Screening of biomarkers of drug resistance or virulence in ESCAPE pathogens by MALDI-TOF mass spectrometry

Samantha Flores-Treviño, Elvira Garza-González, Soraya Mendoza-Olazarán, Rayo Morfin-Otero, Adrián Camacho-Ortiz, Eduardo Rodríguez-Noriega, Adrián Martínez-Meléndez & Paola Bocanegra-Ibarías

Rapid identification and characterisation of drug-resistant bacterial pathogens have an important role in diagnostic and antimicrobial stewardship. Response time in the diagnosis of not only the etiological agent but also in antimicrobial susceptibility results is of utmost importance in patient treatment. In this study, matrix-assisted laser desorption ionisation–time of flight (MALDI-TOF) mass spectrometry (MS) was used to screen for biomarkers of ESCAPE (vancomycin-resistant *Enterococcus faecium*, methicillin-resistant *Staphylococcus aureus*, hypervirulent NAP1/ribotype 027 *Clostridioides (Clostridium) difficile*, multidrug resistant *Acinetobacter baumannii*, multidrug resistant *Pseudomonas aeruginosa*, and carbapenem-resistant *Enterobacteriaceae*) pathogens to predict antimicrobial resistance or hypervirulence. Several biomarkers of drug-resistant genotypes in *S. aureus*, *A. baumannii*, *P. aeruginosa*, and *K. pneumoniae*, as well as hypervirulence in *C. difficile*, were detected. The fastest possible susceptibility testing with MALDI-TOF MS is simultaneous detection of a characteristic drug-resistant peak and species identification in the same spectra generated in routine processing. According to our approach, resistance or virulence biomarker peaks can be identified while performing routine microbiology analysis, and no additional assays nor prolonged incubation time is needed. Outstanding biomarker peaks detected in our study should be further analysed by additional methods to identify the specific proteins involved.

The emergence of drug-resistant bacterial strains has been accelerated by both antibiotic overuse and reduced infection control. In 2013, the Centers for Disease Control and Prevention issued a report with the most concerning drug-resistant threats in the United States. Of those mentioned, the most notable are pathogens predominantly associated with healthcare-associated infections (HAI).

Reports by the Infectious Diseases Society of America have highlighted the importance of specific nosocomial drug-resistant pathogens. Designated by the acronym ESCAPE, this group comprises the following pathogens: vancomycin-resistant *Enterococcus faecium* (VREfm), methicillin-resistant *Staphylococcus aureus* (MRSA), carbapenem-resistant *Klebsiella pneumoniae*, multidrug resistant (MDR) *Acinetobacter baumannii*, MDR *Pseudomonas aeruginosa*, and carbapenem-resistant *Enterobacter cloacae*. These pathogens are of considerable concern due to their high frequency in HAI.

*Clostridioides (Clostridium) difficile* infection is one of the most common HAI s in hospitalised patients receiving antimicrobial therapy. In particular, the infection of hypervirulent *C. difficile* NAP1/ribotype 027 is associated with severe disease and high mortality rates. In recent years, hospitals have reported a rise in *C. difficile* infection rates. Thus, researchers have suggested amending the ESCAPE group by replacing *K. pneumoniae* with *C. difficile* and adding the entire *Enterobacteriaceae* family (instead of only *K. pneumoniae* and *E. cloacae*), changing the acronym to ESCAPE.
| Species, virulence/resistance phenotype (n) | Theoretical peaks (p ≤ 0.05) | Potential biomarkers | Mascot score | Protein sequence coverage (%) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|-------------------------------------------|-----------------------------|----------------------|--------------|-------------------------------|----------------|----------------|---------|---------|
| E. faecium, vancomycin-resistant (34) vs vancomycin-susceptible (30) | 70 | 0 | NA | NA | NA | NA | NA | NA |
| MRSA (36) vs MSSA (31) | 45 | 1 | 4,594 (P) | 0.90 < 0.001 | 50 S ribosomal protein L28 | 26 | 64 | 83.3 | 96.8 |
| C. difficile, 027 ribotype (57) vs non-027 (36) | 76 | 2 | 6,654 (A) | 0.96 < 0.001 | 30S ribosomal protein S20 | 31 | 67 | 100.0 | 91.7 |
| A. baumannii, bla OXA-24 or bla OXA-58 (33) vs non-bla OXA-24 or bla OXA-58 (35) | 2 | 0 | NA | NA | NA | NA | NA | NA |
| A. baumannii, bla OXA-24, OXA-58 (26) vs bla OXA-58 (7) | 6 | 2 | 6,304 (P) | 0.99 < 0.001 | NADH-quinone oxidoreductase subunit K | 29 | 56 | 100.0 | 100.0 |
| K. pneumoniae, carbapenem-resistant (49) vs carbapenem-susceptible (20) | 96 | 0 | NA | NA | NA | NA | NA | NA |
| K. pneumoniae, carbapenem- and colistin-resistant (16) vs carbapenem- and colistin-susceptible (20) | 75 | 1 | 6,100 (A) | 0.86 < 0.001 | Not determined | NA | NA | 93.8 | 62.5 |

Table 1. Distribution of potential peaks detected, protein/peptide assignment and diagnostic utility. A: absence; AUC: area under the curve; MDR: multidrug resistant; MRSA: methicillin-resistant S. aureus; MSSA: methicillin-susceptible S. aureus; NPV: negative predictive value; P: presence; PPV: positive predictive value.

Rapid identification and characterisation of drug-resistant bacterial pathogens have an important role in diagnostic and antimicrobial stewardship. The mission of the Antibacterial Resistance Leadership Group is to reduce the public health threat of antibiotic resistance14, which can be achieved by rapid detection of MDR isolates in clinical diagnostic laboratories. Matrix-assisted laser desorption ionisation–time of flight (MALDI-TOF) mass spectrometry (MS) is a fast and reliable technique that has been implemented in clinical microbiology laboratories to replace or complement conventional phenotypic identification for most isolated bacterial strains13. Although the report of antibiotic susceptibilities in the laboratory is necessary, the rapid determination of antimicrobial resistance or hypervirulence.

Results

Selected genes and virulence factors. Among the E. faecium isolates, 53.1% (34/64) were vancomycin-resistant, and 46.9% (30/64) were vancomycin-susceptible (Table 1). The vanA gene was detected in all (34/34) VREFm isolates, whereas the vanB gene was not detected in any of the strains.

Of the S. aureus isolates, 53.7% (36/67) were MRSA (cefotixin-resistant) and meca positive; 46.3% (31/67) were methicillin-susceptible S. aureus (MSSA) and meca negative (Table 1).

For C. difficile isolates, 53.7% (36/67) were MRSA (cefotixin-resistant) and meca positive; 46.3% (31/67) were methicillin-susceptible S. aureus (MSSA) and meca negative (Table 1).

The following non-027 ribotypes were identified: 001 (n = 17; 47.2%), 106 (n = 4; 11.1%), 003 (n = 3; 8.3%), 176 (n = 3; 8.3%), 002 (n = 1; 2.8%), 012 (n = 1; 2.8%), 014 (n = 1; 2.8%), 017 (n = 1; 2.8%), 019 (n = 1; 2.8%), 020 (n = 1; 2.8%), 076 (n = 1; 2.8%), 220 (n = 1; 2.8%), and 353 (n = 1; 2.8%).

For A. baumannii isolates, 45.6% (31/68) were carbapenem resistant. Class D blaOXA Carbapenemase genes blaOXA-24 and blaOXA-58 were detected in 26/68 and 7/68 isolates, respectively. In addition, blaOXA-58 was detected in two carbapenem-susceptible isolates. As expected, the blaOXA-51 gene was detected in all (68/68) strains. Neither class D blaOXA-24, nor encoding metallo-β-lactamases genes blaOXA-23, blaVIM, and blaIMP genes were detected.

Among P. aeruginosa isolates, 49.1% (27/55) were MDR, and 50.9% (28/55) were non-MDR. The blaVIM and blaIMP genes were not detected.
For *K. pneumoniae* isolates, 72.6% (49/69) were carbapenem resistant, and 27.4% (20/69) were carbapenem susceptible. Of the carbapenem-resistant isolates, 16/49 (32.7%) were colistin-resistant; none of the carbapenem-susceptible isolates presented colistin resistance. The *bla*\_NDM gene was detected in all the carbapenem-resistant isolates. Neither *bla*\_KPC, *bla*\_IMP, *bla*\_VIM, nor *bla*\_OXA-48 were detected.

Potential biomarkers. The generated protein profile of each isolate spot had mass-charge ratio (*m/z*) values between 2,000 and 20,000. A MS spectrum sample of each species is shown in Supplementary Fig. S1. Classification models were generated based on the GA, SNN, and QC algorithms, in which at least one of them yielded adequate cross-validation and recognition capability values for vancomycin-resistant vs vancomycin-susceptible *E. faecium* (90% and 98%, respectively), *C. difficile* ribotype 027 vs non-027 (90% and 96%, respectively), meca positive vs meca negative (87% and 98%, respectively), carbapenem-resistant vs carbapenem-susceptible *K. pneumoniae* (78% and 90%, respectively), MDR vs non-MDR *P. aeruginosa* (76% and 89%, respectively), carbapenem/colistin-resistant vs carbapenem/colistin-susceptible *K. pneumoniae* (67% and 83%, respectively), *bla*\_OXA-24 vs *bla*\_OXA-58 carbapenem-resistant *A. baumannii* (59% and 51%, respectively), and *bla*\_OXA-24/58 positive vs *bla*\_OXA-24/58 negative *A. baumannii* (50% and 72%, respectively). A complete description of all classification analyses is shown in Supplementary Table S2. The ability of each algorithm to classify spectra from the same set of samples was used to perform an external validation. At least one of the models was able to allocate most of the spectra of the validation set to their corresponding phenotype/genotype profile (Table S2).

In the initial screening and after group comparisons (phenotype/genotype absent or present), statistically significant peaks (*p* ≤ 0.05) were detected for carbapenem-resistant *K. pneumoniae* (*n* = 96), *C. difficile* ribotype 027 (*n* = 76), carbapenem and colistin-resistant *K. pneumoniae* (*n* = 75), VREfm (*n* = 70), MDR *P. aeruginosa* (*n* = 61), MRSA (*n* = 45), *bla*\_OXA-58 *A. baumannii* (*n* = 6), and carbapenem-resistant *A. baumannii* (*bla*\_OXA-24 or *bla*\_OXA-58; *n* = 2) (Table 1). The complete list describing all significant peaks is shown in Supplementary Dataset S3. The distribution of each strain after PCA based analysis per phenotype/genotype of all species is shown in Supplementary Fig. S4. The distinguishing capability between drug-resistance (or virulence) phenotype/genotype profiles in the isolates based on their peptide mass fingerprints was highly variable. These statistically significant peaks were selected for further analysis which were the analysis of the peak area/intensity of the spectra, the coefficient of variation and the value of area under the curve (AUC) of each peak. When all three criteria were met, these peaks were considered as "potential biomarkers peaks".

When analysis VREfm, no potential biomarkers were finally selected. For *S. aureus* isolates, one peak with a *m/z* of 4,594 (with a value of area under the receiver operating characteristics curve [AUC-ROC] of 0.82; *p* < 0.001) was selected (Table 1). This peak was present as a singlet in the MSSA isolates; however, in the MRSA isolates, it displayed a splitting pattern as a doublet (Fig. 1). The presence of the peak *m/z* 4,594 as a doublet had a sensitivity of 83.3% and a specificity of 96.8% for the detection of methicillin-resistant phenotype (Table 1). Furthermore, a peak previously reported as a potential biomarker (*m/z* 2,415) was detected in 63.9% (23/36) of the MRSA isolates.

For *C. difficile* isolates, peaks were selected by strain: the peak at *m/z* 6,654 (AUC = 0.96) was present only in non-027 strains, except for the 176 ribotype; and the peak at *m/z* 6,712 (AUC = 0.99; *p* < 0.001) was present only
in 027 and 176 strains (Fig. 2). To detect the hypervirulent 027 ribotype, the absence of peak \( m/z \) 6,654 and presence of peak \( m/z \) 6,712 presented a sensitivity of 100%, specificity of 91.7%, and PPV of 95%.

\( A. \) \textit{baumannii} \( \text{blaOXA-24} \) were compared with \( \text{blaOXA-58} \) isolates, and two peaks were selected: the peak at \( m/z \) 6,304 (AUC = 0.99; \( p < 0.001 \)) was present only in \( \text{blaOXA-58} \)-positive isolates (specificity, and positive predictive value [PPV] of 100%), and the peak at \( m/z \) 6,332 (AUC = 0.99, \( p < 0.001 \)) was present only in \( \text{blaOXA-24} \)-positive isolates (sensitivity, specificity, PPV, and negative predictive value [NPV] of 100%) (Table 1) (Fig. 3). The comparison of \( \text{blaOXA-58} \)-positive with \( \text{blaOXA-58} \)- plus \( \text{blaOXA-24} \)-negative isolates yielded a statistically significant peak at \( m/z \) 6,332 (\( p = 0.016 \)). The comparison of isolates that harboured either \( \text{blaOXA-24} \) or \( \text{blaOXA-58} \) with isolates that harboured exclusively \( \text{blaOXA-51} \) genes showed no potential biomarkers.

\( P. \) \textit{aeruginosa} isolates, three potential biomarkers were detected at peaks \( m/z \) 2,726 (AUC = 0.81; \( p < 0.001 \)), 5,455 (AUC = 0.81; \( p < 0.001 \)), and 5,742 (AUC = 0.84; \( p < 0.001 \)), all of which predominated in the non-MDR isolates (Fig. 4). The presence of all three peaks increased the specificity to 74.1%, PPV to 75%, and NPV to 74.1% for the detection of non-MDR isolates (Table 1).

No potential biomarkers were detected for the carbapenem-resistant \( K. \) \textit{pneumoniae} isolates. However, the absence of peak \( m/z \) 6100 (AUC = 0.88) had a sensitivity of 93.8% and a NPV of 91.7% for the detection of carbapenem and colistin-resistant \( K. \) \textit{pneumoniae} (Fig. 5) (Table 1).

**Peptide/protein identity assignment.** Assignment of possible peptide/protein identity of all the identified potential marker peaks is shown in Table 1. The \( m/z \) 4,594 peak of \( S. \) \textit{aureus} was assigned to the 50S ribosomal protein L28 with MASCOT score of 26 and protein sequence coverage of 64% Both the \( m/z \) 6,654 and 6,712 peaks of \( C. \) \textit{difficile} were assigned to the 30S ribosomal protein (S20 and S21, respectively), with 31–34 score and 67–98% protein sequence coverage. Both the \( m/z \) 6,304 and the \( m/z \) 6,332 peak of \( P. \) \textit{aeruginosa} was assigned to a NADH-quinone oxidoreductase subunit K with a score of 29 and 56% of protein sequence coverage for each peak. In \( P. \) \textit{aeruginosa}, the \( m/z \) 2,726 peak was assigned to a UPF0270 protein PI01_4103 with a score of 20 and 32% of protein sequence coverage, and the \( m/z \) 5,455 peak was assigned to a UPF0391 membrane protein Patl_1732 with a score of 32 and 98% of protein sequence coverage. The \( m/z \) 5,742 peak of \( P. \) \textit{aeruginosa} and the \( m/z \) 6,100 peak of \( K. \) \textit{pneumoniae} could not be determined.

**Discussion**

MALDI-TOF MS is now widely used in clinical microbiology for bacterial identification, taxonomy, and strain typing. One application in development is the detection of antibiotic resistance in bacteria. Thus far, several approaches have been proposed in MALDI-TOF MS to detect antimicrobial resistance such as the detection of the entire cell profile, enzymatic activity by antibiotic hydrolysis, or resistance proteins within the cell. In this study, we identified several potential biomarkers of drug-resistant genotypes in \( S. \) \textit{aureus}, \( A. \) \textit{baumannii}, \( P. \) \textit{aeruginosa}, and \( K. \) \textit{pneumoniae}, as well as hypervirulence in \( C. \) \textit{difficile}, using a direct approach.

Regarding \( C. \) \textit{difficile}, 2 out of 73 peaks reached statistical significance, wherein isolates were divided into 027 and non-027 ribotype groups. The peak \( m/z \) 6,654 was absent in 027 and 176 ribotypes, whereas the peak...
m/z 6,712 was present in both ribotypes. These findings complicate the differentiation of hypervirulent strains (027) from non-hypervirulent strains (176). However, ribotype 176 shares many similarities with 027, such as the presence of binary toxin genes and a single-base-pair deletion at nucleotide 117 in the gene that encodes a negative regulator of toxin production, in such a way that is actually difficult to differentiate both ribotypes in most methodologies. Nevertheless, severe disease outcomes and mortality are associated with ribotype 176 and in clinical epidemiology, hypervirulent strains should be notified. In our work, 027 strains were differentiated from other non-027 strains, such as 001, 002, 003, 012, 014, 017, 019, 020, 076, 106, 220, and 353 ribotypes. It is highly likely (98% of protein sequence coverage) that 30S ribosomal protein was assigned to both peaks, which binds directly to 16S ribosomal RNA. In addition, both peaks had excellent AUC values (0.96–0.99), sensitivity, specificity, PPV, and NPV, all of which indicate a good performance of a biomarker.

MRSA is one of the most studied pathogens of the ESCAPE group regarding MALDI-TOF analysis. A number of studies have used MALDI-TOF MS to discriminate clonal complexes of MRSA (CC5, CC8, CC22, CC30, and CC45). In the present study, the peak m/z 4,594 had a high specificity of 96.8% to detect MRSA isolates. The 50S ribosomal protein was most probably assigned to this peak. Different protein synthesis inhibitors, such as macrolides, ketolides, lincosamides, and B-type streptogramins, prevent the 50S particle from being formed in growing cells. To our knowledge, this is the first report of a potential biomarker peak for MRSA isolates. In addition, a m/z 2,415 peak was detected in the MRSA isolates (not shown). This peak was previously reported in MRSA isolates with the class A mec gene complex, though in our study the characterisation beyond the mecA gene detection was not performed.

Few reports have included the analysis of drug-resistant A. baumannii through MALDI-TOF MS, and although promising, their results were obtained using the carbapenem hydrolysis assay, not the direct approach. In our study, two potential biomarker peaks were detected for MDR A. baumannii: m/z 6,304 and m/z 6,332. Both peaks had high AUC values (0.99) and were present in blaOXA-58- and blaOXA-24-positive isolates, respectively. A NADH-quinone oxidoreductase subunit was likely assigned to both peaks. Type II NADH-quinone oxidoreductases of Gram-negative bacteria can be inhibited by polymyxin B and colistin. The performance of peak m/z 6,332 as a biomarker improved when comparing blaOXA-58-positive with blaOXA-24-negative and blaOXA-24 isolates. Our results indicate the discriminatory power of MALDI-TOF MS to differentiate blaOXA-58 from blaOXA-24 isolates, and this may be applied directly in clinical microbiology.

A signal loss of three peaks was detected for MDR P. aeruginosa isolates in comparison to non-MDR isolates. All three peaks had acceptable individual AUC values (0.81–0.84), and interestingly, the inclusion of all three peaks increased their potential as biomarker peaks for the detection of non-MDR isolates (sensitivity, 75%; specificity, 74.1%; PPV, 75%; and NPV, 74.1%). Two of the peaks were likely assigned to hypotheticals proteins, such as UPF0270 protein Pil01_4103 (m/z 2,726 peak) and UPF0391 membrane protein Patl_1732 (m/z 5,455 peak). In case the identified peaks were proteins, the signal loss could be due to a mutation introducing a premature stop codon or a frame-shift in a gene. As it is, we cannot assume they are indeed associated to proteins and other explanations must be considered. Growth conditions and sample preparation may also contribute to signal loss.
Therefore, loss of a signal does not alone give distinctive information about the strain genotype and cannot be used as a reliable biomarker peak29.

One potential biomarker for carbapenem and colistin resistance in *K. pneumoniae* was identified at peak *m/z* 6,100, which could be used to detect this phenotype, giving its NPV (91.7%). The use and the resistance to polymyxins have increased in carbapenemase-producing *Enterobacteriaceae*. A MALDI-TOF MS–based method that detects polymyxin resistance-related modifications to lipid A (i.e., phosphoethanolamine addition) was previously developed in *E. coli*. This methodology can be performed in 15 min and simultaneously discriminates between chromosome- and plasmid-encoded resistance32,33. This methodology has an advantage over our approach, both in specificity and processing time.

In addition, we did not assess the frequency of the colistin resistance-associated *mcr* gene in our *Enterobacteriaceae* isolates. Our study population contained the *bla*NDM* gene only in the carbapenem-resistant *K. pneumoniae* isolates. Because the *blaKPC* *K. pneumoniae* genotype has been previously identified by MALDI-TOF34, future studies should analyse a higher number of isolates and a wider diversity of drug resistance–associated genes of *K. pneumoniae*.

Previous studies have compared VRE mass spectra among *vanA* or *vanB* positive and negative groups, specifically on *E. faecium* isolates5,21. A number of peaks were detected among *vanA*-positive isolates35,36 as well as...
vanB-positive isolates. In our study, we detected 70 potential peaks with significant difference (p ≤ 0.05), but no final potential biomarkers were detected. Failure to identify lineage-specific biomarker peaks in Gram-positive pathogens such as *E. faecium* has been reported previously.

Changes in the distribution of peaks among strains suggest a peak shift, possibly corresponding to a slight protein weight modification due to an amino acid substitution after gene mutations. This shift can be used as a trustable biomarker to identify the strain genotype (either related to virulence or drug resistance), which was the approach used in this study. According to our approach, resistance or virulence biomarker peaks can be identified while performing routine microbiology analysis, and no additional assays nor prolonged incubation time is needed.

The primary limitation of this study was the lack of validation studies. Predictive biomarkers for MALDI-TOF should be first identified with a training set of data and afterward should be validated on an independent set of data. This validation set should include isolates with geographic, temporal, and clonal diversity to ensure the validity of the peak biomarkers. In addition, outstanding potential biomarker peaks detected in our study should be further analysed by additional methods to identify which proteins/peptide each peak corresponds to.

In conclusion, we identified several potential biomarker peaks suggestive of drug resistance in *S. aureus*, *A. baumannii*, *P. aeruginosa*, and *K. pneumoniae*, as well as hypervirulence in *C. difficile* using MALDI-TOF MS.

**Methods**

**Clinical isolates.** Selected ESCAPE clinical isolates obtained between January 2007 and December 2017 from the Hospital Civil de Guadalajara “Fray Antonio Alcalde” (Jalisco, Mexico) were included. The VITEK 2 system (Biomerieux, Craponne, France) was used according to the manufacturer's instructions, and either identification card for Gram-positive or negative was used for the identification of Gram-negative bacteria. GISAIDs were grown on plates containing Trypticase soy agar with 5% sheep blood and incubated overnight at 35 °C. A suspension of bacteria from pure cultures was performed in a 0.45% sodium chloride solution. This suspension was adjusted to a McFarland standard of 0.5 along with the ID card, introduced into the system, which automatically inoculated the card and incubated at 35 °C until the end of the test.

*A. baumannii* species identification was confirmed by polymerase chain reaction (PCR). Primers and PCR conditions were used as described previously. First, rapid DNA extraction was performed, in which three to five bacterial colonies were suspended in 100 µl of sterile distilled water and heated at 95 °C for 15 min. After centrifugation at 13,500 rpm for 5 min, the supernatant was collected. A pair of primers, forward (F): 5'-CATTATACGGGTAAATAGTG-3' and reverse (R): 5'-AGAGCAGTGCTGACCTGA-3', was used for the amplification of an internal fragment of the 16S-23S rRNA intergenic spacer region of *A. baumannii*. A second pair of primers, F5'-CCTGAACTTCGTGATAAC-3' and R5'-GTTTCTGGGTGCAAAACATTG-3', were used for the amplification of a highly conserved region of the recA gene of *Acinetobacter* genus. The reaction mixture contained 1X PCR buffer, 2 mM MgCl2, 0.2 mM of each dNTP, 200 nM of each primer, 1 U of AmpliTaq polymerase and 2 µl of DNA. PCR was initiated by denaturation for 5 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 30 s at 72 °C with final extension for 7 min at 72 °C and a holding step at 4 °C until analysis. Next, 5 µl of the PCR products were electrophoresed on a 2.0% agarose gel prepared with 0.5X TBE buffer for 1 h at 130 V. The gel was stained with 10 µM ethidium bromide, and the PCR product bands were visualized using a UV transilluminator. Expected PCR products for the 16S-23S rRNA intergenic spacer region and recA gene were 208 and 425 bp, respectively.

Selected clinical isolates included *E. faecium* (n = 64), *S. aureus* (n = 67), *C. difficile* (n = 93), *A. baumannii* (n = 68), *P. aeruginosa* (n = 55), and *K. pneumoniae* isolates (n = 69).

**Antimicrobial susceptibility and phenotypic tests.** Antimicrobial susceptibility was determined using the broth microdilution method according to the 2019 Clinical and Laboratory Standards Institute. Following strains were evaluated for resistance to the corresponding drugs: *E. faecium*, vancomycin; *S. aureus*, ceftoxitin and oxacillin; *A. baumannii*, imipenem and meropenem; *P. aeruginosa*, gentamicin, cefepime, ciprofloxacin, imipenem, and meropenem; *K. pneumoniae*, ertapenem, imipenem, meropenem and colistin. Strains were classified as MDR as described elsewhere, i.e. resistant to three or more antimicrobial classes.

**Genotyping.** *E. faecium* isolates were screened as previously described for the presence of vanA and vanB genes after rapid DNA extraction. Primers F5'-CATGAACTAGAATGAGCATTACTAAT-3' and R5'-CCCTTATTAAAGCCGTGATCAA-3' were used for the amplification of vanA gene (1,030 bp); and F5'-GTGACAAAACCGGAGGGCAGAGGA-3' and R5'-CCGGCCATCTCCTCTGAGAAAAAA-3', for the vanB gene (433 bp). PCR conditions were 10 min at 95 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at 58 °C and 30 s at 72 °C with final extension for 10 min at 72 °C.

The *mecA* gene (147 bp) was screened in all *S. aureus* isolates, using primers F5'-GTGAAAGTATACCAAGTGATT-3' and R5'-ATGGCGGATAGTTGGAAGGAT-3'. PCR conditions were 5 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 1 min at 50 °C and 2 min at 72 °C with final extension for 10 min at 72 °C.

The *blaVIM* gene was amplified in all *A. baumannii* isolates, using primers F5'-GATCCGATTGGAGAAACCAGA-3' and R5'-ATTTGCGGACTTCCTCACC-3'. PCR conditions were 3 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 1 min at 50 °C and 2 min at 72 °C with final extension for 10 min at 72 °C.

The *blaIMP* gene (2,183 bp) was amplified in all *S. aureus* isolates, using primers F5'-GGATTAGTGGGCATATACTC-3' and R5'-AGTTGGCGACTAACAG-3'. PCR conditions were 3 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 1 min at 50 °C and 2 min at 72 °C with final extension for 10 min at 72 °C.
184 bp)⁴⁹, blaVIM, blaIMP, blaNDM, and blaOXA-48 (F′-TTCCGAGGACGGAGCAAATCAG-3′ and R′-GATTGCGCATATGATATCTCGA-3′, 240 bp)⁴⁹. Conditions for multiplex PCR (blaOXA-23, blaOXA-24, blaOXA-51, and blaOXA-53) were 5 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 30 s at 72 °C, and 6 min at 72 °C. For blaNDM and blaVIM, PCR conditions were 5 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 60 s at 60 °C and 45 s at 72 °C, and 5 min at 72 °C. For blaIMP, PCR conditions were 5 min at 95 °C, 30 cycles of 30 s at 95 °C, 40 s at 52 °C and 50 s at 72 °C, and 5 min at 72 °C. For multiplex PCR blaOXA-48 and blaKPC, conditions were 9 min for 15 min, followed by 24 cycles at 95 °C for 1 min, 57 °C for 1 min and 72 °C for 1 min, and 72 °C for 30 min. Capillary electrophoresis was performed at the C. difficile Ribotyping Network Reference Laboratory at Leeds Teaching Hospitals Trust (Leeds, UK).

Mass spectra analysis. All selected strains were prepared for MALDI-TOF MS (Microflex LT system, Bruker Daltonics, Bremen, Germany) according to the manufacturer’s recommendation. For the direct colony method, an overnight culture grown on blood agar at 37 °C under aerobic conditions of each bacterial strain or from a culture grown on C. difficile agar (Neogen Corporation, MI) in anaerobic conditions for up to 24 h, were included. One colony of each bacteria was applied using sterile wooden toothpicks on a 96-spot stainless steel target plate (Bruker Daltonics, Bremen, Germany). After drying, 1 μL 70% formic acid was added and air-dried prior to adding 1 μL of a matrix solution, which was composed of a saturated solution of α-cyano-4-hydroxycinnamic acid 10 mg/mL in 50% acetonitrile (Sigma-Aldrich, Toluca, Mexico) and 2.5% trifluoroacetic acid (Sigma-Aldrich), which enables measurement of peptides and proteins from 0.7 to 20 kDa. Finally, spots were allowed to dry completely, and the target plate was introduced into the equipment. Samples were analysed by MALDI Biotyper 3.0 software for spectra profile match in the database.

All the isolates were classified according to manufacturer’s recommended score identification criteria. A score of 2.0 to 3.0 indicated reliable species-level identification, a score of 1.7 to 1.9 indicated reliable genus identification but questionable species-level identification, and a score of <1.7 indicated an unreliable identification.

Potential marker peaks. The MS spectrum of each spot was analysed by the MALDI Biotyper V3.1.66 with the most updated spectra library, V.7.0 (7,311 spectra). High-quality spectra were captured using the flexControl V.3.4 software (Bruker Daltonics). These spectra were then imported into the ClinProTools V3.0 software (Bruker Daltonics) for recognition of mass spectra patterns between groups and preprocessed using the default parameters. Spectra were smoothed with 10 cycles of the Savitzky/Golay algorithm for 10 cycles with a width of 2 m/z. Baseline subtraction was performed with the Top Hat algorithm. Peak picking was performed on the average spectrum from each group, with a signal to noise threshold of 5. Peak selection was performed using the P-value T-test/ANOVA sort mode.

Group selection was based on the presence or absence of drug-resistance genotypes, except for C. difficile in which hypervirulent ribotypes determination was considered.

Statistical analyses. The selection of potential biomarker peaks of drug-resistant phenotypes was performed using statistical analyses within the algorithms included in the ClinProTools software. Classification models were generated using the genetic algorithm (GA), supervised neural network (SNN), and quick classifier (QC) algorithms. The parameters were maximum of 5 peaks, maximum of 50 generations, k-nearest neighbor classification with 3 neighbors for GA; 1–25 peaks automatically detected for SNN, and QC (automatic peak detection 1–25 peaks, p-value T-test/ANOVA). The recognition capability and cross validation values were calculated to demonstrate the reliability and accuracy of the model.

The Anderson-Darling test was used to determine the distribution of the population (≤0.05: not normally distributed, >0.05, normally distributed). According to the results, either the t-test (normally distributed data) or the Wilcoxon test (not normally distributed data) were used to confirm significant differences between two classes (on the presence or absence of drug-resistance or virulence phenotype/genotypes). As the p-value provides a measure of the strength of an association, the lower it is, the better the respective peak signal can be used to discriminate the groups. If the p-value for the t-test or the Wilcoxon test was ≤0.05, the peak was confirmed to be significantly different.

Peaks that obtained statistical significance were selected for further analysis, which included evaluating the peak area/intensity of the spectra and the coefficient of variation of each peak; this analysis allowed the selection of potential peaks. ClinProTools also calculates a Receiver Operating Characteristic (ROC) curve for each peak. The ROC curve provides an evaluation of the discrimination quality of a peak. In addition, the area under the curve (AUC) was determined and only peaks with AUC ≥ 0.80 values were selected. The principal component analysis (PCA) was used to evaluate the distinguishing capability between drug-resistance (or virulence) phenotype/genotype profiles in the isolates based on their peptide mass fingerprints.

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for the potential biomarker peaks.

Peptide/protein identity assignment. MS spectra were first processed using flexAnalysis software (Bruker Daltonics) in which top-hat baseline subtraction and spectra smoothing were performed. The spectra were then exported to the Biotools software (Bruker Daltonics), in which protein/peptide identification was performed by latter submitting the data to the Mascot Server (Matrix Science, Boston, USA) and run against the SwissProt database for peptide/protein assignment.
50. Bidet, P., Barbut, F., Lalande, V., Burghoffer, B. & Petit, J. C. Development of a new PCR-ribotyping method for Clostridium difficile.

49. Swayne, R. L.

48. Baroud, M.

47. Ellington, M. J., Kistler, J., Livermore, D. M. & Woodford, N. Multiplex PCR for rapid detection of genes encoding acquired metallo-β-lactamases. J Antimicrob Chemother 43, 309–313, https://doi.org/10.1016/j.jantimicag.2010.07.008 (2010).

46. Nordmann, P. & Poirel, L. Emerging carbapenemases in Gram-negative aerobes. J Antimicrob Chemother 65, 2211–2219, https://doi.org/10.1093/jac/dkt390 (2010).

45. Woodford, N.

44. Zhang, K., McClure, J. A., Elsayed, S., Louie, T. & Conly, J. M. Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome mec types I to V in methicillin-resistant Staphylococcus aureus. J Clin Microbiol 43, 5026–5033, https://doi.org/10.1128/JCM.43.10.5026-5033.2005 (2005).

43. Bocanegra-Ibarias, P. et al. Phenotypic and genotypic characterization of vancomycin-resistant Enterococcus faecium clinical isolates from two hospitals in Mexico: First detection of VanB phenotype-vanA genotype. Enferm Infecc Microbiol Clin 34, 415–421, https://doi.org/10.1016/j.eimc.2015.09.011 (2016).

42. Magiorakos, A. P. et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect 18, 268–281, https://doi.org/10.1111/j.1469-0691.2011.03570.x (2012).

41. Clinical and Laboratory Standards Institute. Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Fourth Informational Supplement M100-S24 (2014).

40. Bocanegra-Ibarias, P. et al. Genetic characterisation of drug resistance and clonal dynamics of Acinetobacter baumannii in a hospital setting in Mexico. Int J Antimicrob Agents 45, 309–313, https://doi.org/10.1016/j.ijantimicag.2014.10.022 (2015).

39. Bocanegra-Ibarias, P.

38. 38. 35–38, https://doi.org/10.1016/j.ijantimicag.2011.03.010 (2011).

37. Griffin, P. M. et al. Use of matrix-assisted laser desorption ionization-time of flight mass spectrometry to identify vancomycin-resistant enterococci and investigate the epidemiology of an outbreak. J Clin Microbiol 50, 2918–2931, https://doi.org/10.1128/JCM.01000-12 (2012).

36. Wang, L. J., Lu, X. X., Wu, W., Sui, W. J. & Zhang, G. Application of matrix-assisted laser desorption ionization time-of-flight mass spectrometry in the screening of vanA-positive Enterococcus faecium. Eur J Mass Spectrom (Chichester) 20, 461–465, https://doi.org/10.1255/ejms.1298 (2014).

35. Insufficient discriminatory power of MALDI-TOF mass spectrometry for typing of Enterococcus faecium and Staphylococcus aureus isolates. J Microbiol Methods 100, 58–69, https://doi.org/10.1016/j.mimet.2014.02.015 (2014).

34. Bocanegra-Ibarias, P. et al. Genetic characterisation of drug resistance and clonal dynamics of Acinetobacter baumannii in a hospital setting in Mexico. Int J Antimicrob Agents 45, 309–313, https://doi.org/10.1016/j.ijantimicag.2014.10.022 (2015).

33. Swayne, R. L.

32. underhttps://doi.org/10.1016/j.ijantimicag.2012.05.016 (2012).

31. Ellington, M. J., Kistler, J., Livermore, D. M. & Woodford, N. Multiplex PCR for rapid detection of genes encoding acquired metallo-β-lactamases. J Antimicrob Chemother 59, 321–322, https://doi.org/10.1093/jac/dkl481 (2007).

30. Baroud, M. et al. Underlying mechanisms of carbapenem resistance in extended-spectrum β-lactamase-producing Klebsiella pneumoniae and Escherichia coli isolates at a tertiary care centre in Lebanon: role of OXA-48 and NDM-1 carbapenemases. J Antimicrob Agents 41, 75–79, https://doi.org/10.1016/j.jantimicag.2012.08.010 (2013).

29. Wayne, R. L. et al. for rapid detection of genes encoding five types of non-metallo-β-lactamase-producing Klebsiella pneumoniae in Enterobacteriaceae. Int J Antimicrob Agents 38, 35–38, https://doi.org/10.1016/j.jantimicag.2011.03.010 (2011).

28. Bitde, P., Barbut, F., Lalande, V., Burghoffer, B. & Petit, J. C. Development of a new PCR-ribotyping method for Clostridium difficile based on ribosomal RNA gene sequencing. FEMS microbiology letters 175, 261–266 (1999).

27. Swayne, R. L.

26. Swayne, R. L.

25. Underhttps://doi.org/10.1016/j.ijantimicag.2012.05.016 (2012).

24. Swayne, R. L.

23. 10.1016/j.ijantimicag.2012.05.016 (2012).

22. Swayne, R. L.

21. Swayne, R. L.

20. Swayne, R. L.

19. Swayne, R. L.

18. Swayne, R. L.

17. Swayne, R. L.

16. Swayne, R. L.

15. Swayne, R. L.

14. Swayne, R. L.

13. Swayne, R. L.

12. Swayne, R. L.

11. Swayne, R. L.

10. Swayne, R. L.

9. Swayne, R. L.

8. Swayne, R. L.

7. Swayne, R. L.

6. Swayne, R. L.

5. Swayne, R. L.

4. Swayne, R. L.

3. Swayne, R. L.

2. Swayne, R. L.

1. Swayne, R. L.