous acquisition and spread of resistance to several classes of antibiotics, resulting in the emergence of MDR phenotypes, extensively drug-resistant (XDR), or pandrug-resistant (PDR) phenotypes (15), significantly limiting treatment options:

- MDR: resistant to more than one antimicrobial agent in ≥3 antimicrobial categories;
- XDR: resistant to more than one antimicrobial agent in all but 2 antimicrobial categories;
- PDR: resistant to all antimicrobial agents

AMR is a growing problem that is attracting global attention due to worrying projections. O'Neill et al. have reported that, by 2050, drug-resistant infections could claim 10 million lives annually and place a cumulative USD 100 trillion of global economic output at risk, unless we implement preventive solutions to curb the rise of AMR (16). Recent studies have shown that the worldwide increase in AMR rates is significantly linked to the transmission of MDR bacteria

within healthcare settings (17) as well as to the inadequate and excessive use of antibiotics in humans, animals, and agriculture (11, 18). Other contributing factors include inadequate infection prevention and control (IPC) measures, environmental contamination from pharmaceutical manufacturing and agricultural runoff, insufficient regulation of antibiotic use and prescription practices and limited development of new antimicrobial agents. Socioeconomic factors such as income, education and limited access to clean water and sanitation in particular settings also influence AMR rates, with higher resistance often observed in Low- and middle-income countries (LMIC) settings. Therefore, there is a well-recognized need for comprehensive, multisectoral strategies to address and combat AMR and control HAIs, especially those caused by MDR bacteria (16). At the core of these initiatives is the deployment of robust IPC measures, comprising rigorous compliance with hygiene protocols and isolation procedures. Support through extensive surveillance systems for the prompt identification and monitoring of HAIs and their associated risk factors, antimicrobial stewardship programs, ongoing staff training, and strategic investments in healthcare infrastructure are also essential elements (19, 20). Therefore, major international organizations and health authorities have outlined critical measures to curtail the spread of AMR globally, and identified persistent challenges for their implementation, as detailed in Table 1 (1, 16, 21).

Table 1. Key measures and challenges to tackle HAIs and AMR worldwide.

Key measures	Description	Challenges
Public Awareness	Global awareness of AMR, avoiding unnecessary demand and overprescription of antibiotics.	• Expensive: predicted costs within 40 to 100 million USD/year; • Significant disparities in health literacy and access to information, especially in LMICs.
Hygiene	Improving hygiene measures and sanitation, for infectious diseases prevention and limiting the spread of MDR pathogens.	• Limited resources and lack of awareness, especially in LMICs.
Antibiotic stewardship	Promoting the responsible and appropriate use of antibiotics in human and veterinary medicine, ensuring correct selection, dosage, and duration.	• Easy access to antibiotics without prescription, especially in LMICS; • Off label use and counterfeiting of antibiotics, especially in LMICs; • Excessive use of antimicrobials for infection prevention in animals; • Lack of regulation and oversight in both human and veterinary sectors.
Global surveillance	Establishing a standardized global surveillance data network, to track antibiotic usage, resistance patterns, and infection outbreaks.	• Underreporting; • Lack of investment in surveillance infrastructure and data sharing.
Rapid diagnostics tools	Rapid and accurate diagnostic tools allowing timely identification of pathogens and detection of resistance markers	• High cost; • Limited accessibility; • Lack of standardization; • Reliance on traditional methods, especially in LMICs
Antibiotic Development	Investing in research and developing new antibiotics and alternative therapies, to expand treatment options and to replace ineffective drugs	• High failure rates; • Low innovation due to lack of financial incentives; • Low interest from pharmaceutical companies due to low profitability and high costs

The work undertaken in this thesis focuses on one of these complementary strategies—namely, the development of rapid and precise diagnostic tools—as it plays a pivotal role in guiding appropriate treatment, preventing unnecessary antibiotic use, and ultimately mitigating the dissemination of antimicrobial resistance.

1.2 Diagnostic methods to support the control of HAIs and AMR

In clinical microbiology laboratories, a precise identification of bacterial pathogens at the species and strain levels, and antimicrobial resistance patterns is key for infection management. Discriminatory and accurate diagnostic methods are essential to support empiric and/or targeted therapy, contributing to reducing mortality rates, hospital stays, healthcare costs, and unnecessary antibiotic usage. Furthermore, they play a crucial role in IPC by supporting effective surveillance, detection and management of outbreaks (19). Therefore, in addition to being accurate and discriminatory, diagnostic methods should ideally be simple, fast, and cost-effective to facilitate implementation in routine laboratory contexts and to allow information to be available at the earliest stage possible.

The following sections provide an overview of the main diagnostic approaches used in routine clinical microbiology laboratories to support pathogen identification and infection control according to their specific purpose.

1.2.1 Detection of antibiotic resistance markers

In routine microbiology laboratories, antibiotic susceptibility testing is the gold standard to assess phenotypic resistance, key information to guide empirical therapy. It is commonly performed by semi-automated methods, including VITEK2 or MicroScan or standard agar diffusion (disc diffusion or gradient diffusion) and/or microdilution methods, according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST, www.eucast.org) guidelines (19).

Identification of specific antimicrobial resistance markers is important for surveillance and IPC purposes. Different types of commercial tests are available to detect specific mechanisms of resistance of high clinical relevance. Among these, biochemical assays such as the Carba NP

test (Biomérieux) are commonly used for the rapid identification of carbapenemase producers, while immunochromatographic assays (e.g., O.K.N.V.I. RESIST-5, Coris BioConcept) enable the detection of major carbapenemase families through antibody-based recognition of resistance enzymes. These methods are widely used in hospitals for routine testing of bacterial isolates that have been cultured from clinical specimens, but also samples (e.g., rectal, nasal) to screen carriers of antimicrobial-resistant bacteria. Due to their cost-effectiveness and ease of implementation without the need for specialized expertise, they enable rapid identification (within 2 hours) of clinically significant resistance mechanisms, thereby facilitating timely clinical decision-making, infection control, and surveillance efforts (22, 23).

Molecular-based methods offer a precise characterization of antimicrobial resistance determinants. They provide enhanced sensitivity and specificity for the direct detection of a wide array of specific antibiotic resistance genes and their variants. Commercial real-time polymerase chain reaction (PCR) kits, such as the GeneXpert system (Xpert® Carba R, Cepheid, USA), are widely used in clinical laboratories for the rapid identification of specific antibiotic resistance markers, though their implementation varies with the strategy and capabilities of the hospitals. These platforms are highly automated, require minimal hands-on time, and can deliver results within 2 hours, making them particularly valuable for quick clinical decision-making. Their multiplexing capability allows for the simultaneous detection of multiple resistance genes in a single reaction (24). These tests can be performed on clinical samples such as respiratory samples (e.g., sputum, bronchoalveolar lavage), blood culture bottles, urine, and swabs from skin or soft tissue infections (24-26). They can also be used for screening nasal or gastrointestinal swabs to identify carriers of multidrug-resistant bacteria. Such screening enables the early identification of colonized patients who may serve as reservoirs for transmission and allows for the implementation of targeted IPC measures to prevent the spread of antibiotic-resistant bacteria within healthcare settings (24).

However, these routine phenotypic, immunochromatographic and molecular approaches have some limitations in sensitivity and/or specificity. Their performance can vary significantly across different carbapenemase families, and they may fail to reliably detect enzymes with weak carbapenemase activity, such as OXA-48-like variants (27, 28), certain metallo-beta-lactamases (MBLs) like some rare IMP variants (29, 30) or even certain KPC variants with enhanced capacity to hydrolyze ceftazidime-avibactam and other last-line antibiotics (31, 32). This raises

concerns about the potential underreporting of emerging variants and the limitations of surveillance strategies relying solely on standard molecular screening. In fact, there is considerable genetic diversity within each carbapenemase family (e.g., multiple NDM, KPC, and IMP variants), and their precise identification is relevant since they may differ in their prevalence, transmission dynamics, or hydrolytic capabilities. For this reason, it is important to assess the reliability of phenotypic, genotypic and routine molecular methods for the identification of rare or atypical resistance variants in real clinical microbiology contexts. This will allow empowering healthcare professionals and support diagnostic decisions towards accurate carbapenemase surveillance and outbreak management.

Whole genome sequence (WGS) represents the gold-standard molecular approach, offering complete genotypic profiling of bacterial isolates. WGS enables the identification of all known resistance genes, their variants, and even the discovery of novel resistance determinants, thus providing an exhaustive overview of the genetic basis of resistance within a given isolate (19). However, it is not implemented in the vast majority of routine clinical laboratories due to the complexity, substantial costs, the demand for experienced personnel for technical procedures and time-consuming (3-7 days in the best scenarios) (33, 34).

More recently, innovative approaches using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) are being explored for their potential application in routine AMR detection. These methods are being evaluated for their ability to complement or surpass traditional phenotypic and molecular approaches, offering distinct advantages in terms of speed, versatility, and mechanistic insight (19, 35). By analyzing the mass spectra of bacterial proteins or the degradation products of antibiotics, MALDI-TOF MS can provide results within a few hours. Its competitive advantage lies in its high-throughput capacity, minimal sample preparation, and the possibility of integrating resistance detection into existing identification workflows (36). However, its application for AMR detection remains limited to a few laboratories due to the need of standardized workflows and reliance on in-house protocols and databases for interpretation.

1.2.2 Species-level identification

Molecular methods including the identification of specific genetic markers (e.g.,16S rRNA or *rpoB*) are more accurate and discriminatory for bacterial identification. WGS provides markedly higher resolution since it includes the comparison of hundreds of genetic markers. However, as explained above, it's not yet routinely implemented.

Automated systems, such as the VITEK® 2 (bioMérieux), are fluorescence-based methods employed for the rapid identification of bacteria within routine microbiology laboratories (37). However, there are several reports of limitations in accurately identifying certain clinically important or rare bacteria (37-39). MALDI-TOF MS has been introduced in clinical microbiology and has revolutionized species-level identification by providing rapid and cost-effective results, significantly improving laboratory efficiency and patient care (40, 41). Nonetheless, it exhibits limited discriminatory power in distinguishing closely related species or uncommon organisms (42, 43). This is the case of bacteria belonging to the *Enterobacter cloacae* species complex (*Ec*SC) and the *Klebsiella pneumoniae* species complex (*Kp*SC) (44-47). Recent research efforts have focused on overcoming the species-level identification limitations within the KpSC using MALDI-TOF MS. Rodrigues *et al*. (48) reported the first mass spectrometry biomarkers capable of discriminating all phylogroups within the *K. pneumoniae* complex, thereby enhancing the accurate identification of these pathogens. Difficulties remain in the genus *Enterobacter*, for which no routine methods are currently available for accurate differentiation, including MALDI-TOF MS (44, 49, 50). The *Ec*SC are opportunistic pathogens responsible for hospital-acquired infections, with an increasing occurrence of multidrug-resistant strains (51). *Enterobacter xiangfangensis, E. hoffmannii, Enterobacter roggenkampii, Enterobacter kobei*, and *Enterobacter asburiae* are acknowledged as the most frequently encountered in clinical settings (50, 52, 53). Recently, advances in WGS and the growing availability of genomic data have led to critical taxonomic restructurings within the genus (53). As such, alternative methods are needed for accurate and rapid methods for identification of *Ec*SC.

1.2.3 Strain-typing for outbreak control and surveillance

Bacterial typing methods enable the differentiation of bacterial strains within the same species that derive from a common ancestor by analyzing their genetic, phenotypic, or biochemical features. This information is important for the elucidation of sources and transmission pathways in outbreaks or in endemic scenarios. This supports the implementation of targeted infection control measures to interrupt transmission, control outbreaks, prevent new cases, and ultimately limit the dissemination of MDR bacteria within healthcare facilities. Typing techniques can rely on the differentiation of phenotypic (e.g., biochemical profiles, surface antigen variation), genotypic (e.g., specific genetic markers or the whole genome) and, more recently, spectrometric methods [e.g., Fourier Transform Infrared (FT-IR) spectroscopy] (54). Table 2 presents an overview of the most common phenotypic, genotypic, and spectrometric methods for bacterial strain typing, emphasizing their principles, advantages, and limitations (54-56).

Currently, WGS is widely recognized as the gold standard in bacterial typing, exhibiting the highest discriminatory power among all typing methodologies. When coupled with advanced bioinformatics tools, this approach enables precise characterization of phylogenetic relationships, AMR profiles, and virulence factors. However, as explained above, its implementation as a routine method is challenging (54, 55).

While other established methods, such as pulsed-field gel electrophoresis (PFGE) and multi-locus sequence typing (MLST), continue to play important roles in epidemiological investigations at local or global surveillance, respectively, they also present limitations in terms of turnaround time. Phenotypic methods like serotyping, although standardized and still useful for certain pathogens (e.g., *Salmonella*), are increasingly being replaced by molecular and spectrometric approaches due to their lower resolution and higher labor demands.

In this context, FT-IR spectroscopy stands out by combining several desirable attributes for alternative routine implementation: it is rapid, cost-effective and simple to perform, providing high discriminatory power for bacterial differentiation (55). Therefore, the next chapter will focus specifically on FT-IR spectroscopy, exploring its applications and fundamentals in bacterial typing.

Table 2. Principles, benefits and limitations of main bacterial typing methods.

	Method	Description	Benefits	Limitations
Phenotypic	Serotyping	Analyses specific types of antigens present on cell surfaces (e.g., O- or K-antigen)	• High discriminatory power • Medium reproducibility (depends on organism, protocol and reagents used)	• High labor intensity • Requires experienced personnel • Specialized material (sera) • Time consuming (2-7 days) • Medium costs
Genotypic	MLST	Compares allelic variations of 7 housekeeping internal gene fragments used to define a specific sequence type (ST)	• Medium discriminatory power • High reproducibility • Low labor intensity • Low costs	• Time-consuming (1-5 days) • Requires experienced personnel
	PFGE	Uses a pulsed electric field to generate a fingerprint of large DNA fragments, after enzymatic hydrolysis	• High discriminatory power • Medium reproducibility • Low costs	• High labor intensity • Time-consuming (3-5 days)
	WGS	Determines the entire DNA sequence of a bacterial genome, creating a detailed genetic profile	• High discriminatory power • High reproducibility	• High labor intensity • High costs • Requires experienced personnel • Time-consuming (1-30 days) • Complex interpretation
Spectrometric	FT-IR spectroscopy	Uses IR radiation to cause molecular vibrations, generating fingerprints of the cell composition in biomolecules	• High discriminatory power • Medium reproducibility • Low labor intensity • Low costs	• Complex spectral databases • Environmental and technical standardization • Misidentification if database is incomplete

1.3 FT-IR spectroscopy

1.3.1 Fundamentals

FT-IR spectroscopy uses infrared light to analyze the absorption and emission patterns of a sample, providing information on the excitation and vibrational modes of the molecules present. When applied to a bacterial cell, the result is a highly discriminatory spectrum characterized by specific spectral peaks corresponding to different functional groups, each absorbing radiation at various wavenumber lengths. Therefore, the mid-infrared spectrum (4000-400 cm-1) acquired reveals the composition of nucleic acids (1500-1200 cm-1), proteins (1800-1500 cm-1), lipids (3000-2800 cm-1) and carbohydrates (1200-900 cm-1) that constitute a bacterial cell. Each spectrum serves as a molecular fingerprint of the predominant constituents of bacterial cells, with the carbohydrate region being one of the most discriminating spectral regions (Figure 3) (55).

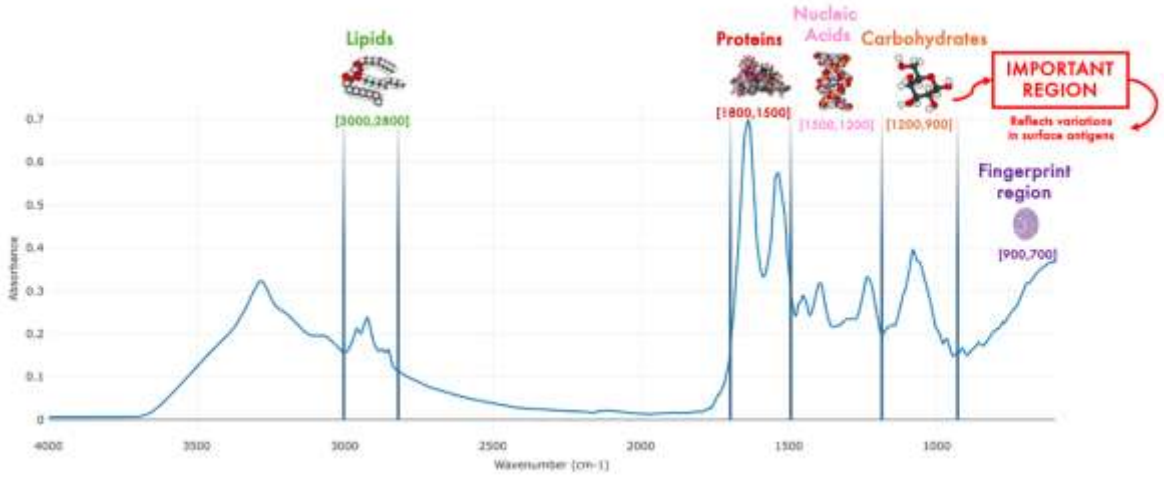


Figure 3. A typical FT-IR ATR bacterial spectrum and its correspondence to each functional group.

(Created with BioRender resources)

FT-IR can be applicable to solid, liquid or gaseous samples and with different modes of acquisition: i) transmission, ii) diffuse reflection and iii) attenuated total reflection (ATR). The transmission mode method entails the direct passage of infrared light through the sample and

the detection of the radiation beam transmitted by the sample. The diffuse reflection mode is based on detecting the scattered beam of infrared radiation after sample interaction, but it is rarely used for bacteria due to its higher cost. The ATR mode detects infrared radiation's reflection when the beam transitions from a high-density crystal, such as zinc selenide (ZnSe) or diamond, to a lower-density material as a bacterial sample. ATR offers simplicity of protocol (comparable to MALDI-TOF MS), and reduced costs (55, 56). Reproducibility is higher as long as culture and environmental conditions (especially humidity) are stable (55), and recent studies demonstrated stability across different culture media and equipment (57).

1.3.2 FT-IR experimental workflow

The experimental workflow to use FT-IR is based on 3 steps: sample preparation, FT-IR spectra acquisition and data analysis. After the FT-IR spectrum is obtained, data analysis will comprise pre-processing, followed by spectral comparisons using statistical and mathematical algorithms, allowing for bacterial typing and discrimination (55, 56, 58).

For the transmission mode sample preparation, we apply a bacterial suspension to optical plates composed of water-insoluble materials (typically ZnSe), followed by drying at 50–60°C or under vacuum before analysis. As for ATR mode, this technique requires minimal sample preparation, allowing for direct application of colonies or suspensions onto the crystal. Afterwards, spectra are generated in only 2-3 minutes, reflecting the molecular vibrations of the characteristic of each sample.

FT-IR spectra undergo pre-processing, which includes removing noise caused by the instrument or environmental factors (e.g., water vapor, CO_2, temperature), correcting the baseline to eliminate background signals or drifts from scattering effects or instrument artifacts, and normalizing to account for variations in sample concentration or path length, enabling meaningful comparisons between spectra (59).

Then, spectra are analyzed to interpret the complex biochemical patterns obtained through multivariate data analysis. Two types of statistical methods can be used for data analysis: supervised and unsupervised. Unsupervised methods such as principal component analysis (PCA) or hierarchical cluster analysis (HCA) identify inherent groupings in datasets without prior sample identity information, enabling categorization of unknown samples according to

intrinsic spectra/sample variation. Supervised methods such as partial least square discriminant analysis (PLS-DA), support vector machines (SVM), Light Gradient-Boosting Machine (LightGBM) and random forest (RF) require prior knowledge about sample identity. These rely on a well-characterized sample database assigned to specific classes to train a classification model, which is then tested with unknown samples to evaluate prediction accuracy (Figure 4).

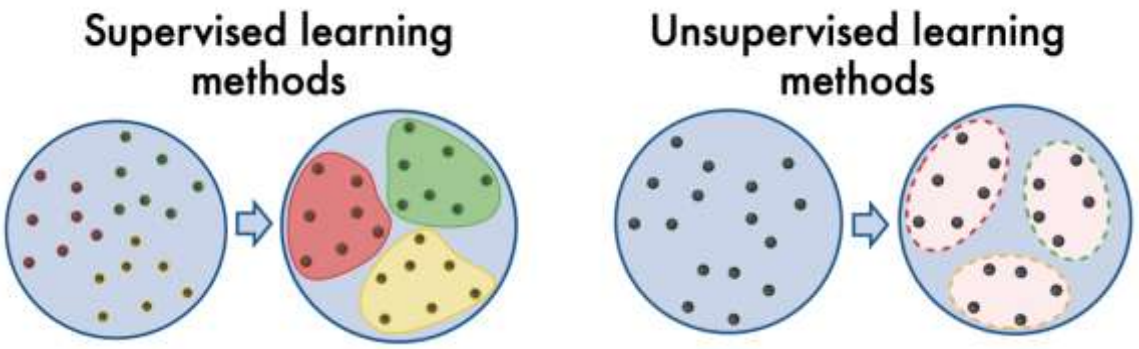


Figure 4. Schematic representation of the difference between unsupervised and supervised learning methods.

(Created with BioRender resources)

1.3.3 Historical evolution and applications of FT-IR

FT-IR spectroscopy was first applied to microorganisms back in the 1950s. However, it was not until the early 1990s that Naumann demonstrated the potential of FT-IR to differentiate bacterial samples, illustrating that each microorganism presents a unique infrared fingerprint spectrum. While for many years the molecular mechanisms underlying this discrimination remained elusive, recent research has clarified that these distinctions are underpinned by variations in bacterial surface structure composition across different species (55).

In subsequent years, studies have established a stronger correlation between FT-IR spectral profiles and reference genotypic methods, including WGS (48, 52-56). Different studies across various bacterial species have shown that this bacterial differentiation relies on molecular differences in bacterial surface antigens (O-antigen, K-antigen). Consequently, comprehensive and representative spectral databases have facilitated the development of accurate classification

models, which were then used to support bacterial typing in both retrospective and prospective studies (57-62).

Building on these advancements, the introduction of the IR Biotyper system by Bruker Daltonics in 2016 marked a pivotal moment in the field, catalyzing a rapid and exponential increase in the use of FT-IR technology for bacterial strain typing. This integrated system, which combines spectrum acquisition in the transmission mode with a smooth workflow for spectral data analysis offered a solution compatible with clinical routine workflows (55, 60). Standardizing sample preparation protocols and automated spectral processing and data analysis has facilitated its use in daily routine clinical microbiology laboratories, leading to a substantial growth in the number of users and research employed in FT-IR bacterial strain typing and outbreak detection (60).

Throughout the years, numerous studies have validated FT-IR spectroscopy as a rapid, cost-effective tool for real-time outbreak detection and IPC decision-making, enabling same-day interventions during critical transmission events. These studies have underscored the efficacy of FT-IR across clinically relevant and frequently multidrug-resistant pathogens like *E. coli*, *K. pneumoniae, Salmonella enterica,* and *A. baumannii* (61-64).

The usefulness and applications of FT-IR spectroscopy have been extensively studied for *K. pneumoniae*. Rodrigues *et al.* (65) correlated the FT-IR spectral profiles with a variation of the polysaccharide composition of the capsule (K antigen). Novais *et al.* (57) developed and validated a classification model for the accurate identification of 33 capsular types (KL)-types associated with globally disseminated MDR clones. More recently, Gonçalves *et al.* (66) demonstrated the practical application of FT-IR spectroscopy as a frontline tool for real-time surveillance and hospital outbreak management.

Although multidrug-resistant *K. pneumoniae* (MDR-*Kp*) has been extensively studied due to its prevalence as a major nosocomial pathogen and its significant impact on global health, the discrimination of hypervirulent *K. pneumoniae* (Hv-*Kp*) remains largely unaddressed. Hv-*Kp* is a highly virulent pathotype of *K. pneumoniae*, known for its ability to cause severe and disseminated infections in healthy individuals, including pyogenic liver abscesses, endophthalmitis, and meningitis, caused by a rapid metastatic spread of the bacteria from the initial infection sites (67). Though more frequent in Asia, Hv-*Kp* strains are increasingly being reported in Europe and worldwide (68). Most infections are caused by strains exhibiting specific

KL types (KL1, KL2, KL5) which are usually antibiotic susceptible (69). More recently, convergent (MDR-Hv-*Kp*) strains are emerging combining both phenotypes (MDR and HV), rendering severe and hard-to-treat infections (70, 71). This situation is worrying not only because of the infection outcome but also because Hv-*Kp* strains are difficult to detect in routine settings. They are well recognized for their abundant capsular production, which can be identified by the mucoidy of the bacterial culture and the string test. A positive string test is considered when a viscous filament exceeding 5 mm in length is observed after stretching a *Klebsiella* spp. colony with a loop on an agar plate. However, not all Hv-*Kp* strains exhibit this phenotype and some MDR-*Kp* may acquire this property (72). As such, this method lacks sensitivity and specificity. Therefore, identification of key virulence markers and *in vivo* evaluation of hypervirulence are the gold standard for identification of Hv-*Kp*. Two main frameworks define key virulence genes: Russo *et al.* (72) focused on *rmpA, rmpA2* (hypermucoviscosity regulators), *iroB, iucA* (iron acquisition), and *peg-344* (putative transporter), while Lam *et al.* (69), via the Kleborate tool, included *ybt* (infection/carriage), *clb* (tissue damage/oncogenic potential), as well as *iro, iuc*, and *rmpA*. It is also important to consistently correlate the results with clinical history and patient outcomes.

Nevertheless, despite their accuracy, this genomic approach is not ideal for real-time clinical decision-making due to its complexity and turnaround time because it requires multiple conventional PCR or WGS. Thus, quick and reliable methods are needed to support the identification of Hv-*Kp* in routine clinical microbiology laboratories. Since this population exhibits specific KL types, we hypothesize that FT-IR spectroscopy might constitute a reliable, rapid and accessible method for Hv-*Kp* identification.

1.4 Purpose of the study

The **main purpose** of this work is to improve diagnostic strategies for accurate, quick and simple detection of clinically relevant pathogens and antimicrobial resistance mechanisms in routine clinical microbiology laboratories. To accomplish this aim, this thesis will extend FT-IR spectroscopy applications to answer specific diagnostic challenges. **Specific objectives are outlined below**:

1. To clarify and identify limitations associated with routine methods for detecting rare or atypical carbapenemases, based on a real clinical case.
2. To assess the potential of FT-IR spectroscopy as an alternative diagnostic method for fast and accessible identification of the main clinically relevant *Ec*SC;
3. To extend the application of FT-IR spectroscopy to Hv-*Kp* strains and provide a universal tool for typing clinically relevant and globally representative MDR-*Kp* and Hv-*Kp*.

2. MATERIALS AND METHODS

2.1. Bacterial isolates

2.1.1. Detection of carbapenemases by molecular routine methods

During a routine microbiological investigation of a real clinical case, involving a two-month-old infant with a urinary tract infection, inconsistencies were observed between commonly used phenotypic and molecular diagnostic methods for carbapenemase detection. These discrepancies prompted further investigation of the limitations inherent in current diagnostic approaches for detecting rare or atypical variants.

One *K. pneumoniae* was detected in a urine culture of a two-month-old infant diagnosed with a urinary tract infection, who was admitted to the hospital on 22nd August 2024. This isolate was resistant to cefotaxime, ceftazidime-avibactam, amoxicillin-clavulanic acid, cefoxitin, cefepime, and ertapenem, and was suspected of producing a carbapenemase. In a subsequent urine culture (10th of September), another *K. pneumoniae* was detected with additional resistance to meropenem and imipenem. The immunochromatographic test (O.K.N.V.I. RESIST-5, Coris BioConcept) yielded positive for IMP carbapenemase, but both Xpert® Carba R (Cepheid, USA) and BD Max™ CPO (BD, USA) were negative.

Two rectal screenings of the infant were done on the 10th and 17th of September to assess gastrointestinal colonization, resulting in the detection of two carbapenem-resistant *Klebsiella* sp.. VITEK MS classified these screening isolates as *K. variicola*, while MALDI-TOF MS identified them as *K. pneumoniae*.

2.1.2. Differentiation of *Ec*SC by FT-IR spectroscopy

Within 25 *Ec*SC described to date (73), five (*E. xiangfangensis, E. kobei, E asburiae, E. roggenkampii, E. hoffmannii*) correspond to ~86% of the isolates identified in clinical settings. A set of 159 representative isolates of these 5 species was selected, preliminarily identified by VITEK 2 and/or MALDI-TOF MS at the hospital of origin. They were collected from two hospitals in northern Portugal, one hospital in Madrid and one hospital in Seville. Most of these

isolates are multidrug-resistant isolates, 65% carbapenemase producers, isolated from either infection or colonization between 2005-2021 (Table 3). Species identification was performed by MALDI-TOF MS, PCR and sequencing of a short (341 bp) of a large (968bp) fragment of *hsp60* and/or whole genome sequencing.

Table 3. Origin and diversity of Enterobacter spp. isolates included in this study.

Species	Porto (nº isolates)	Madrid (nº isolates)	Seville (nº isolates)	Total (nº isolates)	nº of clones	Origin	Date
E. xiangfangensis	14	24	0	38	33	I (n=23) C (n=15)	2005-2021
E. kobei	7	15	10	32	25	I (n=22) C (n=10)	2009-2021
E. asburiae	3	8	14	25	17	I (n=17) C (n=8)	2011-2021
E. roggenkampii	0	33	0	33	23	I (n=19) C (n=14)	2005-2018
E. hoffmannii	4	16	11	31	20	I (n=24) C (n=7)	2006-2021

(I=Infection; C=Colonization)

2.1.3. Identification of globally disseminated Hv-*Kp* and MDR-*Kp*

A collection of 45 Hv-*Kp* strains representing main KL types (K1, K2, K5, n=39) and other variable KL types (KL20, KL30, KL54, KL57, n=6) was gathered. These isolates were recovered from different sources and countries as follows: i) catalog NR-55604 - BEI resources (USA); ii) Child Health Research Foundation (Dhaka, Bangladesh); iii) Institut Pasteur (Paris, France); iv) School of Hygiene & Tropical Medicine (London, UK). All have been previously characterized by WGS and had a virulence score according to Kleborate ≥ 3 and/or were known to cause severe/disseminated infections (69). Our collection includes also 9 putative convergent MDR-Hv-*Kp* isolates with different KL types (KL1, KL20, KL43, KL51, KL62, KL64) (Table 4). These convergent represent both the acquisition of virulence genes by MDR-*Kp* or resistance genes by Hv-*Kp*.

A set of 146 well-characterized MDR-*Kp* strains representative of the ten main clinically relevant KL types worldwide (KL2, KL24, KL25, KL62, KL64, KL102, KL105, KL106, KL107, and KL112) was included for training a supervised model (69, 74, 75) (Table 5). Additionally, a set of 50 MDR-*Kp* isolates from these same 10 KL types, plus 18 MDR-*Kp* isolates belonging to 17 different KL types, were selected for external validation (Table 6).

Table 4. Origin and characteristics of Hv-*Kp* and MDR-Hv-*Kp* isolates included in this study.

	KL type	ST	Origin	Date
Hv-*Kp*	KL1 (n=10)	ST23 (n=9) ST310	BEI CHRF (n=3) IP (n=4) SHTM (n=2)	2007-2016
	KL2 (n=20)	ST25 ST39 ST325 ST493 ST86 (n=6) ST375 (n=2) ST380 (n=4) ST65 (n=2) ST66 (n=2)	BEI (n=3); SHTM (n=5); IP (n=12)	1935-2019
	KL5 (n=9)	ST2701 ST45 ST4948 ST76 ST4039 ST2449 ST968 ST60 (n=2)	IP	2015-2018
	KL20 (n=2)	ST420	CHRF	2007
	KL30	ST43	CHRF	2019
	KL54 (n=2)	ST29 ST1109	IP CHRF	2013-2015
	KL57	ST500	IP	2016
MDR-Hv-*Kp*	KL64 (n=2)	ST147	CHRF SHTM	2007-2019
	KL20	ST268	BEI	2017
	KL51	ST2071	BEI	2016

Table 4. Origin and characteristics of Hv-*Kp* and MDR-Hv-*Kp* isolates included in this study.

	KL15	ST37	BEI	2019
	KL1	ST23	BEI	2016
	KL62 (n=2)	ST48	CHRF	2007
	KL43	ST1399	BEI	2014

(IP: Institute Pasteur; CHRF: Child Health Research Foundation; SHTM: School of Hygiene & Tropical Medicine)

Table 5. Origin and characteristics of MDR-Kp isolates included in this study for a training set.

KL type	ST	Origin	Date
KL2 (n=20)	ST14 (n=13); ST25 (n=7)	Portugal (n=14) Spain (n=6),	2003-2024
KL24 (n=27)	ST15 (n=9); ST45 (n=18)	Portugal	2006-2021
KL25 (n=9)	ST11	Portugal	2018
KL62 (n=13)	ST348 (n=10); ST45 (n=3)	Portugal	2012-2022
KL64 (n=21)	ST147	Portugal (n=20) Spain	2010-2023
KL102 (n=19)	ST307	Portugal	2018-2020
KL105 (n=11)	ST11	Portugal (n=10) Spain	2006-2020
KL106 (n=10)	ST258	Brazil (n=5) Poland (n=4) Greece	2006-2009
KL107 (n=7)	ST258	Poland (n=2); Italy (n=3); Spain (n=2)	2008-2020
KL112 (n=9)	ST15	Portugal	2010-2017

Table 6. Origin and characteristics of MDR-Kp isolates included in this study for the validation set.

KL type	ST	Origin	Date
KL2 (n=5)	ST14; ST25 (n=3) ST395	Middle East; Europe (n=4)	2014-2024
KL24 (n=3)	ST45 (n=3)	North America; Europe (n=2)	2013-2022
KL25 (n=5)	ST134; ST336; ST5445; ST1459 (n=2)	North America (n=3); Europe (n=2)	2003-2024
KL62 (n=8)	ST17; ST45 (n=5); ST348; ST5446	North America (n=2); Europe (n=6)	2013-2023
KL64 (n=5)	ST147	Africa; Middle East; Europe (n=3)	2015-2023
KL102 (n=5)	ST36 (n=2); ST307 (n=2); ST1583	South America; Europe (n=4)	2012-2024
KL105 (n=3)	ST11	North America; Europe (n=2)	2006-2024
KL107 (n=4)	ST258; ST485; ST512; ST4075	North America (n=2); Africa; Middle East	2013-2020
KL112 (n=2)	ST15; ST481	Africa; Europe	2018-2024

Table 6: Origin and characteristics of MDR-*Kp* isolates included in this study for the validation set.

KL14	ST2202	North America (n=13); Europe (n=5)	2003-2015
KL3	ST1842		
KL10	ST107		
KL17 (n=2)	ST322 ST101		
KL23	ST39		
KL30	ST234		
KL31	ST4833		
KL38	ST37		
KL42	ST309		
KL52	ST2623		
KL63	ST111		
KL110	ST1		
KL113	ST1686		
KL119	ST1966		
KL142	ST4270		
KL148	ST1787		
KL154	ST199		

2.2. Detection of carbapenemase encoding genes

Carbapenemase production was initially inferred by the Blue-Carba test previously described by Pires *et al* (76). For the positive control, we used a previously characterized *K. pneumoniae* KPC-3 producer isolate, and for the negative control, we used *E. coli* ATCC 25922.

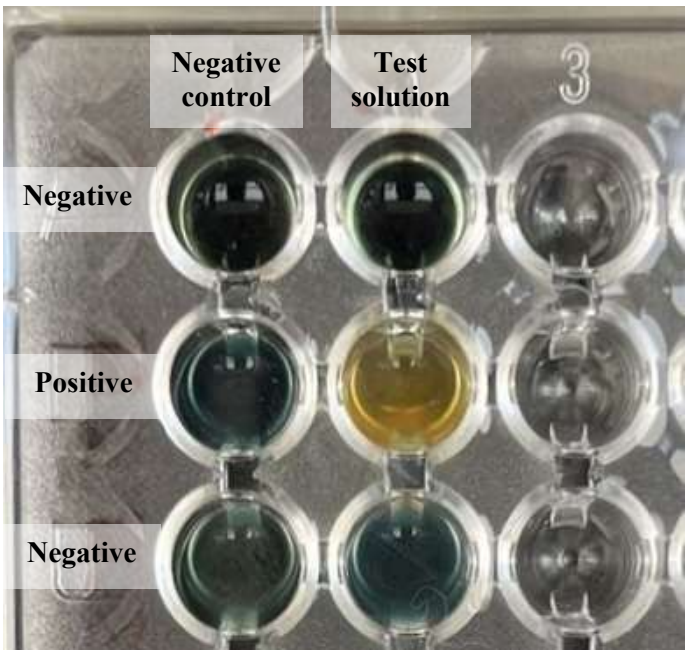


Figure 5. An example of a Blue-Carba test result.

2.3. Molecular techniques

2.3.1. DNA extraction

DNA was extracted by suspending a freshly grown colony in 20 µL of NaOH and SDS, followed by heating at 99 ºC for 15 minutes. Subsequently, 180 µL of Tris-EDTA buffer was added, and the mixture was centrifuged at 13000 rpm for 15 minutes.

2.3.2. DNA amplification and sequencing

Carbapenemase production was confirmed by a multiplex PCR to detect the 5 main

carbapenemase gene families (*bla*$_{KPC}$, *bla*$_{OXA-48}$, *bla*$_{IMP}$, *bla*$_{VIM}$, *bla*$_{NDM}$) and a simplex PCR to detect specifically the *bla*$_{IMP-22}$ variant (Table 7).

Accurate species identification was assessed by PCR of a fragment of 968bp housekeeping gene *hsp60* (Table 7).

After PCR amplification, an electrophoresis 1.5% agarose gel was made with TAE 1X and SYBR® Safe DNA gel stain for visualization of the PCR product. Along with DNA fragments, a weight marker (NZYDNA Ladder VI, NZYtech, Portugal) is loaded and the gel runs at 100 V for 30 minutes. Visualization of the PCR products' integrity was done using a Gel Doc XR+ with Image Lab Software (Biorad, EUA). We then purified the samples to minimize the presence of impurities that could impact the sequencing process, with the kit NZYGelpure (NZYtech, Portugal) according to the manufacturer's instructions. The purified PCR products were sequenced through STAB VIDA (https://www.stabvida.com, Portugal). The obtained sequences were analysed using the Basic Local Alignment Search Tool (BLAST, https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.3.3. *K. pneumoniae* typing

MLST was performed according to the BIGSdb database (https://bigsdb.pasteur.fr), and PCR sequencing of the wzi gene, as previously described (Table 7). The wzi gene is one of six conserved genes of the *cps* cluster, which drives capsule production and attachment to the cell surface. It has been established as a marker for the prediction of the K-type (77).

2.3.4. WGS

WGS was performed on a set of 30 representative isolates of *Ec*SC from the Porto collection (all other isolates from the Madrid and Seville collections had been previously sequenced). The DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA) according to the manufacturer's instructions (78), and the concentration was determined using a Nanodrop instrument (Thermo Scientific Nanodrop 1000 Spectrophotometer). DNA was sequenced using an Illumina Novaseq 6000 (2×300 bp pair-ended runs, ~6 Gb genome, coverage 100×). Data analysis was done using BV BRC software

(https://www.bv-brc.org). The quality and integrity of the sequences were assessed using FastQC utilities of the BV-BRC software. Reads were assembled into contiguous sequences (contigs), which can then be interrogated for gene content and structural variation. These generated files were then uploaded to Pathogen Watch (https://pathogen.watch) and the Center for Genomic Epidemiology website (https://genomicepidemiology.org/services/), to obtain information about resistance genes, precise identification, ST and plasmid content. These analyses provided essential insights into antimicrobial resistance mechanisms, epidemiological lineage assignment, and plasmid content, all of which inform clinical risk, outbreak surveillance, and transmission dynamics.

2.4. Characterization of *K. pneumoniae* extracellular polysaccharides

2.4.1. Sedimentation assay

Sedimentation assay was done to assess bacterial capsule hypermucoviscous phenotype by measuring the resistance to sedimentation after centrifugation, as previously described (79). We started with an overday 1mL LB broth incubation of a fresh colony from a Cled Agar plate. After 6 hours, we dilute it (1:10) into a new Eppendorf with fresh media. The next day, we measured the OD_{600} before and after centrifugation (1,000 g, 5 minutes). Before the 2nd centrifugation, we normalize the OD_{600} to the lowest number obtained from a sample (OD_{min}). The sedimentation value was calculated using the following formula:

$$S = \frac{OD_{600}}{OD_{min}}$$

2.4.2. Capsule extraction and quantification

Capsule extraction was conducted to quantify capsular polysaccharide production utilizing the uronic acid measurement method as previously described (80, 81). We started with an overday 4mL LB broth incubation of a fresh colony from a Cled Agar plate. After 6 hours, we dilute it (1:100) into a new falcon with fresh media. The next day, we measured and adjusted the OD_{600} to 2. We transferred them to a new Eppendorf (500 µL), treated with 100 µL of

Zwittergent, incubated at 56 °C water bath for 20 min and centrifuged (14,000 rpm, 5 min). The supernatant (300 µL) was washed with 1.2mL of cold absolute EtOH and refrigerated on ice for 20 minutes, followed by centrifugation (14,000 rpm, 5 min). Pellet was washed with 70% EtOH, and centrifuged (14,000 rpm, 1 min). After drying the pellet at 56ºC, we resuspend it in 250 µL of double-distilled water and let it hydrate overnight.

The next day, capsule quantification was performed by adding 1.2 mL of 0.0125 M tetraborate, in concentrated H_2SO_4, to 200 µL of our capsular extracts, placed on a 100ºC water bath (5 minutes), and cooled on ice (5 minutes). This solution was incubated with 20 µL of 0.15% m-hydroxybiphenyl in NaOH 0.5%, vortexed (10 minutes) and the OD_{520} was measured. The uronic acid concentration was inferred from a glucuronic acid standard curve.

2.5. FT-IR spectra acquisition

Bacterial isolates were cultured in Mueller-Hinton (Biomérieux, France) and Blood Agar and incubated under standardized conditions (37°C for 18 h). A single bacterial colony was directly applied onto the ATR accessory of the FT-IR instrument (Spectrum 2, PerkinElmer, USA) and allowed to air-dry for 8–10 minutes. Spectral acquisition was performed in triplicate for each strain collected in under five minutes. Spectra were recorded in the range of 4,000 cm^{-1} to 600 cm^{-1}, with a resolution of 4 cm^{-1} and 16 scan co-additions (57, 66).

2.5.1. FT-IR data analysis

Spectral data analysis was conducted using the user-friendly Clover MS Data Analysis software (Clover Bioanalytical Software, Granada, Spain; https://www.clovermsdataanalysis.com). This platform was created to simplify spectral data analysis and machine learning to provide access to users without chemometrics expertise. The analysis workflow was adapted to optimize preprocessing, replicate analysis, and spectral classification for FT-IR-based typing. Spectral data were imported and preprocessed using standard normal variate (SNV) normalization and a Savitzky-Golay filter (window length: 9, polynomial order: 2, derivative order: 2), following previously established methodologies (57, 66).

Table 7. Primers used and amplicon sizes for DNA amplification in this study.

Purpose	Name	Size (bp)	Sequences	Reference
Carbapenemase gene detection	*bla*$_{NDM}$	603	Forward: 5'-ACT TGG CCT TGC TGT CCT T-3' Reverse: 5'-CAT TAG CCG CTG CAT TGA T-3'	(82)
	bla$_{VIM}$	437	Forward: 5'-TGT CCG TGA TGG TGA TGA GT-3' Reverse: 5'-ATT CAG CCA GAT CGG CAT C-3'	
	bla$_{IMP}$	387	Forward: 5'-ACA YGG YTT RGT DGT KCT TG-3' Reverse: 5'-GGT TTA AYA AAR CAA CCA CC-3'	
	bla$_{KPC}$	353	Forward: 5'-TCG CCG TCT AGT TCT GCT GTC TTG-3' Reverse: 5'-ACA GCT CCG CCA CCG TCA T-3'	
	bla$_{OXA-48}$	265	Forward: 5'-ATG CGT GTA TTA GCC TTA TCG-3' Reverse: 5'-CAT CCT TAA CCA CGC CCA AAT C-3'	
	bla$_{IMP-22}$	659	Forward: 5'-TTACTGCCGCAGGAGAGTC-3' Reverse: 5'- TGGGAGCAGGCTGTTAAAGG -3'	(45)
Species identification	*hsp60*	968	Forward: 5'- TCTCCAACATCCGCGAAATG -3' Reverse: 5'- TTACATCATGCCGCCCATNCC -3'	Gonçalves *et al.* unpublished
Typing	*rpoB*	501	Forward: 5'- GTTTTCCCAGTCACGACGTTGTAGGCGAAATGGCWGAGAACCA -3' Reverse: 5'- TTGTGAGCGGATAACAATTTCGAGTCTTCGAAGTTGTAACC -3'	(83)
	gapA	450	Forward: 5'- GTTTTCCCAGTCACGACGTTGTATGAAATATGACTCCACTCACGG -3' Reverse: 5'- TTGTGAGCGGATAACAATTTCCTTCAGAAGCGGCTTTGATGGCTT -3'	

Table 7: Primers used and amplicon sizes for DNA amplification in this study.

	mdh	477	Forward: 5'- GTTTTCCCAGTCACGACGTTGTA CCCAACTCGCTTCAGGTTCAG -3' Reverse: 5'- TTGTGAGCGGATAACAATTTCCCGTTTTTCCCCAGCAGCAG -3'	
	pgi	432	Forward: 5'- GTTTTCCCAGTCACGACGTTGTAGAGAAAAACCTGCCTGTACTGCTGGC -3' Reverse: 5'- TTGTGAGCGGATAACAATTTCCGCGCCACGCTTTATAGCGGTTAAT -3'	
	phoE	420	Forward: 5'- GTTTTCCCAGTCACGACGTTGTAACCTACCGCAACACCGACTTCTTCGG-3' Reverse: 5'- TTGTGAGCGGATAACAATTTCTGATCAGAACTGGTAGGTGAT -3'	
	infB	318	Forward: 5'- GTTTTCCCAGTCACGACGTTGTACTCGCTGCTGGACTATATTCG -3' Reverse: 5'- TTGTGAGCGGATAACAATTTC CGCTTTCAGCTCAAGAACTTC -3'	
	tonB	414	Forward: 5'- GTTTTCCCAGTCACGACGTTGTACTTTATACCTCGGTACATCAGGTT -3' Reverse: 5'- TTGTGAGCGGATAACAATTTCATTCGCCGGCTGRGCRGAGAG -3'	
	wzi	580	Forward: 5'- GTGCCGCGAGCGCTTTCTATCTTGGTATT -3' Reverse: 5'- T GAGAGCCACTGGTTCCAGAAYTTSACCG -3'	(77)

3. RESULTS AND DISCUSSION

3.1. Detection of carbapenemases by molecular routine methods

Carbapenemase production was confirmed in all isolates (2 infection, 2 colonization) by the Blue Carba test. However, multiplex PCR was negative for all main carbapenemase families (*bla*$_{NDM}$, *bla*$_{VIM}$, *bla*$_{IMP}$, *bla*$_{KPC}$, *bla*$_{OXA-48}$). The observation of synergism between ertapenem and EDTA, a metallo-beta-lactamase inhibitor, confirmed the production of a metallo-beta-lactamase. Specific primers were designed (Table 8) to screen for the *bla*$_{IMP-22}$ variant, occasionally reported in our country (84, 85) and the result was positive. Sequencing confirmed the presence of *bla*$_{IMP-22,}$ with 85% aminoacid homology (36 differences) with *bla*$_{IMP-1}$. This divergence might well justify the false negative results obtained by the PCR multiplex and by routine molecular tests (Xpert® Carba R and BD Max™), the latter officially reported to identify common and prototypical variants such as *bla*$_{IMP-1}$ (86).

Infection isolates belonged to sequence type 1079 (ST1079) and capsule typing via *wzi* gene sequencing identified the *wzi*236 allele, corresponding to the KL10. Screening isolates were identified as *Klebsiella variicola*, which also produced IMP-22.

Although relatively uncommon, the ST1079 clone containing both *bla*$_{KPC-3}$ and *bla*$_{IMP-22}$ has been previously reported in Portugal (84), emphasizing the importance of ongoing and accurate epidemiological surveillance, especially in light of the increasing global prevalence of metallo-β-lactamase-producing *Enterobacterales* (87, 88). Therefore, this study confirms that the combination of molecular and cultural methods is crucial to detect rare MBL (*bla*$_{IMP-22}$) and ensure effective diagnosis and early adequate patient treatment. This approach supports infection control measures and prevents underreporting and spread of MDR organisms within healthcare settings.

These findings have been previously reported, and the published article is included in the Appendix.

3.2. Differentiation of *Ec*SC by FT-IR spectroscopy

Initially, for this exploratory study, we evaluated the optimal spectral region and/or combination of regions for species discrimination, the most suitable supervised model for training and species classification, and the software parameters [synthetic minority oversampling technique (SMOTE)] that could provide the highest discriminatory power.

First, the combination of polysaccharide (1200-900 cm^{-1}) and fingerprint (900-700 cm^{-1}) spectral regions showed overall higher balanced accuracy and was subsequently used for analysis. Detailed data on accuracy results for all regions and models are included in the Appendix (Table S1). The polysaccharide region was also reported to be an important region responsible for the discrimination of other species, like *Yersinia enterocolitica* (89), *Campylobacter* spp. (90), *S. enterica* (64) and *K. pneumoniae* (65). The high relevance of the polysaccharide region in *Ec*SC, and among Enterobacteriaceae, could be explained by the presence and diversity of surface polysaccharides (SPs), structures present on the surface of bacterial cells [e.g., capsular polysaccharides (CPS), extracellular polysaccharides (EPS), lipopolysaccharides (LPS)]. To date, little is known regarding the nature and diversity of polysaccharide structures at the surface of *Enterobacter* spp. (91), though their ability to produce EPS is described (92, 93). The fingerprint region has been associated with a variety of different biochemical components, from nucleic acids, proteins, peptidoglycans and membrane lipids and it is expected to reflect inter-strain diversity. In this study, including this region was essential, providing additional discriminatory power.

Secondly, we compared the accuracy of different classification models to differentiate *Ec*SC, PLS-DA, RF and LightGBM, and the latter revealed the highest balanced accuracy (<85%) overall (Table S1). Nevertheless, this accuracy was suboptimal and unbalanced between species since it was high for *E. kobei, E. hoffmannii* and *E. roggenkampii* (74-86%) but particularly low (39%) for *E. asburiae*. This performance may be attributed to class imbalance in the training dataset, where an uneven distribution of isolates among species biases machine learning models towards overrepresented classes, reducing sensitivity for underrepresented species. Consequently, we evaluated whether SMOTE, which generates synthetic spectra for underrepresented species to balance sample distribution across all classes, could enhance model

accuracy. Additionally, model accuracy was evaluated considering all technical replicates independently or as an average.

We compared the performance of 3 LightGBM models: A) 5-species with average technical replicates without SMOTE (Figure 6A), B) 5-species with average technical replicates with SMOTE (Figure 6B), C) 5-species with all technical replicates (Figure 6C).

Comparing Models B and C with Model A, we observed improved classification accuracy. However, it is still possible to observe some overlap between classes. The implementation of SMOTE (Model B) significantly enhanced prediction accuracy (from 68% to 79%), though this performance remains suboptimal for species discrimination. Model C achieved high accuracy levels (99%), suggesting that an increase in the number of isolates per class could improve the model´s accuracy. However, the risk of overfitting needs to be evaluated through validation.

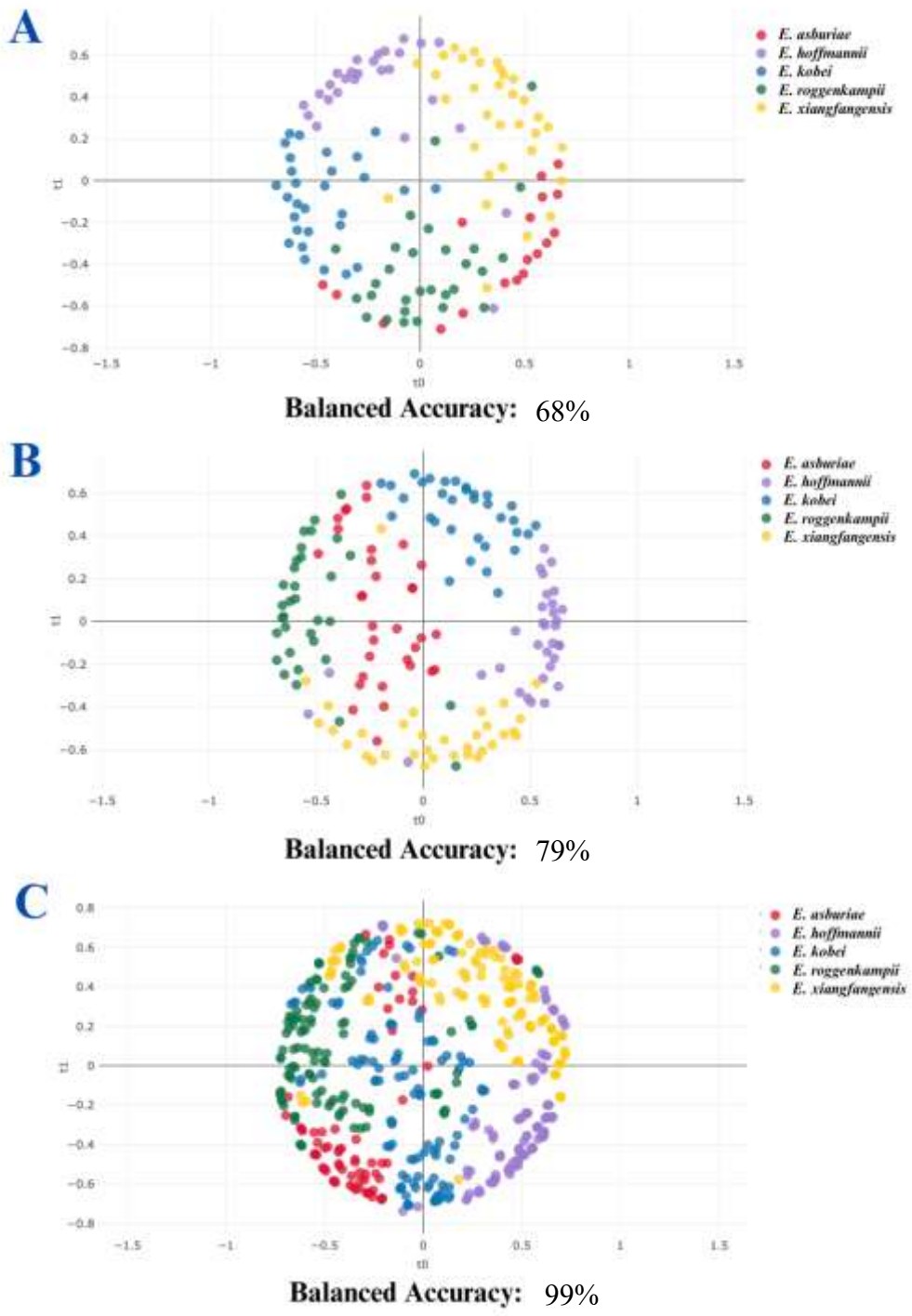


Figure 6. Distance Plot for LightGBM models of the 5 main *Ec*SC. (A) 5-species with average technical replicates without SMOTE (B) 5-species with average technical replicates with SMOTE (C) 5-species with all technical replicates.

The next step was to perform a validation by challenging the three generated models with isolates external to the training dataset. To do this using the same collection of isolates, we built a pipeline in the Clover MS Data Analysis software that automatically and randomly splits the dataset into training and validation sets in an 80-20 proportion (*E. xiangfangensis:* n=8*, E. asburiae:* n=5*, E. roggenkampii:* n=8*, E. hoffmannii:* n=7, *E. kobei:* n=7) for unbiased external validation (Figure 7). This was done for each of the LightGBM models (A, B, C).

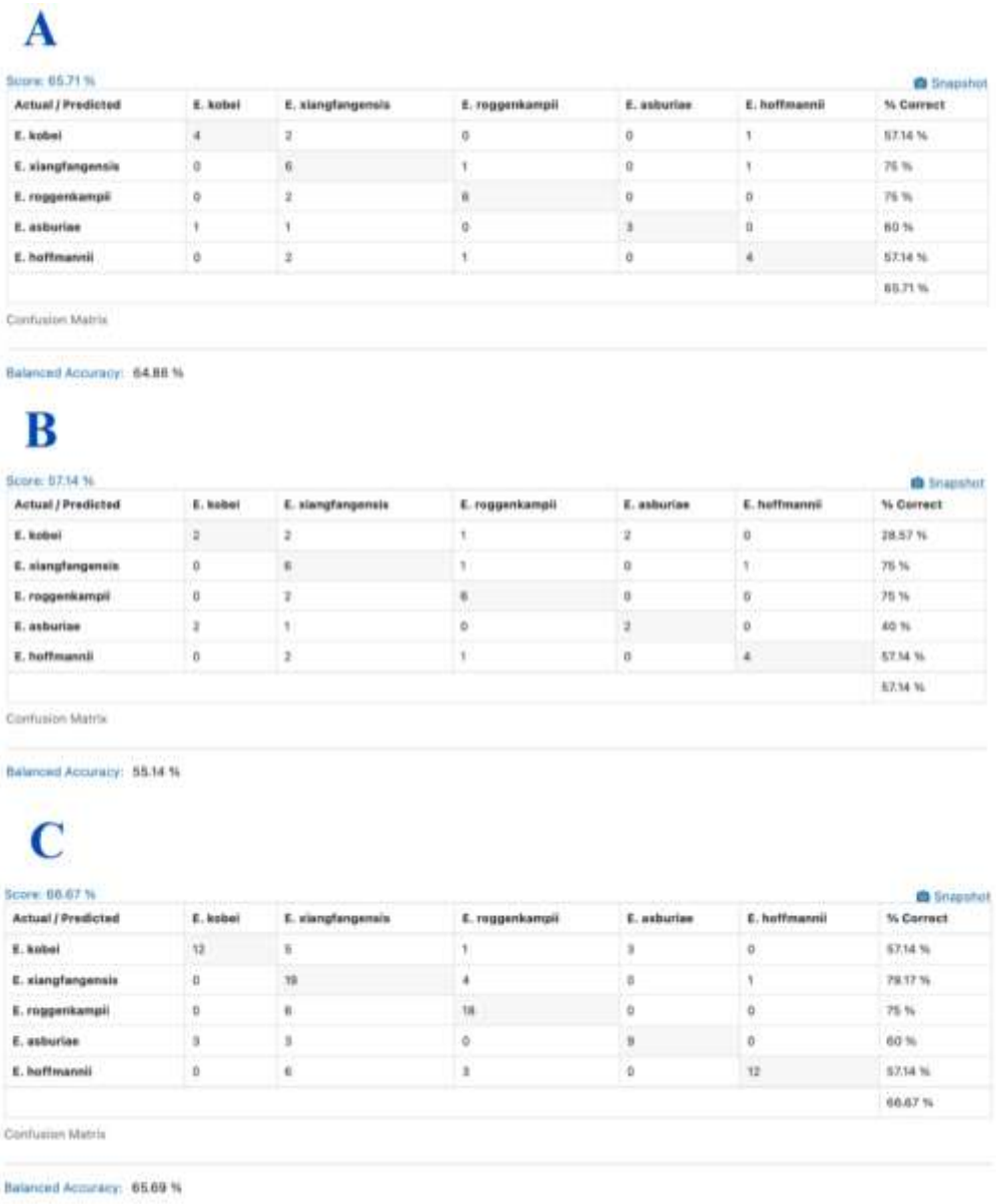


Figure 7. External validation scores of LGBM models of 5 main *Ec*SC. (5-species with average technical replicates without SMOTE (B) 5-species with average technical replicates with SMOTE (C) 5-species with all technical replicates.

The accuracy of Model A was similar to that obtained for internal validation (65%). Model B, which benefited from SMOTE, showed a lower accuracy in validation (decreased from 79% to 55%). As anticipated, Model C, which obtained the highest validation score, experienced the greatest decline in accuracy scores compared to internal validation (decreased from 99% to 66%). Nevertheless, the most clinically relevant *Ec*SC, *E. xiangfangensis*, was the species consistently identified with the highest scores (75-79%) across all models compared to the other species.

Our data suggests that differentiation between the five most clinically relevant *Ec*SC by FT-IR is suboptimal. While Sousa *et al.* (94) successfully discriminated the *A. calcoaceticus—A. baumannii* complex, and Nitrosetein *et al.* (95) successfully discriminated between *E. faecium* and *E. faecalis*, these results indicate that the limitations encountered in the *Ec*SC are not a general constraint of FT-IR for closely related species. Rather, they appear to reflect specific evolutionary and genomic dynamics within this complex that need to be further explored. It is possible that horizontal gene transfer between SP loci has blurred differences between the species (91), but this needs to be confirmed by extensive genomic comparison throughout the genus.

It is also of interest to show that FT-IR outperformed MALDI-TOF MS for species-level identification within the *Ec*SC. We assessed the performance of MALDI-TOF MS against the reference method (WGS) across all five targeted species, using data from the Spanish hospitals, where data for both methods were available. The highest identification rate was observed for *E. kobei* with 80%, followed by *E. asburiae* with 44% and *E. hoffmannii* with 10%. Both *E. xiangfangensis* and *E. roggenkampii* yielded 0. Overall, the identification performance for all five species was 17%. The substantial underperformance of MALDI-TOF MS supports the need for alternative identification methods with greater resolution. Although our FT-IR spectroscopy models did not yield fully satisfactory results for building a robust five-species model as initially anticipated, it outperformed MALDI-TOF MS for most species, with the exception of *E. kobei*.

E. xiangfangensis was consistently identified with the highest scores and it represents the most common species in clinical species (54%) (53). Further optimization of models could improve it's detection, offering a valuable foundation and a noteworthy advancement toward establishing a more robust routine identification method for this species.

3.3. Identification of globally disseminated Hv-*Kp* and MDR-*Kp*

3.3.1. Differentiation of Hv-*Kp* clones by FT-IR spectroscopy

FT-IR spectroscopy was initially tested for differentiation of the three main KL types of Hv-*Kp*: KL1, KL2, and KL5. As expected, the spectral differences among the isolates were more significantly discriminatory in the polysaccharides region (1200-900 cm^{-1}). In this model, PLS-DA was used to perfectly discriminate the 3 clusters corresponding to the 3 Hv-*Kp* KL types, with 100% of total correct KL type predictions (Figure 8). Thus, we show that the most common Hv-*Kp* KL types can be perfectly recognized by FT-IR spectroscopy.

Figure 8. PLS-DA model of 3 typical Hv-*Kp* KL types. (A) Score plot of PLS-DA regression model according to KL types using five components. (B) Confusion matrix for PLS-DA model.

3.3.2. Discrimination of MDR-*Kp* and Hv-*Kp* strains sharing the same KL2 type

FT-IR spectroscopy differentiation is based on KL type composition, and KL types are relevant epidemiological markers of both Hv-*Kp* and MDR-*Kp*. However, a few KL types such as KL2, are shared between the two populations, though they are associated with different sublineages within MDR-*Kp* (ST14) and Hv-*Kp* (ST65, ST86, ST300) (71, 96). Therefore, to test the ability of FT-IR spectroscopy to differentiate between the two pathotypes in isolates sharing the same KL type, we focused our analysis on isolates sharing the same KL2 type. Our PLS-DA model was able to clearly distinguish two clusters belonging to KL2-MDR-*Kp* and KL2-Hv-*Kp* with a balanced accuracy of 94% average correct predictions (Figure 9), when using a combination of the protein (1700-1500 cm^{-1}) and lipid (3000-2800 cm^{-1}) regions.

Figure 9. PLS-DA model of KL2 for MDR-*Kp* and Hv-*Kp*. (A) Score plot of PLS-DA regression model using six components. (B) Confusion matrix for PLS-DA model.

3.3.3. Differentiation of MDR-*Kp* from Hv-*Kp* populations

Based on these results and using the same discriminatory regions identified above (protein (1700-1500 cm^{-1}) and lipid (3000-2800 cm^{-1}) regions), we assessed FT-IR's potential to discriminate MDR-*Kp* from Hv-*Kp* populations using 322 MDR-*Kp* and 76 Hv-*Kp* isolates from our collection. FT-IR spectroscopy was able to successfully cluster MDR-*Kp* from Hv-*Kp* with a balanced accuracy of 96% average correct predictions (Figure 10), independently of each category's diversity.

Figure 10. PLS-DA model of MDR-*Kp* vs Hv-*Kp* populations. (A) Score plot of PLS-DA regression model according to each sublineage, using six components. (B) Confusion matrix for PLS-DA model.

The accuracy of this model was further validated using an external set (n=27) of representative MDR-*Kp* (n=20), Hv-*Kp* (n=7), which included KL types included and not

included in the model. We achieved 100% of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), as all isolates were accurately predicted. These excellent prediction results suggest that this model could be applied in routine settings to putatively identify any Hv-*Kp* isolate. Although the surface discriminatory features on the basis for the discrimination between the two pathotypes are unknown, we could hypothesize that the presence and/or accessibility of specific surface proteins associated with virulence and/or antibiotic resistance mechanisms might differ between them (65).

Additionally, we tested 9 putative convergent MDR-Hv-*Kp* strains and all but one MDR-Hv-*Kp* were predicted to be Hv-*Kp* (Table 8).

Table 8. External validation of the MDR-*Kp* vs Hv-*Kp* populations PLS-DA model.

	KL type	Class Prediction	Accuracy (%)
MDR-*Kp*	KL2	MDR	75
	KL3	MDR	70
	KL10	MDR	81
	KL17 (n=2)	MDR	88-96
	KL23	MDR	61
	KL24	MDR	100
	KL25	MDR	63
	KL30	MDR	91
	KL31	MDR	90
	KL38	MDR	100
	KL62	MDR	92
	KL64	MDR	82
	KL102	MDR	84
	KL105	MDR	99
	KL107 (n=2)	MDR	94-100
	KL112	MDR	87
	KL142	MDR	85
	KL148	MDR	100
Hv-*Kp*	KL2	HV	79
	KL20-ST420 (n=2)	HV	75-86
	KL30	HV	67
	KL54-ST1109	HV	93
	KL54-ST29	HV	58
	KL57	HV	77
MDR-Hv-*Kp*	KL64 (n=2)	HV	70-77
	KL20	HV	68
	KL51	HV	62
	KL15	HV	80
	KL1	HV	68
	KL43	MDR	82
	KL62 (n=2)	HV	76-77

Hv-*Kp* and some convergent MDR-Hv-*Kp* are typically associated with increased capsule production most frequently together with a hypermucoviscous (HMV) phenotype, both linked to the acquisition of the *rmpABCD locus* (97). To test the influence of capsule overproduction on the FT-IR discriminatory profiles observed, we performed tests to quantify capsule production (capsule quantification through uronic acid content) and assess hypermucoviscosity phenotype (sedimentation assay) for the 9 MDR-Hv-*Kp.* The results obtained are presented in Table 9.

Table 9. Sedimentation assay and capsule quantification results for MDR-Hv-*Kp*.

	KL64 (n=2)	KL20	KL51	KL15	KL1	KL43	KL62 (n=2)
Sedimentation assay	0.37-**0.39**	**0.39**	0.29	0.25	0.29	0.29	**0.38-0.43**
Capsular quantification	5.62-**7.33**	**9.68**	6.97	**7.78**	5.07	**9.49**	4.79-5.76

(Isolates possibly hypermucoviscous and/or with overproduction of capsule are highlighted in bold.)

Higher sedimentation values correspond to increased hypermucoviscosity, since HMV isolates pellet less effectively, resulting in a more opaque supernatant and consequently higher OD_{600} readings. Although the data do not follow a linear scale, comparison with previous studies using these methods (98-100) (Ferrador *et al.* unpublished data) suggests that four isolates are likely HMV, while another four potentially overproduce capsule. These properties do not correlate and are not uniformly present across isolates, indicating that identification of convergent MDR-Hv-*Kp* strains depends on molecular features beyond CPS quantity or hypermucoviscosity.

This model provides the first rapid and reliable method for promptly identifying Hv-*Kp* strains in clinical settings, addressing a critical gap, as no current methods can accurately detect them, as most often, these strains remain unidentified. MDR-Hv-*Kp* strains, which are both

highly invasive and resistant to treatment, pose a significant dual threat requiring urgent and aggressive intervention, similar to Hv-*Kp*. Since the model predicts all such strains as Hv-*Kp* at initial screening, none are missed.

While our collection includes diverse KL types, it lacks prominent representatives from these convergent *Kp* strains (ST11, ST258, and ST512) (71, 101). Nonetheless, we do not expect the model's performance to differ from other MDR-Hv-*Kp* strains, as it consistently identified highly diverse clones as Hv-*Kp*, indicating that identification is independent of clonal type.

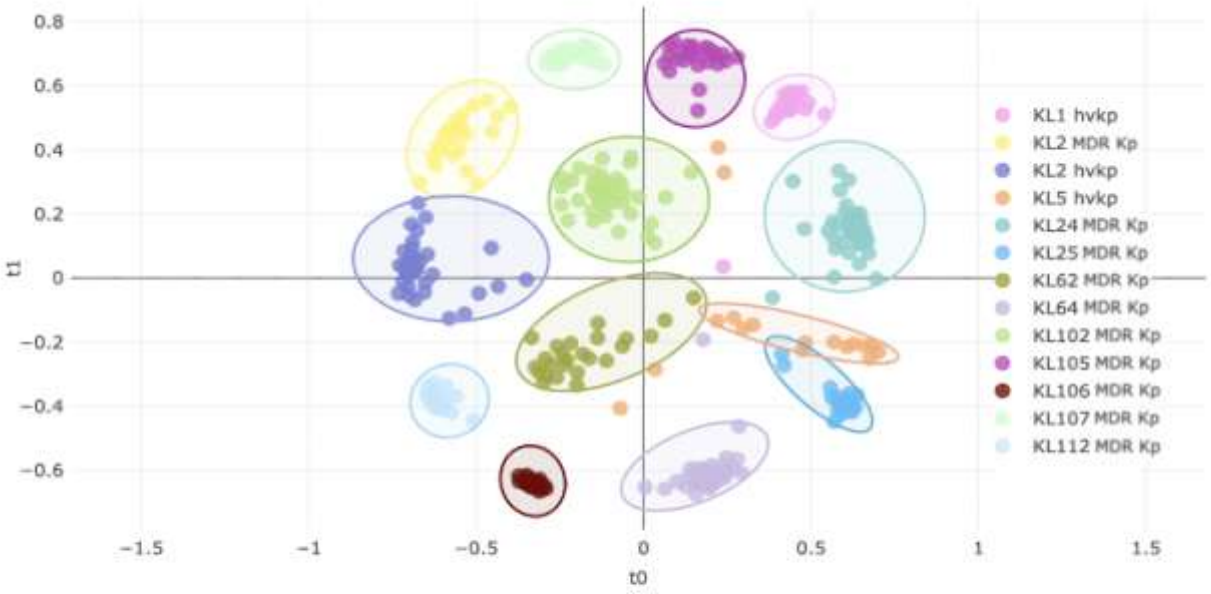
3.3.4. Capsular differentiation of the Top 10 MDR-*Kp* and 3 Hv-*Kp* KL types

Based on previous results, we assessed the feasibility and accuracy of a universal classification model for the identification of globally disseminated MDR-*Kp* and Hv-*Kp* strains. We built a Random Forest model, able to discriminate isolates belonging to the global clinically relevant Top 10 MDR-*Kp* and Top 3 Hv-*Kp*, achieving a balanced accuracy of 97% (88-100% for the different KL types) of average correct predictions, when using the entire spectrum range (4000-600 cm^{-1}) (Figure 11). Moreover, no significant differences were observed between the two different media used.

Validation was performed using an external set of isolates that are representative of diverse KL types, regardless of whether they were included in the model. The positive controls from the validation set (n = 40) represented 9 out of the 13 KL types included in the model. The negative controls of the set (n = 10) belonged to 10 KL types other than those included in the RF model. Most (n = 36) were correctly identified by the model, with a prediction score higher than 30%, similar to the cutoff previously established for MDR-*Kp* (57). Four isolates were misidentified [KL2 (n=2), KL107 (n=2)]. This means that the model generated no false positives (100% specificity, NPV and PPV) but failed to classify some true positives (lower sensitivity, 90%), thereby slightly reducing the overall accuracy to 90%. In clinical settings or an outbreak situation, higher sensitivity is generally more critical to ensure cases are not missed, even if it means accepting a few false positives, as early detection and containment are essential.

Thus, this study provides an universal model for identification of the most frequent MDR-*Kp* and Hv-*Kp* sublineages worldwide and has thus the potential to be used in different contexts and settings to quickly identify high-risk *K. pneumoniae* strains. It is imperative to validate it in real clinical contexts to evaluate its performance as a routine method.

(A)



(B)

Leave one out cross validation

Actual / Predicted	KL1 hvkp	KL102	KL105	KL106	KL107	KL112	KL2	KL2 hvkp	KL24	KL25	KL5 hvkp	KL62	KL64	% Correct
KL1 hvkp	20	0	0	0	0	0	0	0	0	0	0	0	0	100 %
KL102	0	40	0	0	0	0	0	0	0	0	0	0	0	100 %
KL105	0	0	26	0	0	0	0	0	0	0	0	0	0	100 %
KL106	0	0	0	26	0	0	0	0	0	0	0	0	0	100 %
KL107	0	0	0	0	23	0	0	0	0	0	0	0	0	100 %
KL112	0	0	0	0	0	22	0	0	0	0	0	0	0	100 %
KL2	0	1	1	0	0	0	24	1	0	0	0	0	0	88.89 %
KL2 hvkp	0	2	0	0	0	0	2	35	0	0	0	1	0	87.5 %
KL24	0	0	0	0	0	0	0	0	55	0	0	0	0	100 %
KL25	0	0	0	0	0	0	0	0	0	21	0	1	0	95.45 %
KL5 hvkp	0	0	0	0	0	0	0	0	0	0	16	1	0	94.12 %
KL62	0	1	0	0	0	0	0	0	0	0	0	24	0	96 %
KL64	0	1	0	0	0	0	0	0	0	0	0	0	55	98.21 %
														96.99 %

Confusion Matrix

Balanced Accuracy: 96.94 %

Figure 11. Random Forest model of Top 10 MDR-*Kp* and 3 Hv-*Kp* KL types. (A) Distance plot of the Random Forest regression model according to the KL types defined by genomic data. (B) Confusion matrix for the Random Forest model.

4. CONCLUSIONS AND FUTURE PERSPECTIVES

This thesis addressed diagnostic challenges associated with the detection of clinically relevant pathogens and antimicrobial resistance mechanisms within routine clinical microbiology laboratories. It explored and unveiled limitations of current routine methods and demonstrated that integration between different methods is key to supporting accurate diagnosis and validating new approaches. Moreover, it expanded knowledge on FT-IR-based differentiation at species and subspecies levels, highlighting its potential as a rapid, simple, and reliable diagnostic tool to facilitate prompt and precise treatment decisions, reduce transmission of multidrug-resistant pathogens, and enhance patient care in clinical practice.

Regarding **specific conclusions**, the following points emerge:

- There are critical diagnostic limitations in detecting particular MBL ($bla_{\text{IMP-22}}$) by routine methods. These findings emphasized the importance of integrating phenotypic, genotypic and molecular data for an accurate detection of carbapenemases, especially rare or emerging variants that may bypass routine surveillance systems.

- This study revealed limitations of FT-IR spectroscopy for routine identification of the five clinically relevant species within the *Ec*SC, though it outperformed MALDI-TOF MS. Its potential might be improved with a larger collection of isolates together with a comprehensive correlation with the molecular basis for discrimination and/or developing specialized models for identification of *E. xiangfangensis*.

- This study provides the first rapid and reliable approach to identify Hv-*Kp* and MDR-Hv-*Kp* strains in clinical settings, filling a crucial gap in routine microbiology capacity. Based on this knowledge, it extends the application of FT-IR to the accurate identification of clinically relevant MDR-*Kp* and HV-*Kp* sublineages worldwide, of high value in infection control and public health surveillance.

The FT-IR-based models developed in this thesis need to be validated in real-life contexts, thereby ensuring their reproducibility and feasibility for routine diagnostic use.

FT-IR technology demonstrated strong potential for reliable species identification and typing across diverse bacteria. Given the robustness and richness of spectral data, its applicability could be expanded in the future towards the detection of antimicrobial resistance markers and other clinically relevant traits.

5. REFERENCES

1.European Centre for Disease Prevention and Control. 2024. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals, 2022–2023 Stockholm. ECDC. [Available from: https://www.ecdc.europa.eu/sites/default/files/documents/healthcare-associated-point-prevalence-survey-acute-care-hospitals-2022-2023.pdf.

2.Szabó S, Feier B, Capatina D, Tertis M, Cristea C, Popa A. An Overview of Healthcare Associated Infections and Their Detection Methods Caused by Pathogen Bacteria in Romania and Europe. Journal of Clinical Medicine. 2022;11(11):3204.

3.Szabó S, Feier B, Capatina D, Tertis M, Cristea C, Popa A. An Overview of Healthcare Associated Infections and Their Detection Methods Caused by Pathogen Bacteria in Romania and Europe. J Clin Med. 2022;11(11).

4.Organization WHO. Report on the Burden of Endemic Health Care-Associated Infection Worldwide. 2011. [Available from: https://iris.who.int/bitstream/handle/10665/80135/9789241501507_eng.pdf?sequence=1

5.WaterAid. The impact of healthcare-associated infections on patients and hospitals. 2024. [Available from: https://washmatters.wateraid.org/sites/g/files/jkxoof256/files/2024-09/The-impact-of-healthcare-associated-infections-on-hospitals-and-patients-in-low-and-middle-income-countries.pdf

6.Bass LM, de Meireles LHF, Kiriyama EJ, Dos Santos NO, de Sousa AHF, Silva K, et al. Cost savings of a nationwide project preventing healthcare-associated infections in adult, paediatric and neonatal critical care settings in Brazil: a micro-costing study. BMJ Open. 2025;15(4):e097515.

7.Zimlichman E, Henderson D, Tamir O, Franz C, Song P, Yamin CK, et al. Health Care–

Associated Infections: A Meta-analysis of Costs and Financial Impact on the US Health Care System. JAMA Internal Medicine. 2013;173(22):2039-46.

8.Hutton G, Chase C, Kennedy-Walker R, Hamilton H. Financial and economic costs of healthcare-associated infections in Africa. Journal of Hospital Infection. 2024;150:1-8.

9.Morris AJ, Hensen M, Graves N, Cai Y, Wolkewitz M, Roberts SA, et al. The burden of healthcare-associated infections in New Zealand public hospitals 2021. Infection Control & Hospital Epidemiology. 2024;45(10):1176-82.

10.Organization WHO. Global report on infection prevention and control 2024. 2024. [Available from: https://www.who.int/publications/i/item/9789240103986

11.Organization WHO. WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. 2024. [Available from: https://iris.who.int/bitstream/handle/10665/376776/9789240093461-eng.pdf?sequence=1

12.Rice LB. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. The Journal of Infectious Diseases. 2008;197(8):1079-81.

13.Darby EM, Trampari E, Siasat P, Gaya MS, Alav I, Webber MA, et al. Molecular mechanisms of antibiotic resistance revisited. Nature Reviews Microbiology. 2023;21(5):280-95.

14.Cantón R, Ruiz-Garbajosa P. Co-resistance: an opportunity for the bacteria and resistance genes. Curr Opin Pharmacol. 2011;11(5):477-85.

15.Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clinical Microbiology

and Infection. 2012;18(3):268-81.

16.O'NEILL J. Tackling drug-resistant infections globally: Final report and recommendations. 2016. [Available from: https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

17.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. Nature Microbiology. 2019;4(11):1919-29.

18.Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet. 2022;399(10325):629-55.

19.Kaprou GD, Bergšpica I, Alexa EA, Alvarez-Ordóñez A, Prieto M. Rapid Methods for Antimicrobial Resistance Diagnostics. Antibiotics (Basel). 2021;10(2).

20.Organization WHO. Development and implementation of national action plans for infection prevention and control. World Health Organization; 2025. [Available from: https://iris.who.int/bitstream/handle/10665/381706/9789240111943-eng.pdf?sequence=1

21.UNICEF. AMR: The Urgent Threat of Drug-Resistant Infections. 2023. [Available from: https://www.unicef.org/media/144266/file/The-Urgent-Threat-of-Drug-Resistant-Infections:-A-UNICEF-Guidance-Note-on-Antimicrobial-Resistance-2023.pdf

22.Elbehiry A, Marzouk E, Abalkhail A, Abdelsalam MH, Mostafa MEA, Alasiri M, et al. Detection of antimicrobial resistance via state-of-the-art technologies versus conventional methods. Front Microbiol. 2025;16:1549044.

23.Hassall J, Coxon C, Patel VC, Goldenberg SD, Sergaki C. Limitations of current techniques in clinical antimicrobial resistance diagnosis: examples and future prospects. npj Antimicrobials

and Resistance. 2024;2(1):16.

24.March-Rosselló GA. Rapid methods for detection of bacterial resistance to antibiotics. Enfermedades Infecciosas y Microbiología Clínica (English Edition). 2017;35(3):182-8.

25.Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for rapid diagnosis of tuberculosis and detection of rifampin resistance in pulmonary and extrapulmonary specimens. J Clin Microbiol. 2011;49(12):4138-41.

26.Alsuof EA, Alsayed AR, Zraikat MS, Khader HA, Hasoun LZ, Zihlif M, et al. Molecular Detection of Antibiotic Resistance Genes Using Respiratory Sample from Pneumonia Patients. Antibiotics. 2025;14(5):502.

27.Cunningham Scott A, Limbago B, Traczewski M, Anderson K, Hackel M, Hindler J, et al. Multicenter Performance Assessment of Carba NP Test. Journal of Clinical Microbiology. 2017;55(6):1954-60.

28.Lutgring JD, Zhu W, de Man TJB, Avillan JJ, Anderson KF, Lonsway DR, et al. Phenotypic and Genotypic Characterization of Enterobacteriaceae Producing Oxacillinase-48-Like Carbapenemases, United States. Emerg Infect Dis. 2018;24(4):700-9.

29.Notake S, Matsuda M, Tamai K, Yanagisawa H, Hiramatsu K, Kikuchi K. Detection of IMP metallo-β-lactamase in carbapenem-nonsusceptible Enterobacteriaceae and non-glucose-fermenting Gram-negative rods by immunochromatography assay. J Clin Microbiol. 2013;51(6):1762-8.

30.Patidar N, Vyas N, Sharma S, Sharma B. Phenotypic Detection of Carbapenemase Production in Carbapenem-Resistant Enterobacteriaceae by Modified Hodge Test and Modified Strip Carba NP Test. J Lab Physicians. 2021;13(1):14-21.

31.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;60(3):e0224521.

32.Antonelli A, Giani T, Di Pilato V, Riccobono E, Perriello G, Mencacci A, et al. KPC-31 expressed in a ceftazidime/avibactam-resistant *Klebsiella pneumoniae* is associated with relevant detection issues. Journal of Antimicrobial Chemotherapy. 2019;74(8):2464-6.

33.Arbefeville SS, Timbrook TT, Garner CD. Evolving strategies in microbe identification-a comprehensive review of biochemical, MALDI-TOF MS and molecular testing methods. J Antimicrob Chemother. 2024;79(Supplement_1):i2-i8.

34.Żukowska M. Advanced methods of bacteriological identification in a clinical microbiology laboratory. Journal of Pre-Clinical and Clinical Research. 2021;15. 10.26444/jpccr/134646

35.Vrioni G, Tsiamis C, Oikonomidis G, Theodoridou K, Kapsimali V, Tsakris A. MALDI-TOF mass spectrometry technology for detecting biomarkers of antimicrobial resistance: current achievements and future perspectives. Annals of Translational Medicine. 2018;6(12):240.

36.van Belkum A, Chatellier S, Girard V, Pincus D, Deol P, Dunne Jr WM. Progress in proteomics for clinical microbiology: MALDI-TOF MS for microbial species identification and more. Expert Review of Proteomics. 2015;12(6):595-605.

37.Joyanes P, del Carmen Conejo Ma, Martínez-Martínez L, Perea Evelio J. Evaluation of the VITEK 2 System for the Identification and Susceptibility Testing of Three Species of Nonfermenting Gram-Negative Rods Frequently Isolated from Clinical Samples. Journal of Clinical Microbiology. 2001;39(9):3247-53.

38.Ling TK, Liu ZK, Cheng AF. Evaluation of the VITEK 2 system for rapid direct identification and susceptibility testing of gram-negative bacilli from positive blood cultures. J Clin Microbiol. 2003;41(10):4705-7.

39.Ligozzi M, Bernini C, Bonora MG, De Fatima M, Zuliani J, Fontana R. Evaluation of the VITEK 2 system for identification and antimicrobial susceptibility testing of medically relevant gram-positive cocci. J Clin Microbiol. 2002;40(5):1681-6.

40.Guo L, Ye L, Zhao Q, Ma Y, Yang J, Luo Y. Comparative study of MALDI-TOF MS and VITEK 2 in bacteria identification. J Thorac Dis. 2014;6(5):534-8.

41.Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6:791.

42.Rychert J. Benefits and Limitations of MALDI-TOF Mass Spectrometry for the Identification of Microorganisms. Journal of Infectiology. 2019;2:1-5.

43.Berry GJ, Loeffelholz MJ, Williams-Bouyer N. An investigation into laboratory misidentification of a bloodstream *Klebsiella variicola* infection. Journal of Clinical Microbiology. 2015;53(8):2793-4.

44.Pavlovic M, Konrad R, Iwobi AN, Sing A, Busch U, Huber I. A dual approach employing MALDI-TOF MS and real-time PCR for fast species identification within the Enterobacter cloacae complex. FEMS Microbiology Letters. 2012;328(1):46-53.

45.Martins R, Read MA, Neves I, Peixe L, Alves V, Novais Â. Diagnostic challenges in detecting $bla_{\text{IMP-22}}$ carbapenemase in *Klebsiella pneumoniae*. Diagnostic Microbiology and Infectious Disease. 2025;112(2):116772.

46.Bernaschi P, Del Chierico F, Petrucca A, Argentieri A, Atti MCD, Ciliento G, et al. Microbial Tracking of Multidrug-Resistant *Klebsiella Pneumoniae* Isolates in a Pediatric Hospital Setting. International Journal of Immunopathology and Pharmacology. 2013;26(2):463-72.

47.Rodrigues C, Novais Â, Sousa C, Ramos H, Coque TM, Cantón R, et al. Elucidating

constraints for differentiation of major human *Klebsiella pneumoniae* clones using MALDI-TOF MS. European Journal of Clinical Microbiology & Infectious Diseases. 2017;36(2):379-86.

48.Rodrigues C, Passet V, Rakotondrasoa A, Brisse S. Identification of *Klebsiella pneumoniae*, *Klebsiella quasipneumoniae, Klebsiella variicola* and Related Phylogroups by MALDI-TOF Mass Spectrometry. Frontiers in Microbiology. 2018;Volume 9 - 2018.

49.De Florio L, Riva E, Giona A, Dedej E, Fogolari M, Cella E, et al. MALDI-TOF MS Identification and Clustering Applied to *Enterobacter* Species in Nosocomial Setting. Frontiers in Microbiology. 2018;Volume 9 - 2018.

50.Godmer A, Benzerara Y, Normand Anne C, Veziris N, Gallah S, Eckert C, et al. Revisiting Species Identification within the *Enterobacter cloacae* Complex by Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. Microbiology Spectrum. 2021;9(1):10.1128/spectrum.00661-21.

51.Davin-Regli A, Lavigne JP, Pagès JM. *Enterobacter* spp.: Update on Taxonomy, Clinical Aspects, and Emerging Antimicrobial Resistance. Clin Microbiol Rev. 2019;32(4).

52.Ganbold M, Seo J, Wi YM, Kwon KT, Ko KS. Species identification, antibiotic resistance, and virulence in *Enterobacter cloacae* complex clinical isolates from South Korea. Frontiers in Microbiology. 2023;Volume 14 - 2023.

53.Wu W, Feng Y, Zong Z. Precise Species Identification for *Enterobacter:* a Genome Sequence-Based Study with Reporting of Two Novel Species, *Enterobacter quasiroggenkampii* sp. nov. and *Enterobacter quasimori* sp. nov. mSystems. 2020;5(4):10.1128/msystems.00527-20.

54.Ramadan AA. Bacterial typing methods from past to present: A comprehensive overview. Gene Reports. 2022;29:101675.

55.Novais Â, Freitas AR, Rodrigues C, Peixe L. Fourier transform infrared spectroscopy: unlocking fundamentals and prospects for bacterial strain typing. European Journal of Clinical Microbiology & Infectious Diseases. 2019;38(3):427-48.

56.Cunha ÉL. Application of FT -IR Spectroscopy for Strain Typing in Clinical Microbiology [Monograph]: Faculty of Pharmacy, University of Porto, Portugal; 2024.

57.Novais Â, Gonçalves Ana B, Ribeiro Teresa G, Freitas Ana R, Méndez G, Mancera L, et al. Development and validation of a quick, automated, and reproducible ATR FT-IR spectroscopy machine-learning model for *Klebsiella pneumoniae* typing. Journal of Clinical Microbiology. 2024;62(2):e01211-23.

58.Yang H, Shi H, Feng B, Wang L, Chen L, Alvarez-Ordóñez A, et al. Protocol for bacterial typing using Fourier transform infrared spectroscopy. STAR Protocols. 2023;4(2):102223.

59.Lasch P. Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging. Chemometrics and Intelligent Laboratory Systems. 2012;117:100–14.

60.Park S, Ryoo N. Comparative analysis of IR-Biotyper, MLST, cgMLST, and WGS for clustering of vancomycin-resistant *Enterococcus faecium* in a neonatal intensive care unit. Microbiol Spectr. 2024;12(4):e0411923.

61.Sousa C, Novais Â, Magalhães A, Lopes J, Peixe L. Diverse high-risk B2 and D Escherichia coli clones depicted by Fourier Transform Infrared Spectroscopy. Scientific Reports. 2013;3(1):3278.

62.Novais Â, Ferraz RV, Viana M, da Costa PM, Peixe L. NDM-1 Introduction in Portugal through a ST11 KL105 *Klebsiella pneumoniae* Widespread in Europe. Antibiotics. 2022;11(1):92.

63.Silva L, Grosso F, Rodrigues C, Ksiezarek M, Ramos H, Peixe L. The success of particular *Acinetobacter baumannii* clones: accumulating resistance and virulence inside a sugary shield. Journal of Antimicrobial Chemotherapy. 2020;76(2):305-11.

64.Campos J, Sousa C, Mourão J, Lopes J, Antunes P, Peixe L. Discrimination of non-typhoid *Salmonella* serogroups and serotypes by Fourier Transform Infrared Spectroscopy: A comprehensive analysis. International Journal of Food Microbiology. 2018;285:34-41.

65.Rodrigues C, Sousa C, Lopes João A, Novais Â, Peixe L. A Front Line on *Klebsiella pneumoniae* Capsular Polysaccharide Knowledge: Fourier Transform Infrared Spectroscopy as an Accurate and Fast Typing Tool. mSystems. 2020;5(2):10.1128/msystems.00386-19.

66.Gonçalves AB, Alves V, Neves I, Read A, Pinheiro N, Henius AE, et al. Real-time FT-IR typing of *Klebsiella pneumoniae*: a flexible and rapid approach for outbreak detection and infection control. J Antimicrob Chemother. 2025.

67.Paczosa MK, Mecsas J. Klebsiella pneumoniae: Going on the Offense with a Strong Defense. Microbiol Mol Biol Rev. 2016;80(3):629-61.

68.Nagendra D, Chaudhuri S, Gupta N, Shanbhag V, Eshwara VK, Rao S, et al. Prevalence, Risk Factors, and Clinical Outcomes of Hypervirulent *Klebsiella pneumoniae* Strains among *Klebsiella pneumoniae* Infections: A Systematic Review and Meta-analysis. Indian J Crit Care Med. 2025;29(4):370-93.

69.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. Nature Communications. 2021;12(1):4188.

70.Kochan TJ, Nozick SH, Valdes A, Mitra SD, Cheung BH, Lebrun-Corbin M, et al. *Klebsiella pneumoniae* clinical isolates with features of both multidrug-resistance and hypervirulence have

unexpectedly low virulence. Nature Communications. 2023;14(1):7962.

71.Lan P, Jiang Y, Zhou J, Yu Y. A global perspective on the convergence of hypervirulence and carbapenem resistance in *Klebsiella pneumoniae*. Journal of Global Antimicrobial Resistance. 2021;25:26-34.

72.Russo Thomas A, Olson R, Fang C-T, Stoesser N, Miller M, MacDonald U, et al. Identification of Biomarkers for Differentiation of Hypervirulent *Klebsiella pneumoniae* from Classical *K. pneumoniae*. Journal of Clinical Microbiology. 2018;56(9):10.1128/jcm.00776-18.

73.He Y, Xiao Y, Feng Y, Wu S, Wei L, Zong Z. Two novel Enterobacter species, *Enterobacter chinensis* sp. nov. and *Enterobacter rongchengensis* sp. nov., recovered from clinical samples carrying multiple virulence factors. Microbiology Spectrum. 2024;12(8):e00292-24.

74.Ikhimiukor OO, Zac Soligno NI, Akintayo IJ, Marcovici MM, Souza SSR, Workman A, et al. Clonal background and routes of plasmid transmission underlie antimicrobial resistance features of bloodstream *Klebsiella pneumoniae*. Nature Communications. 2024;15(1):6969.

75.Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. Nature Reviews Microbiology. 2020;18(6):344-59.

76.Pires J, Novais Â, Peixe L. Blue-Carba, an Easy Biochemical Test for Detection of Diverse Carbapenemase Producers Directly from Bacterial Cultures. Journal of Clinical Microbiology. 2013;51(12):4281-3.

77.Brisse S, Passet V, Haugaard Anita B, Babosan A, Kassis-Chikhani N, Struve C, et al. wzi Gene Sequencing, a Rapid Method for Determination of Capsular Type for *Klebsiella* Strains. Journal of Clinical Microbiology. 2013;51(12):4073-8.

78. Promega. Wizard® Genomic DNA Purification Kit. 2023. [Available from: https://ita.promega.com/-/media/files/resources/protocols/technical-manuals/0/wizard-

genomic-dna-purification-kit-protocol.pdf

79.Khadka S, Ring BE, Pariseau DA, Mike LA. Characterization of *Klebsiella pneumoniae* Extracellular Polysaccharides. Curr Protoc. 2023;3(11):e937.

80.Domenico P, Schwartz S, Cunha BA. Reduction of capsular polysaccharide production in *Klebsiella pneumoniae* by sodium salicylate. Infection and Immunity. 1989;57(12):3778-82.

81.Blumenkrantz N, Asboe-Hansen G. New method for quantitative determination of uronic acids. Analytical Biochemistry. 1973;54(2):484-9.

82.Bogaerts P, Rezende de Castro R, de Mendonça R, Huang T-D, Denis O, Glupczynski Y. Validation of carbapenemase and extended-spectrum β-lactamase multiplex endpoint PCR assays according to ISO 15189. Journal of Antimicrobial Chemotherapy. 2013;68(7):1576-82.

83.Diancourt L, Passet V, Verhoef J, Grimont Patrick AD, Brisse S. Multilocus Sequence Typing of *Klebsiella pneumoniae* Nosocomial Isolates. Journal of Clinical Microbiology. 2005;43(8):4178-82.

84.Gonçalves D, Cecílio P, Faustino A, Iglesias C, Branca F, Estrada A, et al. Intra- and Extra-Hospital Dissemination of IMP-22-Producing *Klebsiella pneumoniae* in Northern Portugal: The Breach of the Hospital Frontier Toward the Community. Frontiers in Microbiology. 2021;12.

85.Gonçalves AB, Alves V, Neves I, Read A, Peixe L, Novais Â. Ceftazidime-avibactam resistance in *Klebsiella pneumoniae*: a growing public health concern across diverse sublineages, 2019–2024. Journal of Global Antimicrobial Resistance. 2025.

86.Zhang Z, Wang D, Li Y, Liu Y, Qin X. Comparison of the Performance of Phenotypic Methods for the Detection of Carbapenem-Resistant Enterobacteriaceae (CRE) in Clinical Practice. Frontiers in Cellular and Infection Microbiology. 2022;Volume 12 - 2022.

87.Grundmann H, Glasner C, Albiger B, Aanensen DM, Tomlinson CT, Andrasević AT, et al. Occurrence of carbapenemase-producing *Klebsiella pneumoniae* and *Escherichia coli* in the European survey of carbapenemase-producing Enterobacteriaceae (EuSCAPE): a prospective, multinational study. The Lancet Infectious Diseases. 2017;17(2):153-63.

88.Sader HS, Mendes RE, Carvalhaes CG, Kimbrough JH, Castanheira M. Changing Epidemiology of Carbapenemases Among Carbapenem-Resistant Enterobacterales From United States Hospitals and the Activity of Aztreonam-Avibactam Against Contemporary Enterobacterales (2019–2021). Open Forum Infectious Diseases. 2023;10(2).

89.Kuhm Andrea E, Suter D, Felleisen R, Rau J. Identification of *Yersinia enterocolitica* at the Species and Subspecies Levels by Fourier Transform Infrared Spectroscopy. Applied and Environmental Microbiology. 2009;75(18):5809-13.

90.Mouwen DJM, Weijtens MJBM, Capita R, Alonso-Calleja C, Prieto M. Discrimination of Enterobacterial Repetitive Intergenic Consensus PCR Types of *Campylobacter coli* and *Campylobacter jejuni* by Fourier Transform Infrared Spectroscopy. Applied and Environmental Microbiology. 2005;71(8):4318-24.

91.Holt KE, Lassalle F, Wyres KL, Wick R, Mostowy RJ. Diversity and evolution of surface polysaccharide synthesis loci in Enterobacteriales. The ISME Journal. 2020;14(7):1713-30.

92.Dhanya BE, Athmika, Rekha PD. Characterization of an exopolysaccharide produced by *Enterobacter* sp. YU16-RN5 and its potential to alleviate cadmium induced cytotoxicity in vitro. 3 Biotech. 2021;11(12):491.

93.Xie Y, Ye Z, Wan X, Deng H, Sun W, He X, et al. Screening of exopolysaccharide-producing *Enterobacter aerogenes* NJ1023 and its cadaverine biosynthesis promotion. Frontiers in Microbiology. 2023;Volume 14 - 2023.

94.Sousa C, Silva L, Grosso F, Nemec A, Lopes J, Peixe L. Discrimination of the *Acinetobacter*

calcoaceticus–Acinetobacter baumannii complex species by Fourier transform infrared spectroscopy. European Journal of Clinical Microbiology & Infectious Diseases. 2014;33(8):1345-53.

95.Nitrosetein T, Wongwattanakul M, Chonanant C, Leelayuwat C, Charoensri N, Jearanaikoon P, et al. Attenuated Total Reflection Fourier Transform Infrared Spectroscopy combined with chemometric modelling for the classification of clinically relevant *Enterococci*. Journal of Applied Microbiology. 2021;130(3):982-93.

96.Spadar A, Phelan J, Elias R, Modesto A, Caneiras C, Marques C, et al. Genomic epidemiological analysis of *Klebsiella pneumoniae* from Portuguese hospitals reveals insights into circulating antimicrobial resistance. Scientific Reports. 2022;12(1):13791.

97.Russo Thomas A, Marr Candace M. Hypervirulent Klebsiella pneumoniae. Clinical Microbiology Reviews. 2019;32(3):10.1128/cmr.00001-19.

98.Tan YH, Chen Y, Chu WHW, Sham L-T, Gan Y-H. Cell envelope defects of different capsule-null mutants in K1 hypervirulent *Klebsiella pneumoniae* can affect bacterial pathogenesis. Molecular Microbiology. 2020;113(5):889-905.

99.He J, Shi Q, Chen Z, Zhang W, Lan P, Xu Q, et al. Opposite evolution of pathogenicity driven by in vivo wzc and wcaJ mutations in ST11-KL64 carbapenem-resistant *Klebsiella pneumoniae*. Drug Resist Updat. 2023;66:100891.

100.Mike LA, Stark AJ, Forsyth VS, Vornhagen J, Smith SN, Bachman MA, et al. A systematic analysis of hypermucoviscosity and capsule reveals distinct and overlapping genes that impact *Klebsiella pneumoniae* fitness. PLoS Pathog. 2021;17(3):e1009376.

101.Arcari G, Carattoli A. Global spread and evolutionary convergence of multidrug-resistant and hypervirulent *Klebsiella pneumoniae* high-risk clones. Pathog Glob Health. 2023;117(4):328-41.

6. APPENDIX

Diagnostic Microbiology & Infectious Disease 112 (2025) 116772

Contents lists available at ScienceDirect

Diagnostic Microbiology & Infectious Disease

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

Diagnostic challenges in detecting bla_{IMP-22} carbapenemase in *Klebsiella pneumoniae*

Sir,

Carbapenem-resistant *Klebsiella pneumoniae* (CR-KP) prevalence in Portuguese hospitals has risen sharply, with one of Europe's highest rates [1,2], mainly due to KPC or OXA-48 and rarely metallo-β-lactamases (MBL) [3,4]. Here, we report challenges identifying bla_{IMP-22}-harboring *K. pneumoniae* and expose limitations in routine molecular detection methods. A two-month-old infant, delivered via C-section on June 19, 2024 was readmitted on August 22, 2024, after two visits to another hospital's pediatric emergency room. Diagnosed with a urinary tract infection, the infant received cefuroxime for 10 days (5 days IV, 5 days oral). Initial uroculture showed susceptible *Proteus mirabilis* and *Enterococcus faecalis*. On September 3, a *K. pneumoniae* resistant to cefotaxime, ceftazidime-avibactam, amoxicillin-clavulanic acid, cefoxitin, cefepime and ertapenem was identified in the urine culture. The immunochromatographic test (O.K.N.V.I. RESIST-5, Coris BioConcept) detected IMP carbapenemase, while Xpert® Carba R (Cepheid, USA) and BD Max™ CPO (BD, USA) were negative. Blue-Carba [5] confirmed carbapenemase production while a multiplex PCR for carbapenemases was negative [6]. Hence, a PCR using specific primers designed in this study: IMP22-FW (5′-ATTACTGCCGCAG GAGAGTC-3′) and IMP22-RV (5′-TGGGAGCAGGCTGTTAAAGG-3′) identified bla_{IMP-22}, previously described in Portugal [4]. Our results highlight a lack of sensitivity for molecular routine screening methods to identify bla_{IMP} variants other than bla_{IMP-1} [7]. In fact, bla_{IMP-22} is only 86 % homologous to bla_{IMP-1} and failure to identify such rarer and divergent variants might well happen in other PCR-based systems, complicating diagnosis in laboratories with limited molecular capacity. A second urine culture on September 10 identified *K. pneumoniae* additionally resistant to meropenem and imipenem.

K. pneumoniae strains from urine were characterized by MLST (https://bigsdb.pasteur.fr/klebsiella/primers-used/) and sequence type 1079 (ST1079) was identified. *wzi* sequencing revealed *wzi*236 allele and capsular genotype KL10. Though rarely, this clone (with bla_{KPC-3} and bla_{IMP-22}) had been previously reported in the country reinforcing the need for accurate surveillance [4], especially with increasing MBL in the USA and some European countries [8,9]. The infant improved with cefuroxime treatment, and *K. pneumoniae* presence was considered asymptomatic bacteriuria. Rectal screening on August 22 was negative, while two subsequent screenings (September 10 and 17) yielded *Klebsiella* sp. isolates after pre-enrichment step with an imipenem disk in Shaedler broth. These isolates were initially identified as *K. pneumoniae* complex due to discrepancies in species identification (*K. variicola* by VITEK MS and *K. pneumoniae* by MALDI-TOF MS), as reported [10]. PCR sequencing of MLST housekeeping genes confirmed a *K. variicola* carrying bla_{IMP-22}. Source investigation through rectal colonization screening of the parents, both nurses at the hospital, and the five-year-old brother on September 24 by enrichment culture yielded negative results, ruling out family members as sources. Epidemiological investigations revealed that the mother's one-month hospitalization during pregnancy in a high-risk unit sharing facilities with women with diverse ethnic backgrounds potentially contributed to pathogen acquisition.

This study highlights limitations in routine molecular methods for detection of particular MBL (bla_{IMP-22}) risking bla_{IMP} variants underreporting. Finally, combining molecular and cultural methods is crucial for effective diagnosis and surveillance of infections by multidrug-resistant pathogens.

CRediT authorship contribution statement

Rita Martins: Writing – original draft, Methodology, Investigation. **Maria Antónia Read:** Writing – review & editing, Resources. **Isabel Neves:** Writing – review & editing, Resources. **Luísa Peixe:** Writing – review & editing, Funding acquisition. **Valquíria Alves:** Writing – review & editing, Methodology, Conceptualization. **Ângela Novais:** Writing – review & editing, Validation, Supervision, Conceptualization.

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.

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia – FCT, I.P., through the Research Unit on Applied Molecular Biosciences – UCIBIO (project numbers UIDP/04378/2020 and UIDB/04378/2020) and the Associate Laboratory Institute for Health and Bioeconomy – I4HB (project number LA/P/0140/2020). R.M. received financial support from Bill and Melinda Gates Foundation through Grand Challenges (reference INV060596) and Â.N. from Fundação para a Ciência e a Tecnologia – FCT, I.P. through the Scientific Employment Stimulus to Â. N. (DOI: 10.54499/2021.02252.CEECIND/CP1662/CT0009).

References

[1] Control, E.C.f.D.P.a., Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report 2023. 2024: Stockholm: ECDC.
[2] Organization GWH. WHO bacterial priority pathogens list, 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. WHO Bacterial Priority Pathogens List; 2024.
[3] Gorgulho A, et al. Carbapenemase-producing Enterobacteriaceae in a Portuguese hospital - a five-year retrospective study. Germs 2020;10(2):95–103.

https://doi.org/10.1016/j.diagmicrobio.2025.116772
Received 16 December 2024; Received in revised form 15 January 2025; Accepted 24 February 2025
Available online 25 February 2025
0732-8893/© 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

[4] Gonçalves D, et al. Intra- and extra-hospital dissemination of IMP-22-producing *Klebsiella pneumoniae* in Northern Portugal: the breach of the hospital frontier toward the community. Front Microbiol 2021;12:777054.
[5] Pires J, Novais Â, Peixe L. Blue-carba, an easy biochemical test for detection of diverse carbapenemase producers directly from bacterial cultures. J Clin Microbiol 2015;51(12):4281–3.
[6] Bogaerts P, et al. Validation of carbapenemase and extended-spectrum β-lactamase multiplex endpoint PCR assays according to ISO 15189. J Antimicrob Chemother 2013;68(7):1576–82.
[7] Zhang Z, et al. Comparison of the performance of phenotypic methods for the detection of carbapenem-resistant *Enterobacteriaceae* (CRE) in clinical practice. Front Cell Infect Microbiol 2022:12.
[8] Grundmann H, et al. Occurrence of carbapenemase-producing Klebsiella pneumoniae and *Escherichia coli* in the European survey of carbapenemase-producing *Enterobacteriaceae* (EuSCAPE): a prospective, multinational study. Lancet Infect Dis 2017;17(2):153–63.
[9] Sader HS, et al. Changing epidemiology of carbapenemases among carbapenem-resistant enterobacterales from United States hospitals and the activity of Aztreonam-Avibactam against contemporary enterobacterales (2019–2021). Open Forum Infect Dis 2023;10(3). p. ofad046.
[10] Berry GJ, Loeffelholz MJ, Williams-Bouyer N. An investigation into laboratory misidentification of a bloodstream *Klebsiella variicola* infection. J Clin Microbiol 2015;53(8):2793–4.

Rita Martins[a], Maria Antónia Read[b], Isabel Neves[b,c], Luísa Peixe[a], Valquíria Alves[b,c], Ângela Novais[a,*]

[a] *UCIBIO, i4HB, Faculdade de Farmácia, Universidade do Porto, 4050-313 Porto, Portugal*

[b] *Clinical Microbiology Laboratory, Local Healthcare Unit, Matosinhos, Portugal*

[c] *Infection Control Unit, Local Healthcare Unit, Matosinhos, Portugal*

[*] Corresponding author.
E-mail address: asmorim@ff.up.pt (Â. Novais).

Table S1. Detailed results on balanced accuracy and individual accuracy for each species, each spectral region and for each statistical model.

		RF							PLSDA							LGBM						
		E. asburiae	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy
Whole spectra	*E. asburiae*	6	0	5	3	4	33.33 %	65.45	24	2	2	2	3	72.73 %	67.8	5	0	4	6	3	27.78 %	63.26
	E. hoffmannii	1	17	2	2	6	62.96 %		2	21	6	0	0	72.41 %		1	16	1	1	6	66.67 %	
	E. kobei	0	0	23	4	3	79.31 %		1	0	24	2	2	82.76 %		0	0	24	3	2	82.76 %	
	E. roggenkampii	0	1	4	22	2	75.86 %		3	4	1	18	1	66.67 %		0	2	4	21	2	72.41 %	
	E. xiangfangensis	2	1	1	4	25	75.76 %		0	5	5	0	8	44.44 %		4	2	2	3	22	66.67 %	
Proteins	*E. asburiae*	5	3	5	1	4	27.78 %	48,13	22	2	4	3	2	66.67 %	55,15	6	2	5	2	3	33.33 %	51,53
	E. hoffmannii	1	11	4	4	7	40.74 %		8	18	3	3	1	58.17 %		1	13	5	4	4	48.15 %	
	E. kobei	2	4	11	5	7	37.93 %		3	2	21	1	2	72.41 %		1	8	9	7	4	31.03 %	
	E. roggenkampii	2	2	4	19	2	65.52 %		3	2	4	16	2	59.26 %		2	1	3	21	2	72.41 %	
	E. xiangfangensis	1	5	2	2	23	69.7 %		4	4	2	4	4	22.22 %		0	2	5	2	24	72.73 %	
Lipids	*E. asburiae*	3	3	6	2	4	16.67 %	47,46	15	7	3	6	2	45.45 %	42,36	6	2	4	2	4	33.33 %	52,59
	E. hoffmannii	3	14	4	1	5	51.85 %		8	6	11	5	1	20.69 %		1	14	6	2	4	51.85 %	
	E. kobei	1	5	15	5	3	51.72 %		4	4	17	4	0	58.62 %		3	5	13	4	4	44.83 %	
	E. roggenkampii	4	0	4	19	2	65.52 %		2	2	2	19	2	70.37 %		4	1	0	22	2	75.86 %	
	E. xiangfangensis	1	5	6	3	17	51.52 %		4	7	1	3	3	16.67 %		2	8	1	3	19	57.58 %	
Nucleic acids	*E. asburiae*	8	1	7	1	3	33.33 %	52,69	19	2	8	4	2	57.58 %	55,97	9	0	5	1	2	0,0	54,34
	E. hoffmannii	2	14	3	2	7	51.85 %		3	15	5	4	2	51.72 %		2	15	1	2	7	55.56 %	
	E. kobei	3	4	12	5	5	41.38 %		1	3	21	1	3	72.41 %		1	4	14	5	5	48.28 %	
	E. roggenkampii	1	0	2	23	3	79.31 %		2	3	2	19	1	70.37 %		1	3	3	21	1	72.41 %	
	E. xiangfangensis	3	4	3	4	15	57.58 %		5	6	1	1	5	27.78 %		1	6	5	6	15	45.45 %	
Carbohydrates	*E. asburiae*	9	0	3	2	4	0.5	64,78	20	2	6	2	3	60.61 %	65,85	10	1	1	4	2	55.56 %	67,21
	E. hoffmannii	1	18	2	1	5	66.67 %		2	21	5	1	0	72.41 %		1	17	2	3	4	62.96 %	
	E. kobei	2	2	20	5	0	68.97 %		2	3	22	2	0	75.86 %		1	0	22	6	0	75.86 %	
	E. roggenkampii	0	0	4	19	6	65.52 %		3	2	2	19	1	70.37 %		3	0	2	20	4	68.97 %	
	E. xiangfangensis	3	2	0	5	24	72.73 %		4	3	2	0	8	0.5		3	3	1	2	24	72.73 %	

Table S1: Detailed results on balanced accuracy and individual accuracy for each species, each spectral region and for each statistical model.

		RF							PLSDA							LGBM						
		E. asburiae	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy
Fingerprint	*E. asburiae*	7	0	1	5	5	38.89 %		20	3	5	4	1	60.61 %		8	0	3	8	3	33.33 %	
	E. hoffmannii	1	18	2	0	8	66.67 %		5	21	3	0	0	72.41 %		1	21	2	0	3	77.78 %	
	E. kobei	1	0	24	3	1	82.76 %	**69,45**	1	2	24	0	2	82.76 %	**66,12**	4	0	23	1	1	79.31 %	**64,77**
	E. roggenkampii	1	0	2	25	1	86.21 %		5	2	0	19	1	70.37 %		4	0	1	22	2	75.86 %	
	E. xiangfangensis	1	1	3	4	24	72.73 %		3	3	4	0	8	44.44 %		4	5	3	2	19	57.58 %	

		RF							PLSDA							LGBM						
		E. asburiae	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy	*E. asburiae*	*E. hoffmannii*	*E. kobei*	*E. roggenkampii*	*E. xiangfangensis*	% Correct	Balanced accuracy
Carbohydrates + Fingerprint	*E. asburiae*	9	0	2	2	5	0.5		25	2	2	2	2	75.76 %		7	1	1	4	5	38.89 %	
	E. hoffmannii	1	19	2	1	4	70.37 %		1	24	3	0	1	82.76 %		1	20	1	1	4	74.07 %	
	E. kobei	0	0	24	4	1	82.76 %	**71,56**	2	1	24	2	0	82.76 %	**70,85**	1	0	25	3	0	86.21 %	**67,82**
	E. roggenkampii	1	1	2	22	3	75.86 %		3	2	1	20	1	74.07 %		1	0	1	23	4	79.31 %	
	E. xiangfangensis	2	1	2	2	26	78.79 %		4	2	5	0	7	38.89 %		3	4	3	3	20	60.61 %	