

Rapid Detection of VREfm Clusters: FTIR Spectroscopy as a Practical Alternative to Whole-Genome Sequencing

Jun Hao Wang-Wang,^{1,2} Laia Soler,^{1,3} Elisa Martró,^{1,2,3,4,5} Gemma Clarà,¹ Jessica Hidalgo,¹ Irma Casas,⁵ María-José García-Quesada,^{6,7} Montserrat Giménez,^{1,8,9} Verónica Saludes,^{1,2,3,4,5,9} Antoni E. Bordoy,^{1,3,9} and Pere-Joan Cardona^{1,2,3,8,9,10}

¹Microbiology Department, Laboratori Clínic Metropolitana Nord, Germans Trias i Pujol University Hospital, Badalona, Spain, ²Genetics and Microbiology Department, Universitat Autònoma de Barcelona, Bellaterra, Spain, ³Clinical and Experimental Microbiology, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain, ⁴Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Instituto de Salud Carlos III (ISCIII), Madrid, Spain, ⁵Preventive Medicine Department, Germans Trias i Pujol University Hospital, Badalona, Spain, ⁶Infection Control Nursing, Infection Control Team, Germans Trias i Pujol University Hospital, Badalona, Spain, ⁷NURECARE-IGTP Nursing Research Group, Germans Trias i Pujol Research Institute, Badalona, Spain, and ⁸Centro de Investigación Biomédica en Red en Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

Vancomycin-resistant *Enterococcus faecium* (VREfm) has become a significant nosocomial pathogen due to its potential to cause outbreaks. Whole-genome sequencing (WGS) is considered the reference method for determining genomic relatedness among outbreak strains, but its routine use in clinical microbiology laboratories remains challenging. Consequently, faster and simpler typing methods are needed. Fourier transform infrared spectroscopy (FTIR) captures the unique infrared fingerprint of each isolate, enabling the comparison of spectral profiles to infer genomic relatedness. In this study, we evaluated the performance of FTIR for identifying genomic clusters of VREfm in a tertiary hospital, in comparison with three WGS-based methods: core-genome multilocus sequence typing (cgMLST), core-genome single nucleotide polymorphism analysis (cgSNP), and split *k*-mer analysis (SKA). A total of 87 VREfm isolates, collected between April 2020 and October 2023, were typed using both FTIR and WGS. Among these, 56 were associated with three outbreaks in the surgery, nephrology, and oncohematology units, according to conventional epidemiology. Concordance between typing methods was assessed using the Adjusted Rand index (AR) and Adjusted Wallace coefficient (AW). All three WGS-based methods yielded similar clustering results and revealed one monoclonal and two polyclonal outbreaks. Using cgMLST as the reference, an optimal FTIR cutoff range of 0.210–0.227 was determined. FTIR clustering results showed strong concordance with WGS-based methods; however, concordance with SKA was slightly lower. These findings suggest that FTIR provides clustering information comparable to WGS-based methods, providing a rapid and practical alternative to support timely infection control measures during VREfm outbreaks.

Keywords. cgSNP; FTIR; outbreak; SKA; vancomycin-resistant *enterococcus faecium*.

Infections caused by vancomycin-resistant *Enterococcus faecium* (VREfm) pose a considerable challenge due to multidrug resistance [1] and their capacity to spread in hospital environments, where they can cause outbreaks [2] and exacerbate the already significant burden of healthcare-associated infections [3]. In 1988, the first vancomycin-resistant *E. faecium* isolates were reported independently in France and the United Kingdom, marking the emergence of this resistance mechanism as a public health concern [4]. In Europe, the proportion

of VREfm infections rose from 12.3% (2016) to 19.8% (2023), increasing mortality and costs [5]. The rise in VREfm cases has been particularly noteworthy in Germany [6], where the proportion of VREfm cases reached a peak of 26.3% in 2019 but subsequently declined to 12.7% in 2023 [7]. Conversely, the percentage of VREfm in Spain remained below 3% from 2016 to 2022 [8]. Particularly, in the hospitals of Catalonia, there were no reported cases of VREfm until 2020 when the first cases were detected shortly after the beginning of the COVID-19 pandemic. Currently, the VREfm resistance proportion remains around 3% in Catalonia. At Hospital Germans Trias i Pujol the proportions were 8.5% in 2020, 10.0% in 2021, and 1.0% in 2023.

The 2020–2021 increase in our hospital prompted the evaluation of faster typing tools for outbreak management. DNA-based methods such as pulsed-field gel electrophoresis (PFGE) and whole-genome sequencing (WGS) have been widely used [9, 10], being WGS the current gold standard due to its high discriminatory power [11]. However, routine implementation of WGS remains limited by high turnaround times primarily caused by bioinformatic analysis [12]. A lack

Received 08 September 2025; accepted 17 November 2025; published online 24 November 2025

⁹Verónica Saludes, Antoni E. Bordoy and Pere-Joan Cardona contributed equally to this work. Author order was determined on the basis of correspondence role and experience.

Correspondence: Pere-Joan Cardona, MD, PhD, Microbiology Department, Laboratori Clínic Metropolitana Nord, Hospital Universitari Germans Trias i Pujol, Ctra. del Canyet s/n, Badalona, Catalonia 08916, Spain (pcardonai.germantrias@gencat.cat; pj.cardona@gmail.com).

Open Forum Infectious Diseases[®]

© The Author(s) 2025. Published by Oxford University Press on behalf of Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

<https://doi.org/10.1093/ofid/ofaf718>

of a gold standard WGS analysis approach for VREfm outbreak delimitation also exists. Core genome multilocus sequence typing (cgMLST) has been widely used due to its standardized and stable nomenclature and the availability of commercial software tools [13]. Single nucleotide polymorphism (SNP) distances generated from core genome alignments are also usually used for outbreak delimitation. In this case, some studies have proposed that recombination masking may be inaccurate for highly recombinant species such as *E. faecium*, as it can underestimate the number of SNPs and artificially cluster isolates more closely than they really are [14]. Alternatively, split *k*-mer analysis (SKA) may offer greater discriminatory power than cgMLST for VREfm outbreak delimitation, though current evidence is limited [12, 15]. For these methods, determining the adequate clustering threshold to define close genomic relatedness can also be an important limitation. Therefore, selecting an appropriate WGS-based analysis method for VREfm outbreak analysis requires careful consideration.

Fourier-transform infrared (FTIR) spectroscopy has emerged as a phenotypic alternative with resolution suitable for typing and outbreak delimitation [16]. FTIR quantifies absorption by carbohydrates, lipids, and proteins [17, 18]. Each isolate's infrared fingerprint is then compared with the rest so that relationships between bacterial lineages can be established based on spectrum analysis [19]. Previous work supports FTIR for Gram-positive [2] and Gram-negative pathogens with internal validation [20–23].

The aim of the present study was to evaluate the capacity of IR Biotype (Bruker GmbH, Leipzig, Germany) to track emergent VREfm genomic clusters at a tertiary hospital in comparison with three different WGS analysis methods.

MATERIAL AND METHODS

Study Design and Setting

Retrospective cross-sectional study in a tertiary care hospital serving 200 000 inhabitants and reference for 1 200 000 in Barcelona's Northern Metropolitan Area (Spain). We analyzed 87 VREfm isolates from hospitalized patients ($n = 83$) and four environmental samples from the surgery department, collected April 2020–October 2023. Fifty-six isolates were associated with three epidemiologically defined outbreaks (surgery, nephrology, oncohematology). A nosocomial outbreak was declared by the infection control team upon the detection of ≥ 2 clinical or ≥ 3 rectal swab isolates within one month in a single ward. Only isolates from patients who were VREfm negative at the time of admission were considered. Thus, all included isolates were detected after the first 48 hours of hospitalization by routine screening. The date of outbreak closure was set when no VREfm isolates were identified in the same hospital ward for 3 months since the last

identification. The remaining isolates were considered unrelated to these outbreaks (Supplementary Table 1).

Ethics Statement

This study was approved by the Clinical Research Ethics Committee (CEIC) of the University Hospital Germans Trias i Pujol in Barcelona, Spain (PI-24-118).

Routine Microbiological Diagnostics and Antimicrobial Susceptibility Testing

All isolates were firstly identified with matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS, Bruker Daltonik GmbH, Bremen, Germany). The antibiotic susceptibility of vancomycin and teicoplanin was assessed using a gradient test (bioMérieux SA, Marcy-l'Étoile, France), while linezolid susceptibility was evaluated through disk diffusion (Bio-Rad Laboratories Inc., California, United States). For blood cultures, susceptibility to vancomycin, teicoplanin, and linezolid was determined using the VITEK-2 Compact system (bioMérieux SA, Marcy-l'Étoile, France). Interpretation of minimum inhibitory concentration (MIC) values and disk diffusion zone diameters followed the criteria established by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) [24]. All isolates were subsequently preserved in Cryoinstant® Natural storage medium (Scharlab S.L., Barcelona, Spain) at -80°C until further analyses.

Sample Preparation for FTIR Analysis and Spectrum Analysis

All VREfm isolates were thawed at room temperature and cultured in Columbia Agar + 5% sheep blood (bioMérieux SA, Marcy-l'Étoile, France) for 24 hours at 37°C . Then, bacterial isolates were subcultured in Mueller-Hinton agar medium from a single colony (Becton Dickinson GmbH, Heidelberg, Germany) for another 24 hours at 37°C before testing. Preparation of bacterial suspensions was performed as previously described by Wang-Wang *et al.* [23]. Briefly, a loopful of bacterial cells was resuspended in 50 μL of deionized water in a 1.5-mL suspension vial containing metal beads (Bruker GmbH, Leipzig, Germany) and homogenized by vortexing. Subsequently, 50 μL of 70% (v/v) ethanol was added to each vial, followed by a second homogenization step. Then, 15 μL of each bacterial suspension was spotted in quadruplicate onto a 96-spot silicon plate (Bruker GmbH, Leipzig, Germany). For quality control, 12 μL of two infrared test standards (IRTS 1 and IRTS 2, each containing an *Escherichia coli* strain with a defined reference spectrum) were placed in duplicate on the same silicon plate. The plate was then dried at 37°C for 30 minutes prior to insertion into the Biotype system. For spectrum acquisition, all samples were analyzed in transmission mode using the default analysis settings (wavelength region 1300–800 cm^{-1}).

Samples that did not meet the manufacturer's quality criteria were excluded. The acceptance parameters were as follows: absorption between 0.4 and 2, noise ($\times 10^{-6}$) < 300 , signal-to-noise ratio R2 > 200 , signal-to-noise ratio R3 > 40 , water vapor ($\times 10^{-6}$) < 300 , signal-to-water ratio (R2) > 100 , signal-to-water ratio (R3) > 20 , and fringes ($\times 10^{-6}$) < 100 . Based on these criteria, 19 samples required reanalysis until at least three valid spectra were obtained for each isolate. Of these, 13 yielded valid spectra, whereas no valid spectra could be recovered for the remaining six samples, which were therefore excluded from the study. An average spectrum for each individual strain was subsequently generated from the qualified spectra using the OPUS 8.2.28 software (Bruker GmbH, Leipzig, Germany).

The resulting average spectra were then used to construct a dendrogram based on hierarchical cluster analysis (HCA) using the Euclidian metric distance and the average linkage method (UPGMA) without dimensionality reduction techniques. The optimal clustering cutoff within a slightly extended range (0.15–0.25) of that recommended by the manufacturer (0.15–0.20) was calculated. Isolates showing an FTIR spectral distance equal to or below this cut-off value were classified within the same FTIR cluster, whereas those exceeding the threshold were considered FTIR singletons.

Sample Preparation for WGS Analysis and Bioinformatics Analyses

For WGS analysis, all VREfm samples were prepared and sequenced as previously described [23], except for solid sample culture that was performed in Columbia CNA agar (bioMérieux SA, Marcy-l'Étoile, France). The WGS clustering analysis of VREfm isolates was performed based on pairwise genome comparisons using three different genomic approaches: cgMLST, cgSNP, and SKA. Additionally, antimicrobial resistance genes were detected using Abricate v1.0.1 against the NCBI database [25, 26].

Regarding clustering analysis, the observed distributions of pairwise allelic differences and SNPs were used to infer clustering cutoffs. More specifically, for each clustering methodology, the distribution of pairwise distances was modeled applying a Gaussian Mixture Model (GMM) using the Mclust function from the mclust R package (<https://www.r-project.org/>). The clustering cut-off was then set at the 99th percentile of the fit corresponding to the first component (cgMLST: ≤ 11 allelic differences; cgSNP: ≤ 22 SNPs for ST80 and ST80-like isolates, ≤ 9 SNPs for ST117 isolates; SKA: ≤ 30 SNPs; [Supplementary Fig. 1](#)). For cgMLST analysis, genomic clusters were defined as ≥ 2 isolates of the same sequence type (ST) and a genetic below the established threshold of allelic differences. cgMLST analysis was performed with Ridom SeqSphere + software 8.5 (Ridom GmbH, Münster, Germany) using SPAdes v3.15.4 for *de novo* assembly, using the typing scheme of 1423 genes published by Been

et al., [27] and the parameter "pairwise ignoring missing values". A quality criterion of $\geq 97\%$ good cgMLST targets and a coverage of $\geq 30\times$ was applied. From this analysis, ST, clonal complexes (CC) and complex types (CT) were also obtained. For cgSNP analysis, raw reads were trimmed using Trimmomatic v0.39 (LEADING:3 TRAILING:3 SLIDINGWINDOW:5:25 MINLEN:60) [28]. cgSNP analysis was performed separately for ST80 (and ST80-like) and ST117 isolates to provide a greater degree of pairwise SNP resolution, as recommended [16]. To establish the population structure of the isolates, for each predominant ST, trimmed reads were mapped to a close reference genome when possible, using snippy v4.6.0 (mincov 10, minfrac 0.9); genome alignments without masking for recombination were then used to infer a phylogenetic tree using IQtree v2.2.2.3 [29] with 1000 ultrafast bootstraps and a generalized time-reversible model (GTR). snp-dists was used to calculate pairwise SNP distances. Clusters were defined as ≥ 2 isolates of the same ST, with a genetic distance below the established SNP threshold according to ST and a monophyletic origin supported by a bootstrap value $\geq 90\%$. For SKA analysis, SNPs were defined as "number of split kmers found in both samples where the middle base is an A, C, G, or T but differs between files" [30].

Concordance Between Clustering Methodologies

Concordance between FTIR and each WGS method was assessed using Adjusted Rand (AR) and Adjusted Wallace (AW) [31–33], using an online tool (www.comparingpartitions.info) with 95% confidence intervals (95% CI). AR indicates the overall agreement of each cluster composition between two techniques. AW measures the directional probability that two isolates detected in a cluster by a reference technique are also detected in the same cluster using an alternative technique.

RESULTS

Emergence of VREfm

On April 2020, the first VREfm isolates were recovered from rectal swabs of two ICU patients. In November 2020, an outbreak was declared in surgery (three clinical and one surveillance sample). Secondary outbreaks occurred in January and April 2021 in oncohematology and nephrology, respectively. During late 2021, most isolates came from surgery, peaking in September. From February 2022 onward, sporadic cases appeared across wards, mostly not linked to outbreaks ([Figure 1](#), [Supplementary Table 1](#)). All isolates showed phenotypic resistance to vancomycin and all but one (isolate 84) susceptibility to linezolid. Regarding teicoplanin, only isolate 87 was sensitive.

cgMLST and cgSNP Analysis

Initial cgMLST analysis using a cutoff of ≤ 11 allelic differences revealed that most isolates belonged to two differentiated

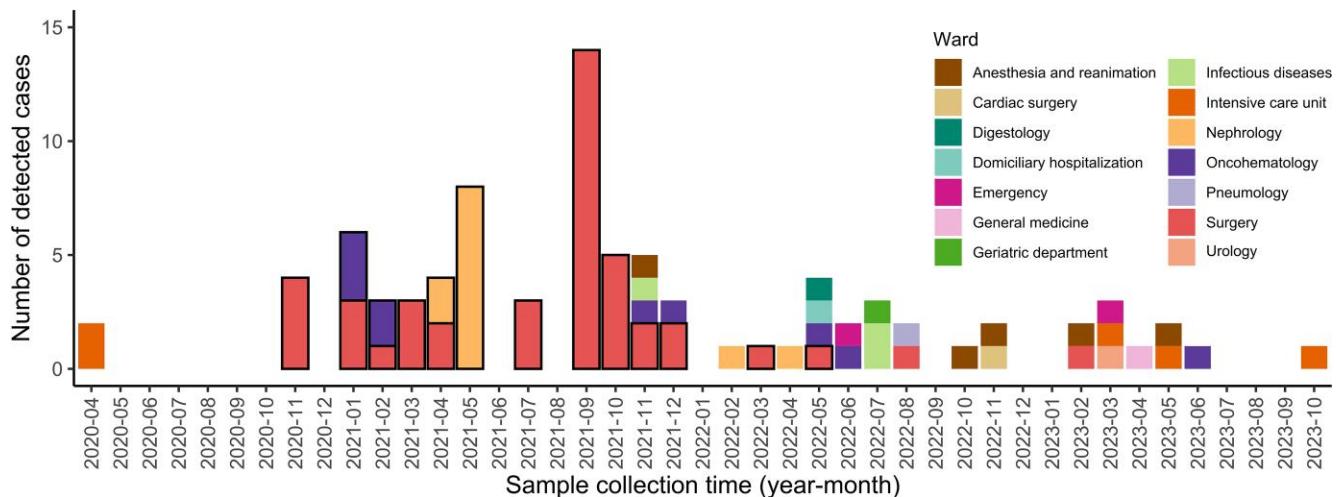


Figure 1. Distribution of VREfm isolates according to their hospital location. Isolates associated with each of the three epidemiological outbreaks (surgery, nephrology and oncohematology wards) are indicated with a black border.

populations: one consisting of ST80 ($n = 69$; of which 66 belonged to CT5967, two to CT6492 and one to CT847) and ST80-like isolates ($n = 4$; three with a novel *gyd* allele and one with a novel *purK* allele; all four belonging to CT5967), and the other comprising isolates of ST117 ($n = 12$). Additionally, one isolate belonged to ST761, and another belonged to novel ST2840 (*atpA* 9, *ddl* 1, *gdh* 1, *purK* 73, *gyd* 12, *pstS* 1, *adk* 1). All typed isolates belonged to clonal complex 17 (CC17). *vanA* was detected in all isolates except 87 (*vanB*); *cfr*(D) was present in isolate 84, matching phenotype. cgMLST grouped all but seven isolates into four clusters (cgMLST_1–cgMLST_4; Figure 2). cgMLST_1 contained 80.5% (70/87), including four environmental isolates, all of which belonged to CT5967. Other clusters had 2–4 isolates each. cgSNP with cut-offs ≤ 22 (ST80/ST80-like) and ≤ 9 (ST117) matched cgMLST clustering, identifying four clusters and seven singletons (Figure 2). Of note, the potential effect of not masking recombination sites on the identified clusters was assessed. Masking for recombination consistently resulted in lower pairwise SNP distances than their unmasked counterparts. However, cluster composition after readjusting clustering cut-offs (ie, ≤ 18 SNPs for ST80 and ST80-like isolates, ≤ 6 SNPs for ST117 isolates) was identical to that obtained without masking for recombination (Supplementary Figs. 2 and 3). Regarding outbreaks identified by conventional epidemiology, nephrology's outbreak was monoclonal (cgMLST_1), whereas surgery and oncohematology were polyclonal but dominated by cgMLST_1 (Supplementary Table 1).

SKA Clustering Reveals Little Subdivision Within cgMLST Clusters

SKA identified five clusters (SKA_1–SKA_5) and eight singletons using a cut-off of 30 SNPs. Among the cgMLST-defined

clusters, only cgMLST_1 was further subdivided by SKA into two distinct clusters (SKA_1 and SKA_2) and one singleton. SKA_1 was the largest cluster, comprising 75.9% (66/87) of the study isolates, while SKA_2 included the remaining three isolates originally assigned to cgMLST_1 (Figure 2). For all other isolates, identical clustering results between cgMLST and SKA were obtained ($AR_{cgMLST-SKA} = 0.850$ [0.708–0.998]). No clear epidemiological link existed among SKA_2 isolates; as they were ≥ 2 months apart and from different wards. Additionally, as the SNP threshold used for SKA clustering considerably differed from those commonly used in the literature (7–12 SNPs) [12–15], a sensitivity analysis was performed on this parameter. Lowering the threshold from 30 to 12 SNPs subdivided cluster SKA_1 into three clusters and one singleton, while not affecting the clustering of the remaining isolates (Supplementary Fig. 4). This resulted in a lower concordance between SKA and cgMLST ($AR_{cgMLST-SKA12} = 0.418$ [0.260–0.581]). Further decreasing the threshold continued to reduce the concordance between SKA and cgMLST (Supplementary Table 2).

FTIR Largely Captures WGS Clusters and Reflects Epidemiological Outbreaks

Given the similar results obtained across the three WGS-based approaches, and to facilitate interpretation, FTIR clustering performance was first compared with cgMLST as the reference method, and subsequently evaluated against SKA (confusion matrices are provided in Supplementary Tables 3–5). For FTIR clustering, an optimal cutoff range of 0.210 to 0.227 was obtained after maximizing AR (Supplementary Fig. 5) using cgMLST as the reference. Using this cutoff, FTIR yielded six clusters (FTIR_1–FTIR_6) plus five singletons (Figure 3).

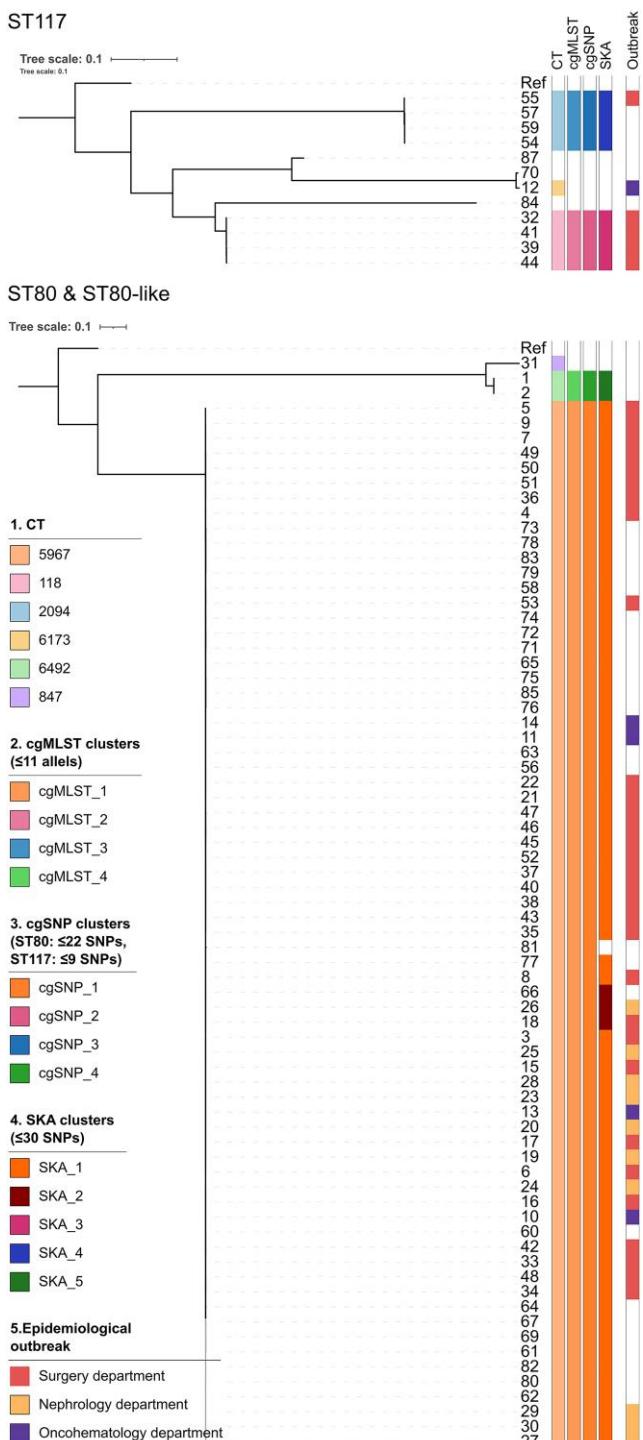


Figure 2. Maximum-likelihood phylogenetic tree (GTR) obtained from cgSNP analysis of VREfm isolates showing clusters identified by cgMLST, cgSNP and SKA approaches as well as epidemiological outbreak. ST80 reference: AUSMDU00004142 (accession no. CP027501.1); ST117 reference: VRE5755 (accession no. DAIGYV000000000.1).

Cluster cgMLST_1 was largely captured by cluster FTIR_1, which included 60 of the 70 cgMLST_1 isolates (85.7%), including all four environmental isolates. However, seven cgMLST_1

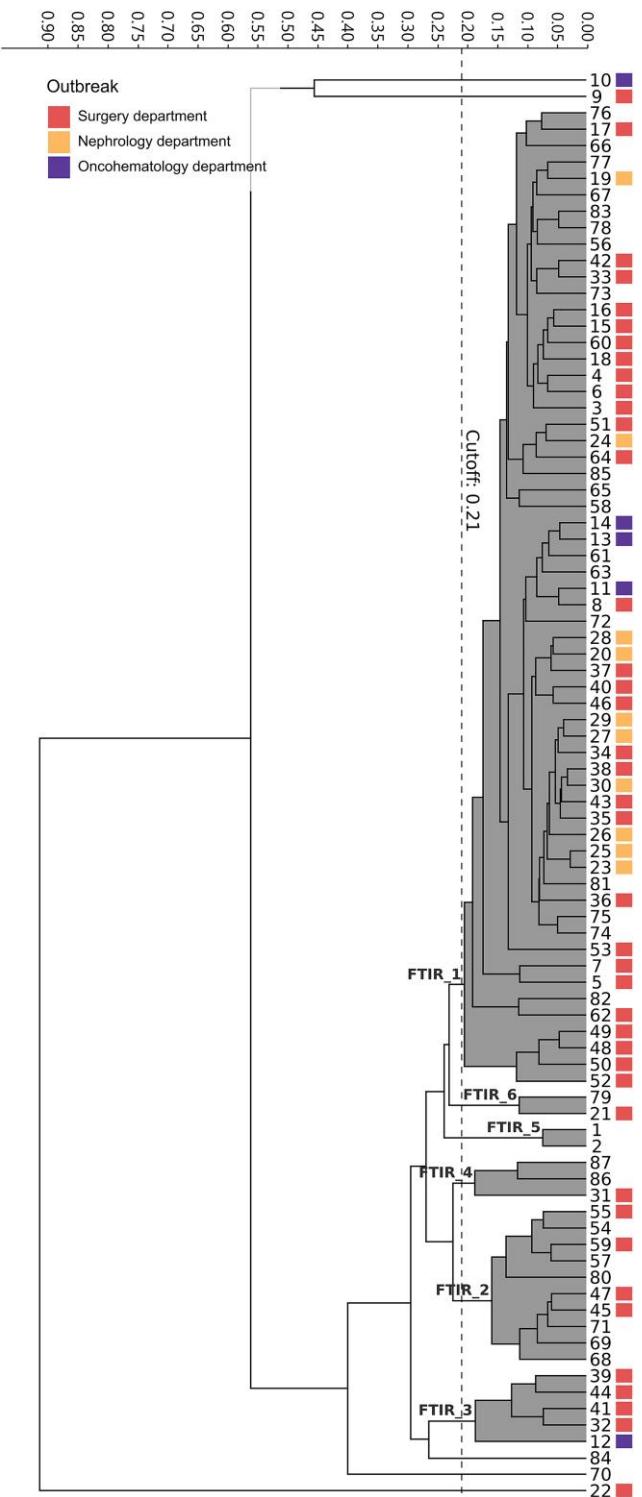


Figure 3. Dendrogram obtained by clustering the FTIR spectra of 87 VREfm. The vertical dashed line indicates the used cutoff value (0.21) for clustering analysis. The resulting FTIR clusters are shadowed in gray. Isolates belonging to epidemiological outbreaks are annotated in color.

isolates were misclassified by FTIR into clusters FTIR_2 ($n = 5$) and FTIR_6 ($n = 2$), and three were identified as singletons

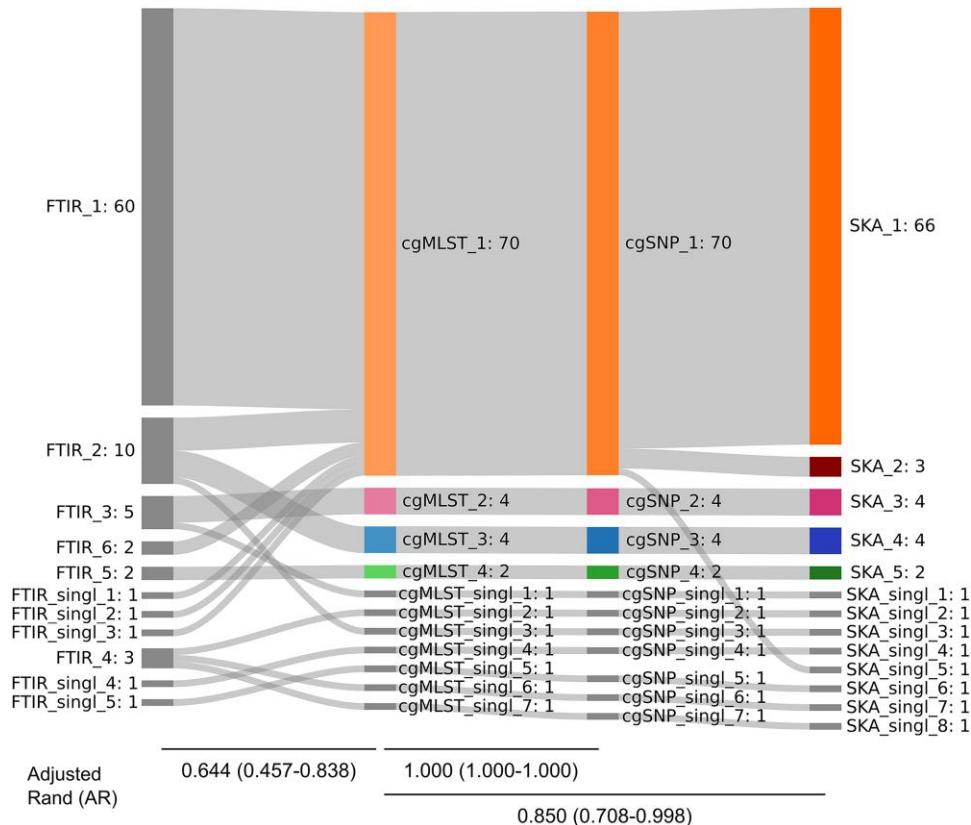


Figure 4. Relationship between Fourier-Transform Infrared spectroscopy (FTIR) and genomic methods (cgMLST, cgSNP, SKA) for genomic clustering of VREfm isolates. Clustering cutoffs (FTIR: ≤ 0.21 , cgMLST: ≤ 11 allelic differences, cgSNP: ≤ 22 SNPs for ST80 and ≤ 9 SNPs for ST117, SKA: ≤ 30 SNPs). Clusters and singletons detected by each method are indicated with a correlative number along with the number of cases in partition.

(Figure 4). FTIR_2 also included all four isolates of cgMLST_3 ($n = 4$, ST117) and isolate 68, of ST2840, thus clustering together isolates of different STs. Conversely, full concordance between cluster cgMLST_4 and FTIR_5 ($n = 2$) was observed. Lastly, three out of five cgMLST singletons were grouped together into cluster FTIR_4. Overall, these results yielded a value of 0.644 (95% CI, 0.457–0.838), and 0.944 (95% CI, 0.928–0.960) for AR and AW respectively, when comparing FTIR versus cgMLST (AW_{FTIR→cgMLST}) as the reference method. An AW value of 0.489 (95% CI, 0.224–0.754) for cgMLST versus FTIR (AW_{cgMLST→FTIR}) as the reference method. The concordance of FTIR with the SKA was slightly lower as SKA further subdivided cluster cgMLST_1 into clusters SKA_1 and SKA_2, and a singleton. Therefore, the comparison of SKA versus FTIR yielded an AR value of 0.543 (95% CI, 0.352–0.745), an AW_{FTIR→SKA} of 0.660 (95% CI, 0.396–0.923) and an AW_{SKA→FTIR} of 0.462 (95% CI, 0.187–0.737). Regarding the three hospital outbreaks, FTIR correctly inferred a true genomic relationship for 33 out of 41 (80.5%) VREfm isolates of the surgery department outbreak: 29 were assigned to cluster FTIR_1 (cgMLST_1) and 4 to cluster FTIR_2 (cgMLST_3); all 10 isolates of the nephrology department outbreak were

grouped into FTIR_1 (cgMLST_1); and, for the oncohematology department outbreak, FTIR clustered 3 of the 5 isolates of cluster cgMLST into cluster FTIR_1.

To further investigate the performance of FTIR in different epidemiological scenarios, we divided the cases in two periods: an “outbreak emergence and peak” period from April 2020 to December 2021 ($n = 60$), when most isolates were associated with outbreaks in the surgery, nephrology and oncohematology wards, and a “sporadic cases” period from February 2021 to October 2023 ($n = 27$). The corresponding AR_{cgMLST→FTIR} for the emergence and peak period was 0.686 (95% CI, 0.462–0.919), slightly higher than that of the post-outbreak period, which was 0.562 (95% CI, 0.215–0.938). This suggests that the performance of FTIR might be slightly higher during active outbreak periods.

DISCUSSION

VREfm is a major cause of a range of infections in humans, including, urinary tract infections, wound infections, bloodstream infections, and endocarditis, and is frequently associated with nosocomial outbreaks. Although our 2020–

2021 outbreaks had limited clinical impact, since most cases represented colonization (rectal swabs) rather than infection and no associated fatalities occurred, they demanded substantial infection-control efforts. These included patient isolation, reinforcement of hygiene measures and enhanced environmental disinfection. Four environmental isolates were also recovered, confirming the persistence of VREfm in the hospital environment and supporting targeted cleaning interventions. These measures were guided by typing results, which enabled the rapid identification of transmission clusters and differentiation between unrelated cases, thereby underscoring the importance of rapid and reliable typing methods for outbreak management. Therefore, in the present study, we evaluated FTIR for identifying VREfm genomic clusters. We show that FTIR clusters VREfm isolates in close agreement with multiple WGS approaches, supporting its use as a practical first-line tool for initial screening and outbreak detection to enable faster infection-control responses. In addition to its value for early outbreak detection, FTIR can be particularly useful for ruling out sporadic cases that are not related to previous outbreak strains, thereby avoiding unnecessary investigations and infection control measures. Moreover, identifying sporadic VREfm cases with FTIR can support antimicrobial stewardship programs by indicating elevated antibiotic pressure and facilitating targeted interventions to reduce unnecessary antimicrobial use in those clinical units.

Previous studies have demonstrated the potential of FTIR as a typing tool for Gram-negative bacteria, including *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and *Acinetobacter baumannii* [20–22]. In line with this, we previously evaluated FTIR for typing *Klebsiella pneumoniae* nosocomial outbreaks, with successful implementation in our hospital setting [23]. More recently, FTIR has also been applied to Gram-positive bacteria, including *Enterococcus faecium*, as demonstrated in recent studies [2, 34, 35]. Here, we provide additional evidence for the expansion of FTIR applicability to outbreak management beyond Gram-negative pathogens.

Most isolates were ST80 and ST117, high-risk clones associated with nosocomial spread [36, 37]. ST80 (83.9% of isolates) predominates in Spain [38], and a recent study in Barcelona reported perfect early-detection concordance between FTIR and genomic methods for an ST80 spread [35]. Although confirmation would be required to determine whether the strains responsible for the outbreaks in our hospital and in the Barcelona study are related, these findings collectively suggest the emergence of ST80 VREfm isolates causing nosocomial outbreaks in this geographic area. Given the rapid dissemination of this ST, continuous surveillance is crucial.

Given the absence of a single gold-standard WGS strategy, we compared cgMLST, cgSNP, and SKA [12, 15]. In the literature, cgMLST analysis has been widely used and is often considered as the standard for outbreak analysis of diverse bacterial

species in many settings. cgMLST uses a defined scheme which comprises a large number of conserved genes, which allows for dataset comparison in a standardized manner. Compared with cgMLST, cgSNP analysis offers higher cluster resolution by relying on SNP differences rather than allele distances. However, its performance is highly dependent on the choice of the reference genome. Using a distantly related reference can artificially inflate the number of detected pairwise SNPs, thereby compromising the fine-scale resolution required for accurate outbreak delineation [14]. More recently, SKA has emerged as a robust method for VREfm outbreak investigation, offering higher discriminatory power than cgMLST [12, 15]. Studies by Higgs *et al.* and Maechler *et al.* have shown that SKA more accurately infers patient-to-patient transmission, outperforming cgMLST at commonly used thresholds of ≤ 25 and ≤ 20 allelic differences, respectively. At these levels, cgMLST often lacks sufficient resolution to differentiate closely related strains. Despite its strengths, a key challenge in SKA—and in WGS-based approaches in general—is defining appropriate clustering cutoffs. These must be interpreted alongside epidemiological data, which are often complex and context-dependent. In our study, clustering thresholds were selected at the 99th percentile first peak in the global distribution of pairwise allelic or SNP differences and were supported by strong spatiotemporal links among isolates. This approach resulted in a clustering cutoff for cgMLST of ≤ 11 allelic differences, improving cgMLST resolution compared with commonly used thresholds of ≤ 25 or ≤ 20 . cgSNP cutoffs differed among STs: ≤ 22 SNPs for ST80 and ST80-like, and ≤ 9 SNPs for ST117. Finally, for SKA, a threshold of 30 SNPs was selected—considerably higher than the 7–12 SNP range commonly reported in the literature [12–15]. However, it is important to note that FTIR performance was still acceptable when compared with SKA clustering at a threshold of 12 SNPs. Besides, our tailored approach led to considerable harmonization across the three WGS analyses, in contrast with previous studies where results varied substantially depending on the method and parameters used [12, 15].

In the present study, an AR of 0.644 (95% CI, 0.457–0.838), an $AW_{FTIR \rightarrow cgMLST}$ of 0.944 (95% CI, 0.928–0.960) and an $AW_{cgMLST \rightarrow FTIR}$ of 0.489 (95% CI, 0.224–0.754) were obtained. Teng *et al.* [16] suggested that an AW of 0.95 or higher should be met when comparing FTIR versus WGS as the gold standard method. This ensures that FTIR clusters the large majority of isolates clustered by WGS without missing potential transmission events. Additionally, an AW of at least 0.5 should be met when using WGS compared with FTIR as a reference method to avoid detection of falsely clustered isolates, which would trigger unnecessary WGS confirmation. In this work, both AW values were marginally below those proposed by Teng *et al.* A slightly lower concordance was observed between FTIR and SKA clustering (AR 0.543 [95% CI, 0.352–0.745]). This value was similar to the results obtained by Park and

Ryoo, who assessed the performance of the FTIR for VREfm clustering in a neonatal intensive care unit outbreak involving four patients. The outbreak was predominantly caused by ST17 and an AR of 0.718 (95% CI, 0.466–0.996) between FTIR and SKA was obtained [39]. In contrast, our study includes three distinct epidemiological outbreaks—one monoclonal and two polyclonal—Involving 87 patients, thereby representing a higher level of epidemiological complexity. Thus, our results taken together with those reported by Pitart *et al* [35]. Highlight a potentially broader applicability of FTIR beyond monoclonal outbreak scenarios. It is also important to note that it has been suggested that the performance of FTIR decreases as the dataset size increases [16], which could be partially the cause of the lower performance of FTIR observed in the present study in comparison to the previously published literature. Consequently, further investigation of the performance of FTIR using large datasets of different STs and geographical areas is still needed.

As with WGS, the lack of a standardized clustering cutoff across different laboratories represents a key limitation of FTIR. Before implementing FTIR for early detection of genomic clusters, each laboratory must validate its own clustering cutoff, which may vary based on local epidemiology and the culture media used. The manufacturer recommends a cutoff range of 0.15 to 0.20 for *E. faecium*. Our findings, consistent with previously published data, suggest that the optimal cutoff lies within or near this recommended range [2, 35, 39]. Another limitation of FTIR is its lower discriminatory power relative to WGS-based methods. FTIR provides a global spectral fingerprint that reflects multiple cellular components, including polysaccharides, fatty acids, and proteins. However, the wavelength range applied in our study (1300–800 cm⁻¹) primarily targets the polysaccharide region of the bacterial envelope. Consequently, isolates with identical genomic backgrounds may exhibit differences in their polysaccharide profiles, potentially leading to misclassification [22].

Our study also has several limitations. Firstly, isolates were predominantly ST80 and ST117 and originated from a single hospital. This limited genetic diversity and single-center design restrict the generalizability of our findings, as FTIR performance may differ when applied to other VREfm sequence types, geographic regions, or epidemiological contexts. As FTIR spectral signatures can be influenced by strain-specific phenotypic characteristics, further validation using a more genetically diverse set of isolates is warranted. Expanding FTIR analyses to multicenter collections could not only strengthen methodological validation but also facilitate the early detection of new emerging VREfm lineages, as recently proposed in large-scale FTIR surveillance studies [40]. Secondly, despite the acknowledged inherent differences between genotypic and phenotypic methods, the reasons behind the lack of

concordance between FTIR and cgMLST for some isolates were beyond the scope of this study and remain unclear.

Collectively, our results showed that FTIR clustering provides results comparable to those obtained from WGS analyses. While FTIR delivers more limited information, its speed and simplicity make it a viable method for real-time outbreak management. FTIR represents a cost-effective approach (approximately 17€ per sample compared with 70€ for WGS considering only reagents) and does not require highly specialized technical personnel. In terms of turnaround time, WGS results from a positive culture may take up to approximately a week, whereas FTIR can deliver results in about 3 hours.

Based on these findings, we propose a practical workflow for outbreak management. In suspected outbreak situations, FTIR could be used as an initial screening tool to rapidly assess the relatedness of isolates. If FTIR results indicate clonal spread, infection control measures to tackle active transmission should be initiated. However, in the case of epidemiologically linked cases not clustered by FTIR, WGS analysis should be performed to assess genomic relatedness and guide further interventions. Thus, FTIR can serve as a frontline tool for outbreak confirmation in settings where WGS is not immediately accessible, enabling faster and more targeted outbreak responses.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. We thank the Germans Trias i Pujol Research Institute (IGTP) Translational Genomics Core Facility and staff for their contribution to this publication.

Data availability. The raw reads analyzed for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under BioProject PRJEB82335.

Financial support. This work was supported by the Departament de Salut de la Generalitat de Catalunya [grant number PERIS SLT028/23/000129 to A. E. B.]

Potential conflicts of interest. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kjær Hansen S, Andersen L, Detlefsen M, et al. Using core genome multilocus sequence typing (cgMLST) for vancomycin-resistant *Enterococcus faecium* isolates to guide infection control interventions and end an outbreak. *J Glob Antimicrob Resist* 2021; 24:418–23.
2. Scheier TC, Franz J, Boumasmoud M, et al. Fourier-transform infrared spectroscopy for typing of vancomycin-resistant *Enterococcus faecium*: performance analysis and outbreak investigation. *Microbiol Spectr* 2023; 11:e0098423.
3. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 2022; 399:629–55.
4. Leclercq R, Derlot E, Duval J, Courvalin P. Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N Engl J Med* 1988; 319: 157–61.

5. European Centre for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-net)—annual epidemiological report 2023. Stockholm: ECDC, 2024.
6. Cimen C, Berends MS, Bathoorn E, et al. Vancomycin-resistant enterococci (VRE) in hospital settings across European borders: a scoping review comparing the epidemiology in the Netherlands and Germany. *Antimicrob Resist Infect Control* 2023; 12:78.
7. Cimen C, Voss A, Hellkamp J, Hamprecht A, Berends MS. Temporal trends of enterococcal and vancomycin-resistant *Enterococcus faecium* bacteraemia in the northern Dutch-German cross-border region: a 10-y multicentre analysis (2013–2022). *J Glob Antimicrob Resist* 2025; 42:187–94.
8. European Centre for Disease Prevention and Control and World Health Organization. Antimicrobial resistance surveillance in Europe 2023–2021 data. Solna: ECDC, 2023.
9. Bertelli C, Greub G. Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clin Microbiol Infect* 2013; 19:803–13.
10. European Centre for Disease Prevention and Control. Monitoring the use of whole-genome sequencing in infectious disease surveillance in Europe. Solna: ECDC, 2018. 31.
11. European Centre for Disease Prevention and Control. ECDC Strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations—2019–2021. Stockholm: ECDC, 2019.
12. Higgs C, Sherry NL, Seemann T, et al. Optimising genomic approaches for identifying vancomycin-resistant *Enterococcus faecium* transmission in healthcare settings. *Nat Commun* 2022; 13:509.
13. Deurenberg RH, Bathoorn E, Chlebowicz MA, et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol* 2017; 243:16–24.
14. Gorrie CL, Da Silva AG, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe* 2021; 2:e575–83.
15. Maechler F, Weber A, Schwengers O, et al. Split k-mer analysis compared to cgMLST and SNP-based core genome analysis for detecting transmission of vancomycin-resistant enterococci: results from routine outbreak analyses across different hospitals and hospitals networks in Berlin, Germany. *Microb Genomics* 2023; 9:1–14.
16. Teng ASJ, Habermehl PE, van Houdt R, et al. Comparison of fast Fourier transform infrared spectroscopy biotyping with whole genome sequencing-based genotyping in common nosocomial pathogens. *Anal Bioanal Chem* 2022; 414: 7179–89.
17. Naumann D, Helm D, Labischinski H. Microbiological characterizations by FT-IR spectroscopy. *Nature* 1991; 351:81–2.
18. Griffiths PR, de Haseth JA. Fourier Transform infrared spectrometry. Hoboken, New Jersey: Wiley, 2007; 222:297–302.
19. Deidda F, Bozzi Cionci N, Cordovana M, et al. Bifidobacteria strain typing by Fourier transform infrared spectroscopy. *Front Microbiol* 2021; 12:1–10.
20. Martak D, Valot B, Sautet M, et al. Fourier-Transform InfraRed spectroscopy can quickly type gram-negative Bacilli responsible for hospital outbreaks. *Front Microbiol* 2019; 10:1440. doi:10.3389/fmicb.2019.01440
21. Novais Â, Freitas AR, Rodrigues C, Peixe L. Fourier transform infrared spectroscopy: unlocking fundamentals and prospects for bacterial strain typing. *Eur J Clin Microbiol Infect Dis* 2019; 38:427–48.
22. Rakovitsky N, Frenk S, Kori H, et al. Fourier transform infrared spectroscopy is a new option for outbreak investigation: a retrospective analysis of an extended-Spectrum-Beta-lactamase-producing *Klebsiella pneumoniae* outbreak in a neonatal intensive care unit. *J Clin Microbiol* 2020; 58:e00098–20.
23. Wang-Wang JH, Bordoy AE, Martínez E, et al. Evaluation of Fourier transform infrared spectroscopy as a first-line typing tool for the identification of extended-Spectrum β-lactamase-producing *Klebsiella pneumoniae* outbreaks in the hospital setting. *Front Microbiol* 2022; 13:897161.
24. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters, version 13.1. Data access December 2023. Available at: <http://www.eucast.org>
25. Seemann T. Abricate. Github. Available at: <https://github.com/tseemann/abricate>
26. Feldgarden M, Brover V, Haft DH, et al. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob Agents Chemother* 2019; 63: e00483–19.
27. de Been M, Pinholt M, Top J, et al. Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. *J Clin Microbiol* 2015; 53:3788–97.
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 2014; 30:2114–20.
29. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020; 37:1530–4.
30. Harris SR. SKA: split kmer analysis toolkit for bacterial genomic epidemiology. *bioRxiv* [Preprint]. 2018. 453142. Available at: <https://www.biorxiv.org/content/10.1101/453142v1.abstract>
31. Santos JM, Embrechts M. On the use of the adjusted rand Index as a metric for evaluating supervised classification, eds. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). Berlin, Heidelberg: Springer, 2009:175–84.
32. Severiano A, Pinto FR, Ramirez M, Carriço JA. Adjusted wallace coefficient as a measure of congruence between typing methods. *J Clin Microbiol* 2011; 49: 3997–4000.
33. Hoffman M, Steinley D, Brusco MJ. A note on using the adjusted Rand index for link prediction in networks. *Soc Networks* 2015; 42:72–9.
34. Wong RCW, Cheung IYY, Lai CKC, et al. Potential use of Fourier-transform infrared spectroscopy as a rapid screening tool for investigating nosocomial outbreaks of ST-80 vancomycin-resistant *Enterococcus faecium*. *J Hosp Infect* 2025; 155:161–9.
35. Pitart C, Piquet M, Burgwinkel T, et al. Early identification of the nosocomial spread of vancomycin-resistant *Enterococcus faecium* by Fourier-transform infrared spectroscopy and performance comparison with PFGE and WGS. *Emerg Microbes Infect* 2024; 13:1–11.
36. Freitas AR, Pereira AP, Novais C, Peixe L. Multidrug-resistant high-risk *Enterococcus faecium* clones: can we really define them? *Int J Antimicrob Agents* 2021; 57:106227.
37. Falgenhauer L, Fritzenwanker M, Imrzalioglu C, et al. Near-ubiquitous presence of a vancomycin-resistant *Enterococcus faecium* ST117/CT71/vanB –clone in the rhine-main metropolitan area of Germany. *Antimicrob Resist Infect Control* 2019; 8:128.
38. Tedim AP, Ruiz-Garabajosa P, Rodríguez MC, et al. Long-term clonal dynamics of *Enterococcus faecium* strains causing bloodstream infections (1995–2015) in Spain. *J Antimicrob Chemother* 2017; 72:48–55.
39. 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; 12:31–.
40. Lurie-Weinberger MN, Temkin E, Kastel O, et al. Use of a national repository of Fourier-transform infrared spectroscopy spectra enables fast detection of silent outbreaks and prevention of spread of new antibiotic-resistant sequence types. *Antimicrob Resist Infect Control* 2025; 14:34.